A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

http://wrap.warwick.ac.uk/116336

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications
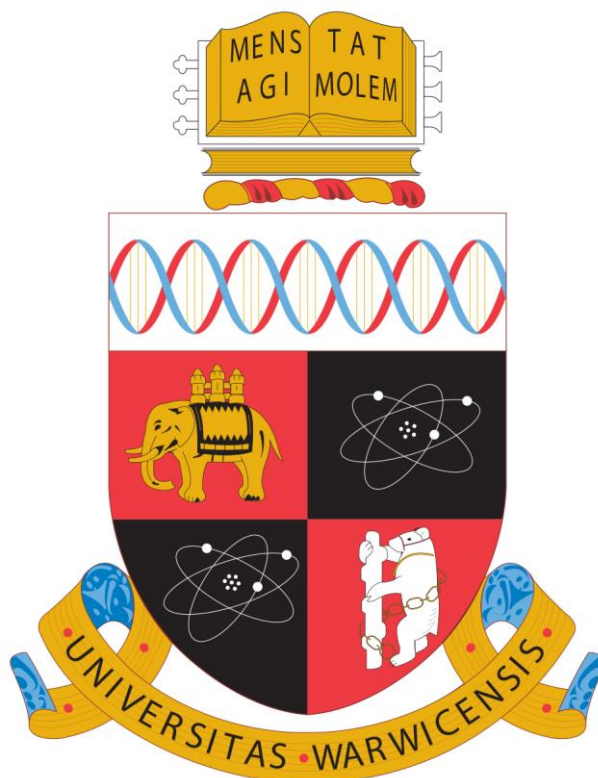
Advanced New Technologies for Protein Analysis by Mass Spectrometry

Pui Yiu Lam

A thesis submitted for the degree of

Doctor of Philosophy

Department of Chemistry

University of Warwick

July 2018

Table of Contents

Table of Contents

Table of Contents

Table of Contents

Pui Yiu (Yuko) Lam

Email: p.y.lam@warwick.ac.uk

Mobile: +44 (0)7413 669964

Education

Ph.D. Chemistry

University of Warwick   2014.10 – 2018.09

Thesis title: Advanced new technologies for protein analysis by mass spectrometry

Supervisor: Prof. Peter O'Connor

M.Sc. Analytical Science

University of Warwick   2010.10 – 2011.09

Graduate with Distinction

Thesis title: Use of high resolution mass spectrometry for analysis of polymeric excipients in drug delivery formulations

Supervisor: Prof. Peter O'Connor

B.Sc. Science (Chemistry & Finance)

University of Hong Kong   2007.09 – 2010.06

Graduate with 2.1 honour

Publications

Accept papers

- van Agthoven, M. A.; Lynch, A. M.; Morgan, T. E.; Wootton, C. A.; **Lam, Y. P. Y.**; Chiron, L.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Can Two-Dimensional IR-ECD Mass Spectrometry Improve Peptide de Novo Sequencing? *Anal. Chem.* **2018**, 90 (5), 3496-3504.
- Floris, F.; van Agthoven, M. A.; Chiron, L.; Wootton, C. A.; **Lam, P. Y. Y.**; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Bottom-Up Two-Dimensional Electron-Capture Dissociation Mass Spectrometry of Calmodulin. *J. Am. Soc. Mass. Spectrom.* **2018**, 29 (1), 207-210.

- Zhang, P.; Chiu, C. K.; Huang, H.; **<u>Lam, Y. P.</u>**; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; Sadler, P. J., Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells. *Angew. Chem. Int. Ed.* **2017**, 56 (47), 14898-14902.
- Wootton, C. A.; **<u>Lam, Y. P.</u>**; Willetts, M.; van Agthoven, M. A.; Barrow, M. P.; Sadler, P. J.; Peter, B., Automatic assignment of metal-containing peptides in proteomic LC-MS and MS/MS data sets. *Analyst* **2017**, 142 (11), 2029-2037.
- Wei, J.; Antzutkin, O. N.; Filippov, A. V.; Iuga, D.; **<u>Lam, P. Y.</u>**; Barrow, M. P.; Dupree, R.; Brown, S. P.; O'Connor, P. B., Amyloid Hydrogen Bonding Polymorphism Evaluated by 15N {17O} REAPDOR Solid-State NMR and Ultra-High Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Biochem.* **2016**, 55 (14), 2065-2068.
- Simon, H.; van Agthoven, M.; **<u>Lam, P.</u>**; Floris, F.; Chiron, L.; Delsuc, M.-A.; Rolando, C.; Barrow, M.; O'Connor, P., Uncoiling collagen: a multidimensional mass spectrometry study. *Analyst* **2016**, 141 (1), 157-165.
- Floris, F.; van Agthoven, M.; Chiron, L.; Soulby, A. J.; Wootton, C. A.; **<u>Lam, Y. P.</u>**; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., 2D FT-ICR MS of calmodulin: a top-down and bottom-up approach. J. *Am. Soc. Mass. Spectrom.* **2016**, 27 (9), 1531-1538.
- Perez Hurtado, P.; **<u>Lam, P. Y.</u>**; Kilgour, D.; Bristow, A.; McBride, E.; O'Connor, P. B., Use of high resolution mass spectrometry for analysis of polymeric excipients in drug delivery formulations. *Anal. Chem.* **2012**, 84 (20), 8579-8586.

<u>Submitted paper</u>
- "Exploring the Aggregation Mechanisms of Amyloid Protofibrils Using Electron Capture Dissociation Tandem Mass Spectrometry" by **<u>Lam, Y.P.Y.</u>**; Wootton, C.A.; Hands-Portman, I.; Wei, J; Chiu, C.K.C.; I. Romero-Canelon Lermyte, F.; Barrow, M. P.; and O'Connor, P. B. Submitted to Analytical Chemistry, 2018.
- "Does Deamidation of Islet Amyloid Polypeptide Accelerate Amyloid Fibril Formation?" by **<u>Lam, Y.P.Y.</u>**; Wootton, C.A.; Hands-Portman, I.; Wei, J; Chiu, C.K.C.; I. Romero-Canelon Lermyte, F.; Barrow, M. P.; and O'Connor, P. B. Submitted to Chemical Communications, 2018.
- "The Inhibition Mechanisms of Human Islet Amyloid Polypeptide Fibrillisation" by **<u>Lam, Y.P.Y.</u>**; Chiu, C.K.C.; Wootton, C.A.; Song, J.I.; Li, M.; Hands-Portman, I.; Barrow, M. P.; and O'Connor, P. B. Submitted to Chemical Science, 2018.

- "Two-dimensional mass spectrometry: new perspectives for tandem mass spectrometry" by van Agthoven, M.A.; **Lam, Y.P.Y.**; O'Connor, P.B.; Rolando, C.; Delsuc; M.A. Submitted to European Biophysics Journal, 2018.
- "Discovery of a natural product-based Factor Xa inhibitor as an anticoagulant agent" Chen, P.; Xhang, D.W.; Li, M.; Wu, Q.; **Lam, Y.P.Y.**; Guo, Y.; Chen, C.; Hu, R.; Bai, N.; Malhotra, S.; Li, W.; O'Connor, P.B., Fu, H. Submitted to Journal of Medicinal Chemistry, 2018.

Manuscripts in preparation
- "The Next Dimension in Proteomics" by **Lam, Y.P.Y.**; Wootton, C.A.; Chiu, C.K.C.; Morgan, T.E.; Gavard, R.; van Agthoven, M.A.; Li, M.; Barrow, M. P.; and O'Connor, P. B.

Awards

Academic Awards
- **Educational Grant**
  From The Great Britain – China Educational Trust (GBCET) for 2017/18
- **Research Development Fund (RDF) Strategic Award**
  From University of Warwick for 2016/17
- **Foundation Main Grant**
  From Funds for Women Graduates organisation for 2016/17
- **Educational Trust**
  From Sir Richard Stapley Education Trust for 2016/17
- **Travel Grant**
  From Royal Society of Chemistry (RSC) Analytical Trust fund and International Mass Spectrometry Foundation (IMSF) for the 2018 International Mass Spectrometry Conference (IMSC)
- **Travel Grant**
  From Royal Society of Chemistry (RSC) for the 2017 Analytical Research Forum (ARF)
- **Travel Grant**
  From American Society for Mass Spectrometry (ASMS) and British Mass Spectrometry Society (BMSS) for the 65th American Society for Mass Spectrometry (ASMS) conference
- **Travel Grant**
  From British Mass Spectrometry Society (BMSS) for the 9th Isolated Biomolecules and biomolecular Interaction (IBBI) conference

Presentation Awards

- **Flash Presentation Runner Up Prize**
  At the 2017 Analytical Research Forum (AFR)
- **Best Talk Award**
  At the 14th East Midlands Proteomics Workshop (EMPW)
- **Bordoli Prize for Best Young Researcher Poster Presentation**
  At the 36th British Mass Spectrometry Society (BMSS) conference

Conference Presentations

Oral Presentations

- **22nd International Mass Spectrometry Conference (IMSC 2018) –**
  Florence, Italy**;** Presented in the Non-covalent Interactions session
- **15th Uppsala Conference on Electron Capture and Transfer Dissociation Mass Spectrometry** – Leeds, UK
- **65th American Society for Mass Spectrometry Annual Conference (ASMS 2017) –** Indianapolis, USA; Presented in the Native MS in Structural Biology session
- **12th European Fourier Transform Mass Spectrometry Conference (EFTMS 2016) –** Matera, Italy; Presented in the Protein Analysis session
- **14th East Midlands Proteomics Workshop (EMPW 2015) –** Coventry, UK

Poster Presentations

- **16th East Midlands Proteomics Workshop (EMPW 2017) –** Nottingham, UK
- **38th British Mass Spectrometry Society Annual Conference (BMSS 2017) –** Manchester, UK
- **2017 Analytical Research Forum (ARF 2017) –** London, UK; Poster Presentation plus 5 minutes flash talk
- **15th East Midlands Proteomics Workshop (EMPW 2016) –** Nottingham, UK
- **37th British Mass Spectrometry Society Annual Conference (BMSS 2016) –** Eastbourne, UK
- **9th Isolated Biomolecules and biomolecular Interaction Conference (IBBI 2016) –** Oxford, UK
- **36th British Mass Spectrometry Society Annual Conference (BMSS 2015) –** Birmingham, UK

Acknowledgements

It is a great pleasure to take this opportunity to thank all the people who helped in my 4-year PhD degree at the University of Warwick. Foremost, I would like to thank my supervisor Prof. Peter O'Connor for his guidance and support during my PhD. The knowledge and the enthusiasm from Prof. O'Connor in the scientific research are the strongest guiding forces in these 4 years of study which allow me to achieve higher during my PhD. Prof. O'Connor is a great academic who always reminds me to pursue the fact of science not the expectation; during project discussion, he always keeps an open mind and judges the project solely based on the data obtained. Prof. O'Connor also supports me to attend lots of national as well as international conferences during my PhD which helps to broaden my horizons in different research areas as well as enhance my networking in the science field. Aside from academics, Prof. O'Connor also provides lots of advice and assistance when I faced difficulties in life. I am very honoured to be his student for my masters as well as PhD degrees, and I am looking forward to continuing to work with him in the future.

I would like to acknowledge all the past and present members in Peter O'Connor group as well as Mark Barrow's group. I have learnt not only the science, but also the interesting cultures, languages, as well as viewpoints from different group members. I really enjoy working in such a pleasant and collaborative environment. I would like to specially thank Dr. Christopher A. Wootton, Mr. Cookson K.C. Chiu, and Miss Meng Li for the help and support during my PhD. I would also like to thank all the co-authors and colleagues who have been involved in the projects that presented in this thesis.

I would like to take this opportunity to thank my parents Percy Lam and Kewinda Wong for their endless support and love in these years, so that I can

continue to pursue my career in science. I would also like to thank my sister Hitomi Lam as well as my partner Jonathan Wong for their support and advice during my PhD.

At last, I would like to thank the following organisations: University of Warwick, Funds for Women Graduates, The Great Britain-China Educational Trust, and The Sir Richard Stapley Educational Trust for providing funding to support my PhD degree.

Declaration

I hereby declare that except where specifically references/stated are made to other sources, the thesis entitled "Advanced New Technologies for Protein Analysis by Mass Spectrometry" is the original work of the named Author. It has been composed by myself and co-authors where stated and has not been submitted in whole or in part for any other degree, diploma, or qualification.

Pui Yiu Lam

July 2018

Thesis Summary

The work presented herein applies advanced techniques in an ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) to study one of the most amyloidogenic amyloid proteins − human islet amyloid polypeptide (hIAPP) and we also demonstrate alternative analytical techniques for proteomics.

Exploring the aggregation and inhibition pathways of amyloid proteins is critical for future therapeutic development. Previous studies demonstrated the early, soluble oligomers of amyloid proteins are more toxic to neuro cells than mature amyloid fibrils due to their ability to penetrate and destroy neuro cell membranes. Current analytical techniques for amyloid protein studies, including nuclear magnetic resonance (NMR) spectroscopy, fluorescence spectrometry, transmission electron microscopy (TEM), and atomic force microscopy (AFM), are difficult to perform as they require molecular isolation and focus on studying individual target early oligomers from a mixtures of oligomers. Previous literature, however, has demonstrated the ability of MS in observing as well as isolating a specific oligomer from an *in vitro* amyloid protein sample. Thus, MS has been applied together with electron-based dissociation technique in this thesis to determine the aggregation as well as inhibition pathways of hIAPP.

The results herein demonstrate the region around Gly-33 and Ser-34 of hIAPP is critical for hIAPP aggregation since it is the minimum interaction region between the two hIAPP units which was shown by the electron-based dissociation spectra of the dimer as well as the trimer of hIAPP. hIAPP has also shown to be deamidated rapidly within a month at Asn-21, Asn-22, and Asn-35 residues, regarding there are seven potential deamidation sites within the sequence. The deamidated hIAPP species with iso-aspartic acid residue products at the

deamidation sites shows a significant change in the amyloid fibril morphology which may help to explain the previously observed acceleration of aggregation rate by inducing 5% of deamidated hIAPP into non-deamidated hIAPP solution.

The methodologies applied in studying the aggregation pathway of hIAPP are further used to explore the inhibition pathways between hIAPP and potential inhibitors. From the results, two different inhibition pathways are proposed to effectively prevent the formation of hIAPP fibrils. Inhibitors interacting specifically at the critical aggregation region around Gly-33 and Ser-34, are shown to prevent the formation of hIAPP fibrils by forming hetero-complexes (a complex composes of hIAPP and inhibitor). The alternative inhibition pathway requires an inhibitor which interacts non-specifically with hIAPP and triggers the formation of non-toxic amorphous aggregates faster than the generation of normal, toxic hIAPP oligomers. These two proposed inhibition pathways show an effective inhibition effect on formation of hIAPP fibrils, however, only the non-specifically interaction pathway was effective in preventing the aggregation of the deamidated from hIAPP since the protein sequence as well as structure are distorted in the deamidated hIAPP which significantly affects the specific interaction between the deamidated hIAPP and inhibitors.

The *in vitro* studies propose the potential aggregation and inhibition pathways for hIAPP; it is critical, but unfortunately very challenging, to study these molecular interactions *in vivo* due to the solvents and column chemistry required by the current liquid chromatography (LC) approach which is necessary for complex mixture studies. In order to study the *in vivo* molecular interactions, an alternative analytical technique, therefore, is required.

Herein, two dimensional mass spectrometry (2DMS) is applied to study proteomic samples ranging from standard protein digests to whole cell lysate

digest without online LC separation. Various sample preparation and 2DMS acquisition methods have been applied to optimise the proteomic outputs from 2DMS. From the results, 2DMS shows a similar capability in studying proteomics compared to the standard LC tandem MS (MS/MS) techniques. The peptides assigned in 2DMS experiments are more hydrophilic, basic, and short, which are complementary to those assigned in the LC MS/MS technique, thus more proteins can be covered by combining the results obtained from 2DMS and LC MS/MS, in which a deeper proteome coverage is achieved. 2DMS is shown to be a viable alternative analytical tool for proteomics, in which is also potential for the *in vivo* study of molecular interactions in the coming future.

The final section of this thesis provides a conclusion on all the works that have been demonstrated herein and discusses the future research which we needed to improve/enhance the current knowledge on amyloid protein studies using MS as well as 2DMS for proteomics.

Abbreviations

2D-LC	Two dimensional liquid chromatography
2DMS	Two dimensional mass spectrometry
2D-PAGE	Two dimensional polyacrylamide gels
3-APS	3-amino-1-propane sulfonic acid
5CP	5 combined proteins
ABC	Ammonium bicarbonate
ACN	Acetonitrile
AD	Alzheimer's disease
AFM	Atomic force microscopy
Asn	Asparagine
$A\beta_{1-40}$	Amyloid beta 1-40
BISA	1H-Benzimidazole-2-sulfonic acid
BSA	Bovine serum albumin
CAD	Collisionally activated dissociation
CCS	Collision-cross section
CD	Circular dichroism
CE	Capillary electrophoresis
CEM	Chain ejection model
CHCA	α-Cyano-4-hydroxycinnamic acid
CI	Chemical ionisation
CLZ	Chicken lysozyme
CO_2	Carbon dioxide
CRM	Charge residue model
Cys	Cysteine
DC	Direct current
DHB	2,5-Dihydroxybenzoic acid
DTT	Dithiothreitol
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EDTA	Ethylenediaminetetraacetate
EGCG	(-)-Epigallocatechin 3-gallate
EI	Electron ionisation
EID	Electron induced dissociation
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FA	Formic acid
FDR	False discovery rate
FTICR	Fourier transform ion cyclotron resonance

Abbreviations

FTMS	Fourier transform mass spectrometry
FWHM	Full width half maximum
GC	Gas chromatography
Gln	Glutamine
HD	Huntington's diseases
HDX	Hydrogen-deuterium exchange
HHA	Human haemoglobin
hIAPP	Human islet amyloid polypeptide
HILIC	Hydrophilic interaction liquid chromatography
HLZ	Human lysozyme
HPLC	High performance liquid chromatography
HSA	Human serum albumin
i.d.	Internal diameter
IAA	Iodoacetamide
IEM	Ion ejection model
IM-MS	Ion mobility-mass spectrometry
IR-ECD	Infrared electron capture dissociation
IRMPD	Infrared multiple photon dissociation
kV	Kilovolts
LC	Liquid chromatography
M	Mega-word
m/z	Mass-to-charge
MALDI	Matrix assisted laser desorption
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
nESI	Nano electrospray ionisation
nLC	nano-LC
P_1	Excitation pulse applied in two dimensional mass spectrometry
P_2	Excitation pulse applied in two dimensional mass spectrometry
PD	Parkinson's disease
ppm	Parts-per-million
PTM	Post translation modification
RF	Radio frequency
rIAPP	Rat islet amyloid polypeptide
RP	Reversed-phase
S/N	Signal-to-noise
SA	3,5-Dimethoxy-4-hydroxycinnamic acid
SAX	Strong anion exchange chromatography

Abbreviations

SCX	Strong cation exchange chromatography
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodcyl sulfate polyacrylamide
SORI-CID	Sustained off-resonance irradiation collision-induced dissociation
SPE	Solid phase extraction
SPIKE	Spectrometry Processing Innovative Kernel
SSNMR	Solid-state nuclear magnetic resonance
SWIFT	Stored waveform inverse Fourier transform
T	Tesla
t_1	Delayed period applied in two dimensional mass spectrometry
T2D	Type II diabetes (for Chapter 1 -3)
TEM	Transmission electron microscopy
ThT	Thioflavin T
TOF	Time-of-flight
UVPD	Ultraviolet photodissociation
μM	Micro-molar

Tables and Illustrated Material

List of Figures and Tables

Chapter 1 Introduction

Tables and Illustrated Material

Tables and Illustrated Material

Tables and Illustrated Material

mutant $(isoD)_3$hIAPP monomer unit and a 3-APS molecule.

Tables and Illustrated Material

Chapter 4 The Next Dimension in Proteomics

Figures and Tables in Supplementary Information

Tables and Illustrated Material

Tables and Illustrated Material

Chapter 5 Conclusion and Future Works

- Summarised figure for Chapter 2

- Summarised figure for Chapter 3

- Summarised figure for Chapter 4

Chapter 1 Introduction

Mass spectrometry (MS) is an important technique to study biological molecules because of its accuracy, sensitivity, speed, and the capability of simultaneous analysis.[1, 2]   MS is commonly applied to provide sequence and modification information,[3, 4] as well as structure and quantification analysis of biological molecules.[5, 6]

This thesis focuses on the applications of advanced technologies in mass spectrometry for protein analysis, which includes the application of tandem MS to explore the aggregation and inhibition mechanisms of amyloid proteins, as well as development of a new analytical technique which is suitable for the study of fragile amyloid aggregates in a complex mixture.

This chapter contains three sections, including introduction of mass spectrometry (section 1.1), applications of MS of amyloid proteins, deamidation, as well as proteomics studies (section 1.2), and an overview of the thesis (section 1.3).

1.1 Introduction of Mass Spectrometry

MS is a powerful analytical instrument which provides molecular mass information of gas state molecules by measuring the mass-to-charge (m/z) ratios of ions.   The first MS, which was also named a "parabola spectrograph", was invented in 1912 by J.J. Thomson and the mass spectra of O_2, N_2, CO, CO_2, and $COCl_2$ were obtained.[7]   MS has been rapidly developed in the last century.[8]   MS is currently one of the most common analytical instruments in an array of studies, including metabolic,[9] proteomics,[3] genomics,[10] and petrolomics.[11]   In general, MS contains three major components which are the ionisation source, the mass

analyser, and a detector.   The key ionisation techniques, mass analysers, and fragmentation methods for these studies will be discussed in this section.

1.1.1 Key Ionisation Techniques in Mass Spectrometry

In a mass spectrometer, only charged gas state molecules can be detected, thus it is important to ionise analytes prior to MS detection.   In general, ionisation techniques can be classified into two categories which are hard and soft ionisation. Hard ionisation indicates the ionisation process is energetic, fragments and molecular ions from the analytes are obtained simultaneously.[8]   Fragments generated from hard ionisation are useful for elucidating the structures of analytes; these ionisation methods; however, are not suitable for a mixture of analytes as complex spectra with multiple fragments from each analyte are generated which is challenging for data analysis.   On the other hand, molecular ions are almost exclusively generated in soft ionisation which leads to less complicated spectra. This method is more suitable for the studies of mixtures of analytes, particularly when it couples with isolation and fragmentation (section 1.1.3) techniques in MS.

1.1.1.1 Electron Ionisation

Electron ionisation (EI) is one of the classic ionisation techniques for MS and was first introduced by Dempster in 1918.[12]   In general, an EI source contains a heated filament, an anode, an inlet for gaseous samples, and an ionisation chamber (see Figure 1.1).   In EI, electrons are generated from a heated filament, accelerated to ~ 70 eV, and moved towards the anode in an ionisation chamber.[8] Vaporised molecules are introduced from the inlet which is orthogonal to the heated filament, then collided with the electron beam to remove an electron, and

result in the formation of radical cations ($M^{+\bullet}$).   The overall EI process can be represented by the following equation:

$$M + e^- \rightarrow M^{+\bullet} + 2e^-$$ [Eq.1]

M represents the vaporised molecule, e^- represents the electron, and $\bullet$ represents the radical.
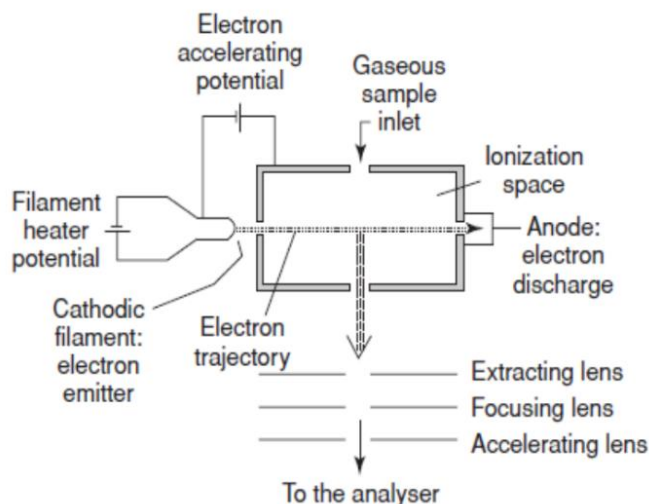
Figure 1.1.   Schematic representation of an EI source (Reproduced from De Hoffmann *et al.* 2007).[8]

The energetic radical cations generated from EI are often not stable which result in further dissociation and generate fragment ions, thus, EI is a hard ionisation technique.   The extensive fragment ions generated during ionisation provide extra information for structure elucidation of the analytes (see Figure 1.2). EI is commonly applied to separated, volatile, small molecules;[13] the complex spectra generated from EI, however, limited application to small molecules (~ 600 Da).   Gas chromatography (GC) is the most common separation techniques coupled with EI for clinical, environmental, food/drug analysis, and other applications.[14]
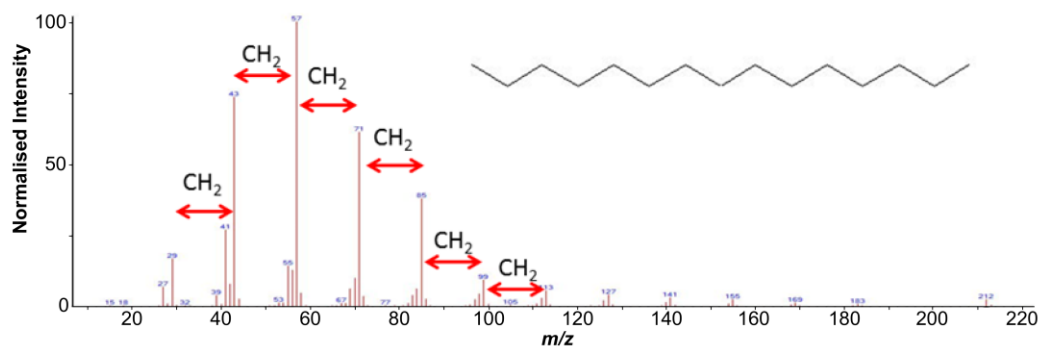
Figure 1.2. EI-MS spectrum of pentadecane ($C_{15}H_{32}$) (Reproduced from the tutorial material provided by SCION Instruments, West Lothian, United Kingdom).

1.1.1.2 Chemical Ionisation

Chemical ionisation (CI) was first introduced by Talrose in 1952 and the ionisation method was further developed by Munson and Field in 1966.[15, 16] Proton transfer is the most common ionisation reaction applied in CI. Reagent gas molecules, i.e. methane, isobutane, and ammonia etc., are first ionised using EI.[8] The protons generated in the EI process are then transferred to the analyte molecules if the molecules contain a higher proton affinity than the applied regent gas. Taking methane (CH_4) as an example, the overall CI process can be represented by the following equations:

Step 1: EI of reagent gas – methane (CH_4)

$$CH_4 + e^- \rightarrow CH_4^{+\bullet} + 2e^- \qquad\qquad [Eq.2]$$

$$CH_4^{+\bullet} + CH_4 \rightarrow CH_5^+ + CH_3^{\bullet} \qquad\qquad [Eq.3]$$

Step 2: Proton transfer

$$M + CH_5^+ \rightarrow MH^+ + CH_4 \qquad\qquad [Eq.4]$$

M represents the vaporised analyte molecule, e^- represents the electron, and $\bullet$ represents the radical.

Compared to EI, CI is a milder ionisation technique which results in less

extensive fragmentation during ionisation.[8]   The molecular ions (MH$^+$) generated by CI are even-electron molecules which are usually more stable than the odd-electron radical cations (M$^{+\bullet}$) produced by EI.   Thus, the extent of fragmentation generated in CI is less than with the EI process, which results in a relative higher intensity of molecular ion in the CI spectrum.   CI also requires volatile and less complex samples, just like EI.[15]

1.1.1.3 Matrix Assisted Laser Desorption

Matrix assisted laser desorption (MALDI) was first introduced by Karas and Hillenkamp in 1985.[17-18]   It was then applied to a broad range of non-volatile and thermally labile samples, including polymers, glycoproteins, and large inorganic molecules. [17, 19]   In 1988, Tanaka *et al.* showed that protein and polymer with *m/z* over 100,000 could be analysed by MALDI MS,[20] who was then awarded with the Nobel prize of Chemistry.   Since then, MALDI was further applied in large molecule analysis, such as DNA and proteins.[21, 22]

Currently, the mechanism of MALDI, especially the origin of ions generated in MALDI, is still under debate.[23-25]   In general, desorption and ionisation are the two important steps to generate a MALDI spectrum.   Analytes are first mixed with matrix solution which contains small organic molecules.   2,5-dihydroxybenzoic acid (DHB), 3,5-dimethoxy-4-hydroxycinnamic acid (SA), and α-cyano-4-hydroxycinnamic acid (CHCA) are the examples of common matrices for MALDI for protein and peptide analysis (see Figure 1.3).[26, 27]   The common features of these matrices are composed of an aromatic ring for absorbing laser energy and a functional group for providing a proton for analytes.   When samples (mixed of analytes and matrix) are irradiated with laser, matrix molecules absorb a large amount of energy which result in an ablation of a portion of matrix

crystals.[28, 29]   Neutral analyte molecules are vaporised together with the expanding matrix plume, protons are then transferred from the matrix to the analyte molecules, and result in the formation of ionized molecules (see Figure 1.4).
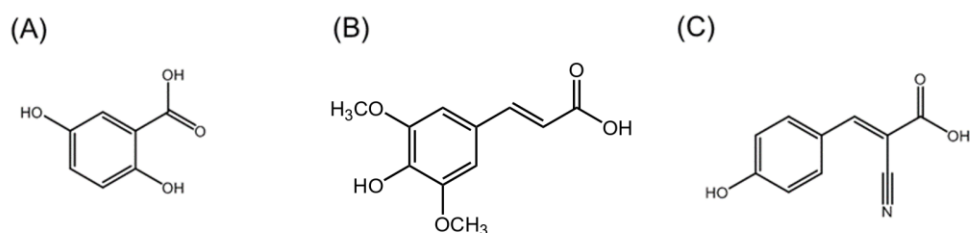
Figure 1.3.   Common matrices (A) 2,5-dihydroxybenzoic acid (DHB), (B) 3,5-dimethoxy-4-hydroxycinnamic acid (SA), and (C) α-cyano-4-hydroxycinnamic acid (CHCA) to use in protein and peptide analysis with MALDI source.
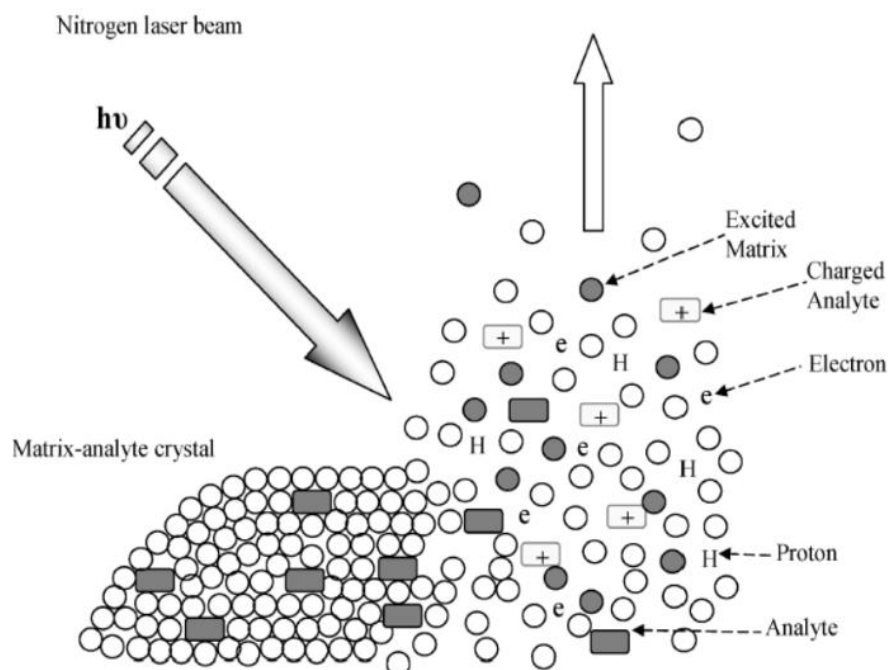
Figure 1.4.   The principle of MALDI (Reproduced from EI-Aneed *et al.* 2009).[29]

MALDI is a soft ionisation technique which does not induce fragmentation during ionisation, but excess irradiation from laser energy can still lead to fragmentation of molecular ions. In general, only singly-charged ions are formed during MALDI which provides an advantage in spectrum analysis due to less complexity. Low shot-to-shot reproducibility is one of the major disadvantages of MALDI, which is contributed mostly by the non-uniform distribution of matrix and analyte within the layers.

1.1.1.4 Electrospray Ionisation

Electrospray ionisation (ESI) was first introduced by John Fenn in 1984 who was then rewarded with a Nobel Prize in Chemistry in 2002 for this invention and its development.[30] ESI is generated at atmospheric pressure by applying a strong electric field to a liquid that flows out slowly from a capillary tube (see Figure 1.5).[31-33] The electric field is formed by applying several kilovolts (kV) potential difference between the capillary tube and the counter electrode, which are normally separated a few centimetre. Charges are then continuously generated and accumulated at the liquid surface that located at the tip end of the capillary (Taylor cone) under the electric field, which leads to the formation of highly charged droplets. The radial dispersion of the droplets is reduced by using a low flowrate, coaxial gas (sheath gas) and the solvent on the droplets is continuously evaporated by a heated inert gas, i.e. nitrogen, or a heated capillary. When the solvent has evaporated, the sizes of the droplets decrease while the number of charges retained on the droplets remains the same, which leads to an increase of the charge per droplet volume, resulting eventually in a "Coulombic repulsion". When the "Coulombic repulsion", which is generated by the charges on the droplet, is greater than the surface tension (Rayleigh limit), the droplets explode into

multiple smaller sized droplets.[34]   The daughter droplets generated continue to undergo further Coulombic explosions and then travel along the electric field and eventually form desolvated ions, which can be detected by the mass analyser. The detailed mechanism of generating ions from the charged droplets is still under debate, the ion ejection model (IEM), the charge residue model (CRM), and the chain ejection model (CEM), however, are the most commonly accepted models to explain the ions generation of small molecules, large folded proteins, and unfolded proteins respectively (Figure 1.6).
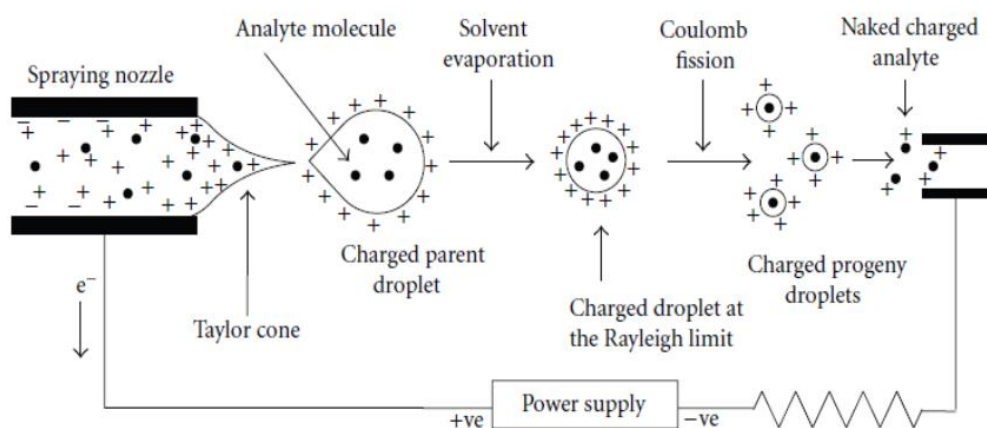
Figure 1.5.   The principle of ESI source (Reproduced from Banergee *et al.* 2012).[33]
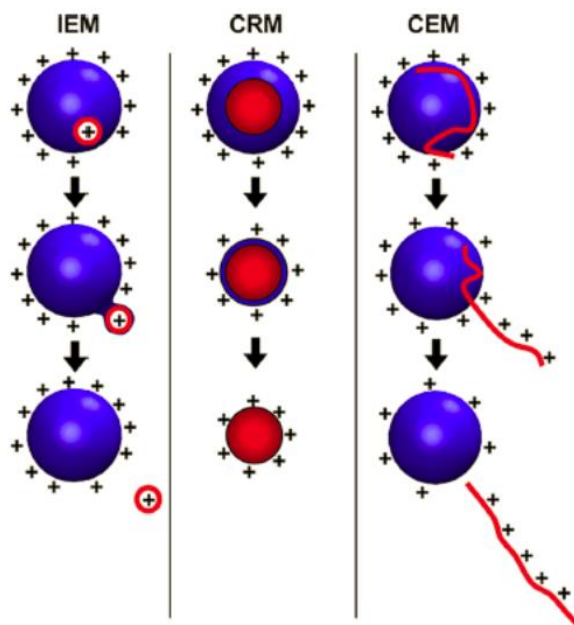
Figure 1.6. ESI models, including ion ejection method (IEM), charge residue method (CRM), and chain ejection method (CEM), that are proposed to explain the formation of gas phase ions from charged droplets (Reproduced from Konermann *et al.* 2012).[35]

ESI allows the ionisation of very fragile analytes without disrupting the non-covalent interactions between molecules.[36, 37] Compared to MALDI, ESI generates multiply charged ions through protonation or deprotonation, which allows large biomolecules to be analysed at a low *m/z* range and can improve sensitivity. Furthermore, ESI can easily couple with high performance liquid chromatography (HPLC) and capillary electrophoresis (CE), which are the two common methods for the separation of biomolecules.[38, 39] ESI, therefore, has a wide range of applications including proteins,[40] DNA,[41] polymers,[42] and small polar molecules.[43] The major disadvantages of ESI include sensitivity to salt contaminants and inability to effectively ionise many non-polar compounds.

1.1.1.5 Nano Electrospray Ionisation

Nano electrospray ionisation (nESI) is a nano flow analogue of ESI (section 1.1.1.4) to increase sensitivity of ESI.   The orifice of the nESI emitter (~ 0.5 –5 µm) is much smaller than ESI, the droplets formed in the ionisation process, therefore, are initially much smaller.[44]   The smaller droplets formed during the nESI process are believed to allow a more efficiency desolvation which results in higher efficiency of converting charged droplets to gas phase ions, leading to improvement in ionisation efficiency compared to the ESI source.   A schematic of nESI is shown in Figure 1.7.
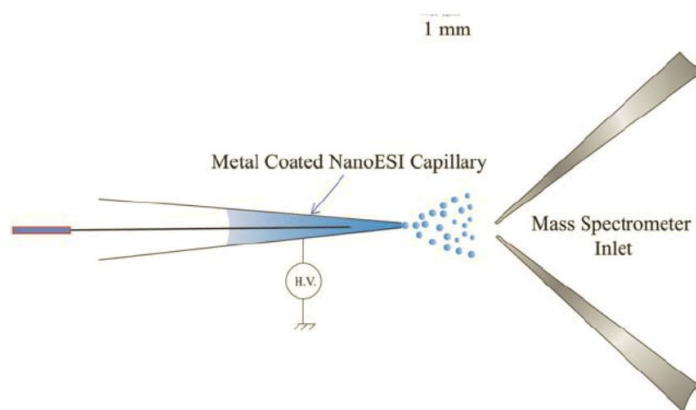
Figure 1.7.   The schematic diagram of nESI (Reproduced from Mandal *et al.* 2012).[45]

In addition to the difference in sample consumption, lower voltages (0.5 – 1.5 kV) can be used in the nESI process.   Furthermore, the source temperature applied in the nESI can be lower than the ESI as the droplets formed are smaller which requires less desolvation.   This feature is of particular importance for the studies of native proteins using MS as high source temperature can unfold/denature the proteins.[46]

Chapter 1 Introduction

1.1.2 Mass Analysers

Once ions are generated from the ionisation source, they are transmitted down to the mass analyser by electric potentials, separated according to their *m/z*, then detected and measured by the mass analyser.   In this section, three major mass analysers, quadrupole, time-of-flight (TOF), and Fourier transform ion cyclotron resonance (FTICR) MS, will be discussed.

Different mass analysers have different principles to separate ions, for example, quadrupole separates ions according to their *m/z* trajectory stabilities, TOF separates ions according to their flight velocities (flight time), and FTICR separates ions according to their resonance frequencies.   Mass analysers can be further classified into two groups according to the ion transmission properties. Some analysers, i.e. quadrupole, allow only ions with a particular *m/z* to pass through at a particular time, and some analysers, i.e. TOF and FTICR, allow the concurrent detection of all ions.

Mass range, speed of analysis, mass accuracy, and resolving power are the four main characteristics for evaluating the performance of a mass analyser. Mass range is the limitation of *m/z* range that a mass analyser can measure in a single scan.   Speed of analysis is the rate required to measure over a particular mass range in a mass analyser.

Mass accuracy is the accuracy of the *m/z* measured by a mass analyser which is calculated by the difference between theoretical (exact) and measured *m/z*. Parts-per-million (ppm) is the common unit for the expression of mass accuracy (see Eq. 5).   Mass accuracy is also affected by the stability and the resolving power of a mass analyser (if other species interfere with the measurement).   Mass analysers with high mass accuracies help to improve elemental composition determination.[47]

$$\text{Error (ppm)} = \frac{(measured\ m/z - theoretical\ m/z)}{theoretical\ m/z}\ x\ 10^6 \qquad \text{[Eq. 5]}$$

Resolving power is the ability of a mass analyser to distinguish two signals from two ion packets with a small difference in their *m/z* ratios. The resolving power can be calculated by m/$\triangle$m$_{50\%}$ (see Figure 1.8), where m is the *m/z* value of the peak and $\triangle$m$_{50\%}$ is the peak width measured at half maximum (FWHM). A greater resolving power indicates the mass analyser has a higher ability to distinguish ions with small *m/z* differences, therefore, a high resolving power mass analyser is useful in complicated sample measurement.[48, 49]

m = 976.724449
Δm = 0.04
Resolving power: 24,418
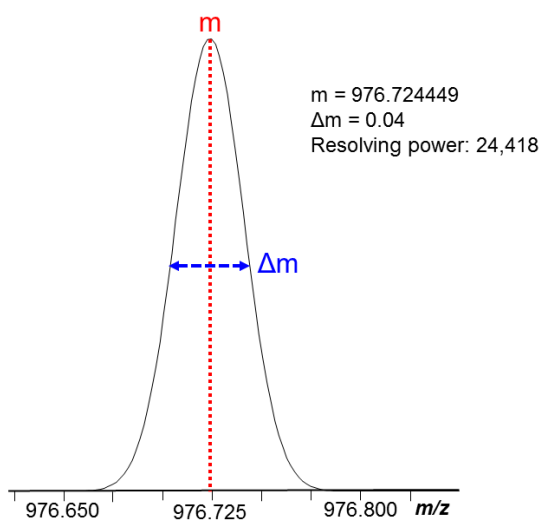
Figure 1.8. Resolving power calculated by Full Width Half Maximum (FWHM).

1.1.2.1 Quadrupole Mass Analyser

The principle of the quadrupole analyser was first introduced by Paul and Steinwegen in 1953.[50] The quadrupole was further developed into a commercial instrument by Shoulders, Finnigan, and Story.[51]

A quadrupole consists of four circular rods which are parallel to each other and the rods are connected oppositely in pairs (see Figure 1.9). Ions are injected and travel along the z axis, and oscillate in the x and y plane. The oscillation of ions is influenced by the electric field generated by direct current (DC), voltage

(U), and radio frequency (RF) amplitude (V) on the rods. The potentials on the different pairs of rods are set to be out of phase by 180° and rapidly oscillate between the pairs of rod, which results in the formation of low field regions (saddle points) to transmit particular *m/z* ions (see Figure 1.10).[8, 52, 53] Only particular ions with the stability of the trajectories at the particular U and V values can be transmitted and detected, while other unstable ions are discharged onto the rods, neutralised and cannot be detected.
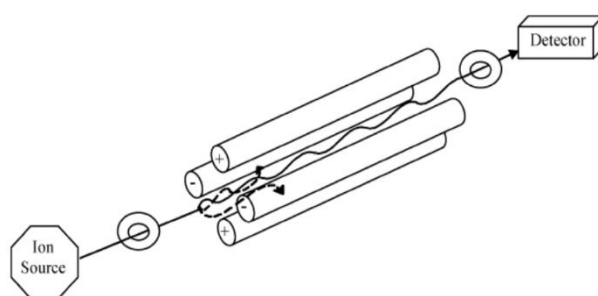
Figure 1.9. The layout of a quadrupole which contains four parallel rods, connecting diagonally in pairs of positive and negative terminals (Reproduced from EI-Aneed *et al.* 2009).[29]
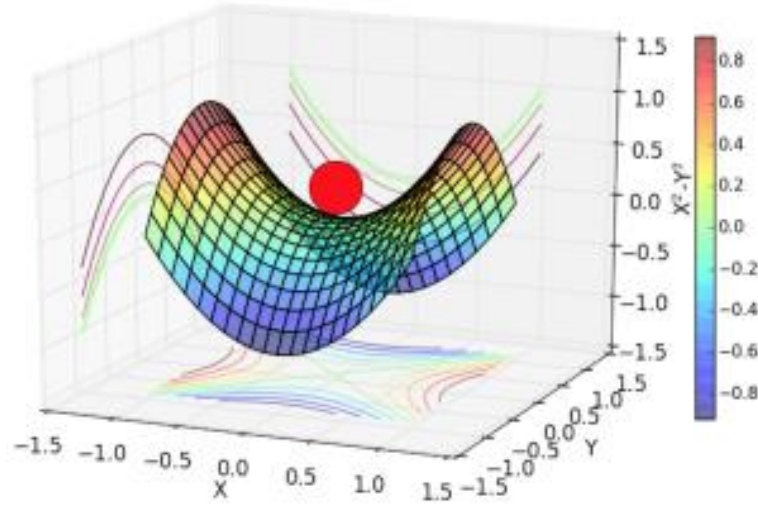
Figure 1.10. The formation of saddle points by rapidly oscillating the electric potential on the opposite pair of rods in order to retain the positive charged ion in the centre of the quadrupole (See https://aquadrupauliontrap.wordpress.com/the-mechanical-analogue for information, accessed 02nd Oct, 2018).

Ions are accelerated before entering the quadrupole, then remain at the same velocity in the z axis during the transmission through quadrupole. The trajectory stability of ions in the xz and yz planes can be represented by the following equations:[54]

$$a = \frac{8qU}{mr_0^2 \omega^2} \qquad \text{[Eq.6]}$$

$$q = \frac{4qV}{mr_0^2 \omega^2} \qquad \text{[Eq.7]}$$

where a and u are directly proportional to the U and V values respectively, q is the ion charge, m is the ion mass, r_0 is the field radius, and ω is the angular frequency. For different *m/z* ions, the a and q values obtained from equation 6 and 7 are different which results in generating various stability areas. Since the U and V values are the two variables for different *m/z* ions, a stability area as a function of U and V values containing various *m/z* ions can be plotted (see Figure 1.11). As

shown in Figure 1.11, each *m/z* ion is stable under its stability area and is discharged outside the area. During data acquisition, the quadrupole varies the U and V values.
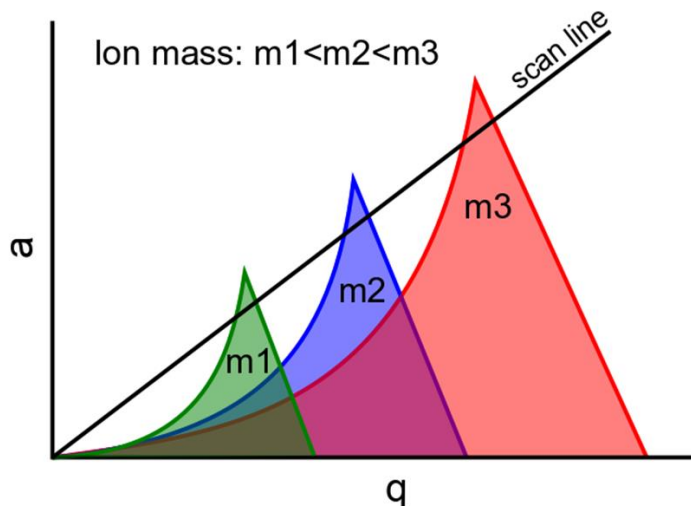
Figure 1.11. Representation of the stability diagram of ions in a quadrupole. Ions are represented by m1, m2, and m3 with increasing the mass respectively. Each ion is stabilised under its area, i.e. m1 ion is only stabilised in the green area, and being transmitted down the quadrupole to the detector. The quadrupole scans according to the scan line and only a particular ion can be transmitted to the detector each time.

The quadrupole is commonly used as an ion guide, collision cell, and mass analyser by varying the U/V value. When a quadruple is an ion guide or collision cell, the U value is set to zero and the quadrupole is RF-only, which allows all ions to pass through the quadrupole down to the low mass cut off (LMCO). When a quadrupole is a mass analyser, the U/V value is adjusted to a particular ratio that only specific *m/z* ion can pass through the quadrupole, the resolving power of quadrupole, therefore, can be adjusted by varying the U/V value.

In general, the mass range of a quadrupole is approximately one decade , the

FWHM resolution is usually around 2000 (at *m/z* 1000), the mass accuracy is around 100 ppm, and the operating pressure is around 10^{-5} Torr.[8]   The quadrupole is widely applied in different research studies because of its robustness, reliability, small footprint, and lower-cost production.[55, 56]   Quadrupoles are also commonly used as an ion guide and collision cell in order to couple with other mass analysers, e.g. TOF and FTICR, to achieve tandem MS (section 1.1.3).[57, 58]   The major disadvantages of a quadrupole are low resolving power and limited mass range.

1.1.2.2 Time-of-flight Mass Analyser

The principle of the time-of-flight (TOF) mass analyser was first introduced by Stephens in 1946 and was further designed into the first commercial instrument (linear TOF mass analyser) in 1955 by Wiley and McLaren.[59, 60]

In TOF-MS, ions generated from an ion source are accelerated through the flight tube by an electric field which is caused by the potential difference between the electrode and extraction grid.   To a first-order approximation, all ions obtain same kinetic energy in the acceleration region, and drift in a flight tube which is a free-field region according to their *m/z*, and reach the detector at different times, allowing separation of ions by *m/z* ratios (see Figure 1.12).   The process of ion drifting in TOF mass analyser can be represented by the following equations:

Kinetic energy obtained by the ion:

$$E = \frac{1}{2}mv^2 = qV \qquad\qquad\qquad\qquad \text{[Eq. 8]}$$

Velocity of the ion (rearranged from Eq. 8):

$$v = \sqrt{\frac{2qV}{m}} \qquad\qquad\qquad\qquad \text{[Eq. 9]}$$

where E is the kinetic energy obtained by the ion, m is the ion mass, v is the ion velocity, q is the ion charge, and *V* is the electric potential generated from the

potential difference between electrode and extraction grid.
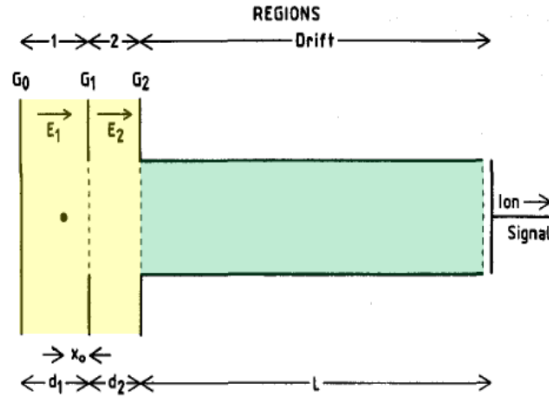
Figure 1.12. The outline of a linear TOF (Reproduced from Opsal *et al.* 1985).[61] Ions (black dot) are accelerated in the yellow region by the electric potential, drift through the green region (flight tube), and reach the detector (right).

Eq. 9, it shows the velocity of ion is inversely proportional to the mass$^{1/2}$ which indicates heavier ions have a lower velocity, thus they require a longer time to drift through a flight tube and vice versa. The time for ion to reach a detector can be calculated by the following equation:

$$t = \frac{L}{v} \qquad \text{[Eq. 10]}$$

where t is the ion drift time across the flight tube and L is the length of flight tube. The time for ions to reach a detector can be presented in terms of ion mass by combining Eq. 9 and Eq. 10 as follows:

$$t = \sqrt{\frac{mL^2}{2qV}} = \sqrt{\frac{m}{z}} \sqrt{\frac{L^2}{2eV}} \qquad \text{[Eq. 11]}$$

where z is the number of charge on the ion and e is the charge constant.

Previous results showed that a linear TOF mass analyser could detect protein samples in $100-200$ attomole amounts.[62, 63] In a linear TOF mass analyser, there is, in theory, no upper mass limit which makes it suitable for the detection of large biomolecules and polymers.[64, 65] Furthermore, the analysis speed of TOF is very

fast, in which a spectrum can be obtained in ~100 micro-seconds.[8]

In reality, the kinetic energies obtained by each ion are slightly different, which results in the spreading of ions (with same *m/z*) drift times and causes peak broadening and decreases mass resolution.   In order to minimise the spreading of ion's (same *m/z*) drift time, delayed pulsed extraction and reflectron are the two common techniques that applied to correct the kinetic energy differentiation between ions (see Figure 1.13).   In delayed pulsed extraction, a delayed pulse is applied to all ions and ions with lower kinetic energy stay in the acceleration region longer than higher kinetic energy ions which results in lower kinetic energy ions obtain more energy before entering the flight tube.   Since the initial lower kinetic energy ions receive more energy in the acceleration region, the velocities of these ions increase and catch up with the initial higher kinetic energy ions in the flight tube, both ions reach the detector at the same time.
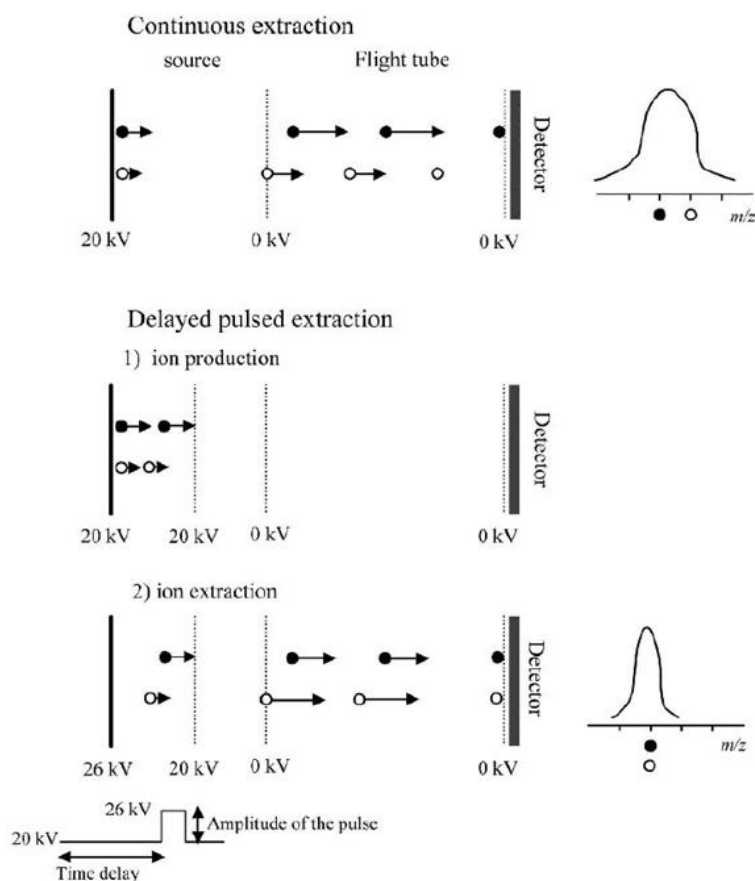
Figure 1.13. The implication of a delayed pulsed extraction in a TOF mass analyser which improves the mass resolution (Reproduced from De Hoffmann *et al.* 2007).[8] The black dot represents an ion with higher kinetic energy, while the white dot represent an ion with lower kinetic energy. The white dot (low kinetic energy ion) stays longer in the acceleration region which gains more energy before entering the flight tube, as a result the velocity of white dot is higher than the black dot and both dots reach the detector at the same time.

The reflectron is also used to correct the kinetic energy spreading between ions. This idea was first introduced by Mamyrin in 1973.[66] The reflectron, placed opposite to the source and detector, acts like a mirror which reflects the ions coming from the source to the detector which is placed next to the ion source (see

Figure 1.14).   The ions with higher kinetic energy penetrate the reflectron more deeply than the ions with lower kinetic energy, the higher kinetic energy ions, therefore, stay in the reflectron longer than the lower kinetic energy ions which results in both ions reach the detector at the same time.   With the applications of delayed pulse extraction and reflectron, the accuracy of the TOF mass analyser can be improved from 200 ppm to 10 ppm and the FWHM resolution is enhanced from 5,000 to 20,000 (at *m/z* 1,000).[8]
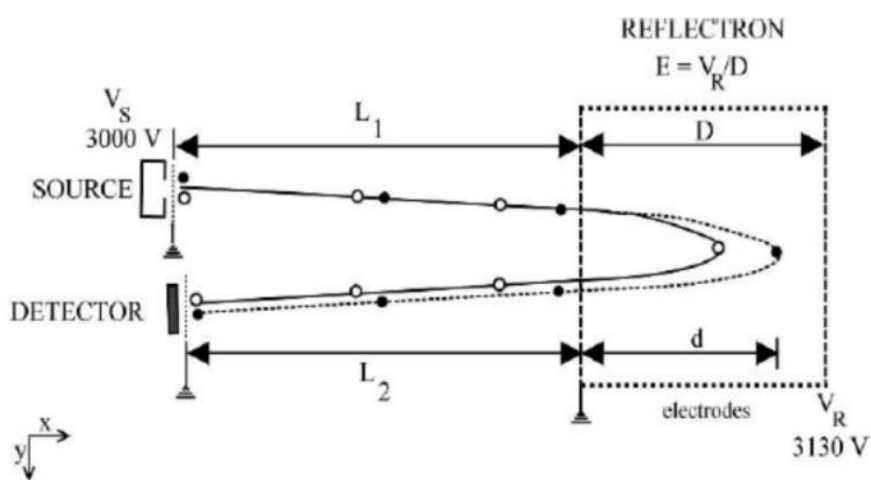
Figure 1.14.   The application of a reflectron in a TOF mass analyser which helps to improve the mass resolution (Reproduced from De Hoffmann *et al.* 2007).[8] The black dot represents an ion with higher kinetic energy, while the white dot represents an ion with lower kinetic energy.   Since the black dot contains higher kinetic energy, it penetrates the reflectron deeper than the white dot, which results in arriving the detector at the same time.

1.1.2.3 Fourier Transform Ion Cyclotron Resonance Mass Analyser

The Fourier transform ion cyclotron resonance (FTICR) mass analyser was first introduced by Marshall and Comisarow in 1974.[67]   FTICR has been widely applied in scientific research due to its high resolving power and mass accuracy.[68-

[71] Herein, the major principles of FTICR and its applications will be discussed.

Cyclotron motion. In a FTICR mass analyser, an FTICR cell is placed in superconducting magnet which generates a homogenous magnetic field across the cell. When ions in the FTICR cell move perpendicular to the magnetic field, the ions will experience the "Lorentz force" which is perpendicular to the ion motion as well as the magnetic field (see Figure 1.15). The Lorentz force can be calculated by the following equation:

$$F = qv \text{ x } B \qquad\qquad\qquad [Eq.12]$$

where F is the Lorentz force experienced by the ion, q is the ion charge , v is the ion velocity, and B is the magnetic field strength of the magnet.

Under the influence of magnetic field, ions move around the centre of the magnetic field axis which results in an orbital movement, and this motion is called "cyclotron motion". The theoretical cyclotron frequency of an ion can be represented by the following equations:

$$\omega_c = \frac{qB}{m} \qquad\qquad\qquad [Eq.13]$$

$$f_c = \frac{\omega_c}{2\pi} = \frac{qB}{2\pi m} = \frac{1.535611 \times 10^7 B}{m/z} \qquad\qquad\qquad [Eq.14]$$

where ω_c is the ion cyclotron frequency, m is the ion mass, f_c is the cyclotron frequency expressed in Hertz. Eq. 14 shows the cyclotron frequency of an ion is inversely proportional to its m/z ratio which indicates the heavier ion rotates with a lower cyclotron frequency and vice versa. Because of this relationship, the cyclotron motion can be used to distinguish different m/z ratios and thus be used as a mass analyser.
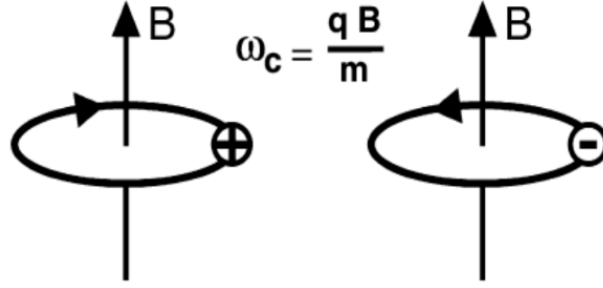
Figure 1.15. The cyclotron motion of a positively charged ion (left) and a negatively charged ion (right) in the presence of magnetic field (B) (Reproduced from Marshall *et al.* 2002).[69]

In general, the ICR frequencies of ions range from a few kHz to a few MHz. Eq. 14 clearly shows that the cyclotron frequency of an ion is independent of its velocity and kinetic energy, which is a key feature for FTICR to achieve a high resolving power. In other mass analysers, i.e. TOF, delayed pulsed extraction and reflectron are required to correct the energy spreading issue which causes poor resolution (section 1.1.2.2).[8]

Trapping motion. If the magnetic field defines the z axis of Cartesian coordinates, ions orbit in the x and y plane inside an FTICR cell because of the influence of the magnetic field mentioned above, in order to prevent the ions "escaping" via the z axis dimension, two trapping plates with low electric potential are applied to create a potential well. When the energies of the ions are lower than the trapping voltage, the ions are trapped in the ICR cell and oscillate back and forth between the plates. The trapping motion frequencies of ions are around several kHz which can be represented by the following equation:

$$\omega_z = \sqrt{\frac{2qV_{trap}\alpha}{ma^2}} \qquad\qquad \text{[Eq.15]}$$

where ω_z is the ion oscillating frequency in z axis induced by the trapping motion,

q is the ion charge, V_{trap} is the trapping potential applied in the trapping plates, and α is the value determined by the trapping geometry (type of ICR cells), and a is the plate-to-plate distance in the cell.

Magnetron motion. The trapping motion induced by the two trapping plates is an effective method to prevent ions to escape from the z axis inside an FTICR cell. However, it also induces the magnetron motion, which is an undesirable side effect generated from the trapping motion. In between the two trapping plates, an electrostatic field is generated, ions in the FTICR cell experience the magnetic force as well as the electric force, which results in a second motion in the x and y plane which is named the magnetron motion. The magnetron motion induced by the trapping plates can be represented by the following equation:

$$\omega_m = \frac{V_{trap}\alpha}{2\pi B a^2} \qquad\qquad [Eq.16]$$

where ω_m is the magnetron frequency of the ion, V_{trap} is the trapping potential applied in the trapping plates, α is the value determined on the trapping geometry (types of ICR cells), B is the strength of magnet, and a is the plate-to-plate distance in the cell.

Compared to the cyclotron motion, the magnetron motion acting on an ion is much slower (with the frequency <100 Hz). Since the cyclotron motion and magnetron motion are acting in the x and y plane, the frequency measured by an FTICR cell is a "reduced" cyclotron frequency which can be represented by the following equation:

$$\omega_{measured} = \omega_c - \omega_m \qquad\qquad [Eq.17]$$

where ω_c is the cyclotron frequency and ω_m is the magnetron frequency.

A high trapping potential is preferred to effectively trap the ions in an FTICR

cell, however, Eq. 16 and Eq.17 show that a high trapping potential increases the magnetron motion which results in a larger offset of the ideal cyclotron frequencies of the ions. Thus, a balance between high ion trapping efficiency and interrupting the ideal cyclotron frequency is required. The summarised ion motions, cyclotron motion, trapping motion, and magnetron motion, in a penning trap are shown in Figure 1.16.
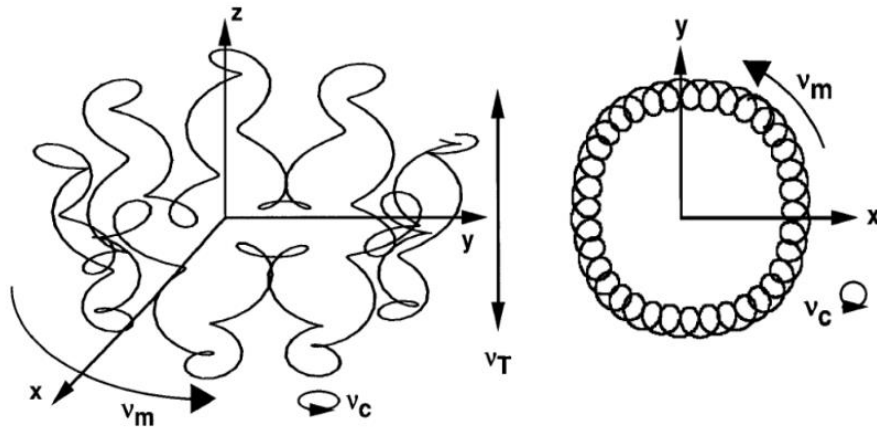
Figure 1.16. The cyclotron motion (υ_c), trapping motion (υ_m), and magnetron motion (υ_m) in a penning trap (Reproduced from Marshall *et al.* 1998).[70]

Excitation and detection in FTICR cell. When ions enter an FTICR cell, the cyclotron radii of the ions are too small to generate a detectable signal. Furthermore, because the initial phase of the ions is random, the cyclotron frequencies of the ions in the cell are incoherent, the current generated on one detector is balanced out by the other detector plate, which results in zero net current generated on the detector plates. Thus, an excitation pulse which is ideally generated by using a spatially uniform electric field with an amplitude (E_0) that is perpendicular to the magnetic field is applied to the ions. The electric field used for excitation has the same frequency as the cyclotron frequency of an ion (resonant). During excitation, the size of the ion cloud remains the same while

the ion cloud accelerates with a spiral trajectory,[72] which results in a larger, detectable cyclotron radius (see Figure 1.17). The radius of the post-excited ion cloud can be represented by the following equation:

$$r = \frac{E_0 T_{excite}}{2B}$$ [Eq. 18]

where r is the radius of the post-excited ion packet, E_0 is the electric field applied to excite the ions, T_{excite} is the excitation duration, and B is the magnetic field.
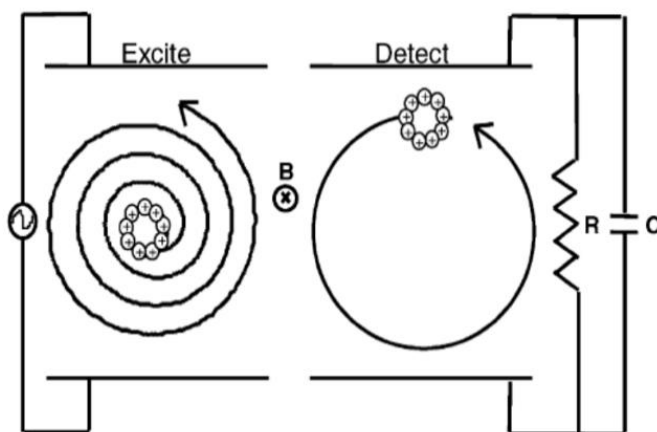
Figure 1.17. Representation of an ion cloud in an FTICR cell when excited (left) showing acceleration with a spiral trajectory into a large, detectable orbit (right) (Reproduced from Marshall *et al.* 2002).[69]

According to Eq. 18, the radii of the post-excited ion clouds are independent of *m/z*; therefore, all ions can be excited to the same cyclotron orbits in the FTICR cell. Frequency-sweep excitation (chirp) and stored waveform inverse Fourier transform (SWIFT) are the two common excitation methods applied in the FTICR cell to increase the cyclotron radius of the ion cloud for detection. The chirp, which was first proposed by Marshall in 1974,[73] excites the ions over a broad frequency range by slowly sweeping through the relevant frequencies with a low excitation voltage (see Figure 1.18A).[74] However, the excitation amplitude of chirp is non-uniform particularly towards the ends of the frequency range, which

results in influencing the observed relative abundances of the ions, as well as, the poor mass selectivity at the beginning and the end of the sweep frequencies which limits the performances of chirp in an FTICR cell. The SWIFT was then proposed by Marshall which reversed the excitation process.[75] The desired excitation mass is first defined and converted back into a frequency domain spectrum, and then stored in memory to generate the time domain waveform to excite the desired *m/z* ion; as a result, SWIFT provides the flattest broadband excitation (see Figure 1.18B). SWIFT is further applied to isolate the desired *m/z* ions in an FTICR cell by over exciting the undesired *m/z* ions to an orbit larger than the FTICR cell, thus only desired *m/z* ions are retained in the cell (see Figure 1.18C).
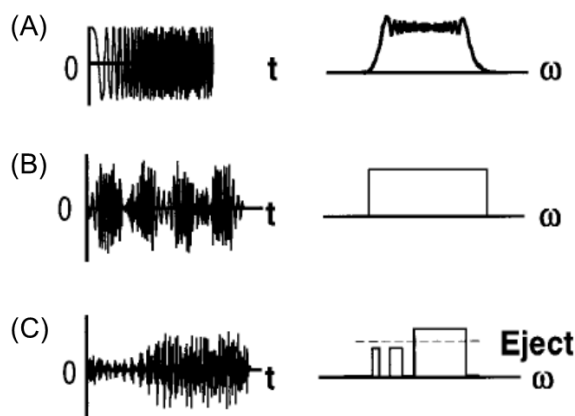
Figure 1.18. The excitation methods applied in FTICR cell using (A) frequency-sweep excitation (chirp) and (B) stored waveform inverse Fourier transform (SWIFT). (C) Ions ejection from the FTICR cell by applying the SWIFT excitation (Reproduced from Marshall *et al.* 1998).[70]

The excited ion cloud circulates between the two detection plates which generates a differential current. The detected ICR signal is proportional to the differential current which can be represented by the following equation:[76]

$$\frac{\Delta Q}{dt} = -\frac{2q(\frac{dy}{dt})}{d}$$

[Eq. 19]

where q is the charge of the ion, y is the position of ion in the y-axis, and d is the

distance between the ion and the detection plate.

According to Eq. 19, the ICR signal increases proportionally with the post-

excitation radius as dy/dt is increased with a larger radius.   Ions with different

m/z are excited into the same orbit in the FTICR cell (Eq. 18) and the detection in

the ICR cell is non-destructive which allows the simultaneous detection of a wide

m/z range in the cell.   Furthermore, the FTICR signal is also increased according

to the number of charges on the ion, thus, the sensitivity is enhanced for multiply-

charge ions.[69, 70]

FTICR cells.   Cubic and cylindrical cells are two common geometries in

FTICR cell design.[70]   The cubic cell contains three sets of plates which are placed

opposite to one another and arranged like a cubic box.   The three sets of opposed

plates are used for excitation, detection, and ion trapping (see Figure 1.19A).

There are two common types of cylindrical cells which are the open

cylindrical cell and the cylindrical cell with end-caps, one of which has segmented

end-caps which is known as the infinity cell.[77]   For the open cylindrical cell, the

middle part of the cell contains two sets of opposed, cylindrical plates which are

used for excitation and detection, and there are two extra sets of cylindrical plates

placing to the left and the right of the middle part of the open cylindrical cell which

are used for ion trapping (see Figure 1.19B & 1.19C).[78]   The open cylinder cell

design allows effective ion trapping in the z axis with the minimum influence of

the ion cyclotron frequencies.   For the infinity cell, the middle part of the cell also

contains two sets of opposed, cylindrical plates which are used for excitation and

detection, two end plates are placing to the left and right of the cylindrical excitation and detection plates (see Figure 1.19B).[77]   The trapping voltages are carefully divided across the trapping plates in order to trap the ions in the z axis, but minimise z-axis ejection during excitation and drastically increase sensitivity.
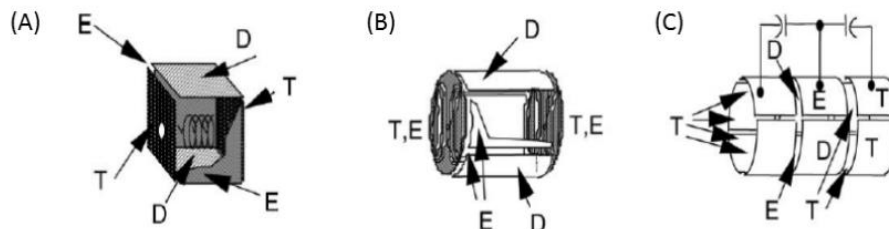
Figure 1.19.   The configurations of (A) cubic, (B) infinity, and (C) open-ended cylindrical cells (Reproduced from Marshall *et al.* 2002).[69]   E is the plate for excitation, D is the plate for detection, and T is the plate for trapping.

Resolving power.   As mentioned above (Section 1.1.2), resolving power is useful to determine the capability of a mass analyser to separate the ions with a close *m/z* differences.   The resolving power of FTICR can also be represented as follows:[79, 80]

$$\frac{m}{\Delta m} = \frac{1.274 \; x \; 10^7 \, zBT_{aqn}}{m} \qquad \text{[Eq. 20]}$$

where m is the mass of ion, z is the charge of ion, B is the magnetic field strength, and T_{aqn} is the transient length.

According to Eq. 20, the resolving power of FTICR is determined by the length of transient and the strength of magnetic field.   Ideally, the transient length of an ion can last forever in a FTICR cell; however, the coalescence as well as the ion cloud radius (and detectable signal) decays with acquisition time as there are collisions between the ions and the residual background gas molecules, and the ion cyclotron frequency is interrupted which results in the decay of the transient signal.

Thus, the pressure of a FTICR cell is always required to be maintained at $< 10^{-10}$ torr in order to reduce the number of background gas molecules in the cell.

Furthermore, the resolving power increases proportionally to the strength of magnetic field according to Eq. 20. The strengths of magnets used in FTICR experiments range from 1.5 to 21 Tesla (T), and a recent study of a complex crude oil sample showed that a 21 T FTICR MS contains 2.2-fold higher mass-resolving power than a 9.4 T FTICR MS.[81]

Dynamic range. To generate a detectable image current in FTICR, currently the lowest required amount of ions required is ~ 187.[69, 70] The peak intensity in FTICR MS spectra is directly proportional to the number of ions in a FTICR cell, thus more ions are always preferred in order to generate a more intense peak with a significant signal-to-noise (S/N) ratio for peak assignment. However, when more than ~ 10,000 ions are trapped in the FTICR cell, Coulombic repulsions between ions become significant which results in shifting and broadening the peaks in FTICR mass spectrum depending on the strength of the magnetic field.[70] This phenomenon is also known as the space charge effect. Furthermore, peak coalescence can occur when two ions contain similar *m/z* ratio and both of the ion clouds are significant large enough to coalesce.[76] The space charge and peak coalescence effects in FTICR are inversely proportional to the magnetic field strength, indicating a higher magnet field strength is always desired to achieve a greater dynamic range in FTICR cell.

1.1.3 Tandem Mass Spectrometry

Tandem mass spectrometry (MS/MS) involves at least two stages of MS detection, precursor ions are either dissociated into fragment ions or involved in a

reaction which leads to the change of mass or charge of the ions.[8]   Fragment ions

generated in MS/MS provide extra information for molecular assignment and

structural elucidation, thus MS/MS is applied to proteomics studies.[82-84]

Many fragmentation techniques have been developed and are widely applied

in proteomics studies to sequence biomolecules.   Different fragmentation

techniques show advantages in certain application areas.   Common

fragmentation techniques include collisionally activated dissociation (CAD),[82]

infrared multiple photon dissociation (IRMPD),[85] sustained off-resonance

irradiation collision-induced dissociation (SORI-CID),[82] ultraviolet

photodissociation (UVPD),[86] electron capture dissociation (ECD),[87] electron

transfer dissociation (ETD),[83] electron induced dissociation (EID),[88] and electron

detachment dissociation (EDD).[89]   The possible fragment ions of a biomolecule

can be summarised by the Roesptorff nomenclature (see Figure 1.20).[90]   In this

section, the three major fragmentation techniques, including CAD, IRMPD, and

ECD, are discussed.
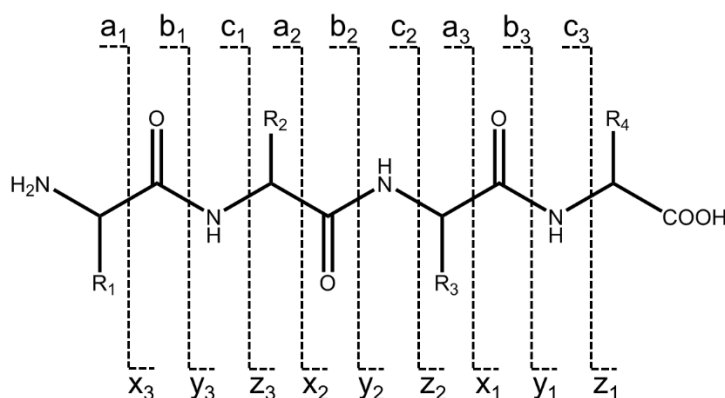
Figure 1.20.   The possible fragment ions of a biomolecule.

1.1.3.1 Collisionally Activated Dissociation

Collisionally activated dissociation tandem MS (CAD MS/MS) is the most

common fragmentation technique applied in proteomics studies as CAD is compatible with most of the mass analysers, e.g. quadrupole, TOF, and FTICR MS. CAD MS/MS was first introduced by Jennings in 1968.[91]

In CAD MS/MS, precursor ions are accelerated by the electric fields in the ion optics and then collided with a non-reactive gas molecule, usually helium, argon, or nitrogen. The kinetic energy from the precursor ion is converted into internal energy and results in fragmentation. The internal energy is rapidly distributed across the precursor ion, and causes the "slow heating" effect in ions, usually breaking the weaker bonds available, thus is known as an ergodic process. In a protein/peptide, therefore, the fragmentation occurs at the weakest amide bond of the ion. The fragmentation at amide bonds results in generating b and y ions which contain the N-terminus and C-terminus of the protein/peptide respectively (see Figure 1.21).[92] CAD MS/MS does not only generate b and y ions, but also produces the loss of neutral molecules, i.e. H_2O and NH_3, which increase the complexity of a MS spectrum.
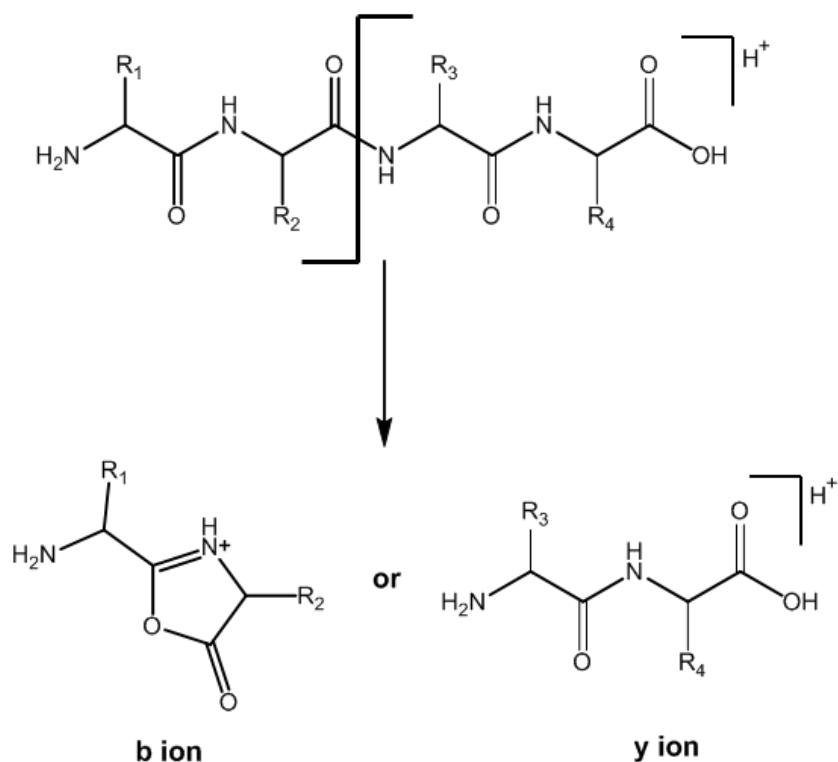
Figure 1.21. Diagram to show the formation of the b and y ions in CAD MS/MS fragmentation.[92]

CAD MS/MS is commonly applied in determining sequence information about biomolecules. Since CAD MS/MS is an ergodic fragmentation process, in which the weakest bonds are dissociated first, it is not suitable to apply CAD MS/MS to study of non-covalent interactions between biomolecules, as well as fragile post translation modifications (PTMs), i.e. glycopeptide and phosphopeptide.

1.1.3.2 Infrared Multiple Photon Dissociation

Infrared multiple photon dissociation tandem MS (IRMPD MS/MS) is another fragmentation technique that is commonly used to study biomolecules.[85, 93] Compared to CAD MS/MS, IRMPD MS/MS is usually applied on ion trapping

based instruments, i.e. ion trap, linear ion trap, and FTICR MS, only.[80, 94]

In IRMPD MS/MS, photons generated from an IR laser are absorbed by the precursor ion, which excites the ion to a higher energy state (vibrational state) and leads to a dissociation. Just like CAD MS/MS, IRMPD MS/MS is an ergodic process, the energies from the photons are re-distributed among the ion molecules which results in fragmentation of the weaker bonds (amide bonds in biomolecules). Thus, b- and y- ions are generated during IRMPD MS/MS (see Figure 1.21).

The carbon dioxide (CO_2) laser is the most common laser used in IRMPD MS/MS as it is widely available, relatively cheap, as well as providing a wavelength at 10.6 μm which is suitable to be absorbed by a broad range of biomolecules. Since each photon generated by a CO_2 (~25 W) laser only provides ~ 0.1 eV of energy, a high laser power or a long irradiation time (0.1 – 1 second) is required to produce effective IRMPD fragmentation. A prerequisite of IRMPD is that the ions of interest must be able to absorb the photons, i.e. they must have a chromophore.

1.1.3.3 Electron Capture Dissociation

Electron capture dissociation tandem MS (ECD MS/MS) was first introduced in 1998 by Zubarev *et al.* to fragment multiply charged proteins.[95] ECD fragmentation is not as common as CAD and IRMPD fragmentations because ECD MS/MS is usually implemented in FTICR MS,[83] even though recently ECD has also applied in the other types of mass analysers.[96, 97]

Compared to CAD and IRMPD MS/MS, ECD MS/MS is believed to be a non-ergodic process in that the precursor ion is dissociated before the energy is randomised in the ions.[98] In ECD MS/MS, low energy electrons, ~1-3 eV, are generated and captured by a multiply charged precursor ion and the N-C_α bond of

the ion is cleaved, even-electron c ions (with N-terminus) and odd-electron $z^\bullet$ ions

(with C-terminus) are formed (see Figure 1.22A).   At least one electron is

captured by the precursor ion during the ECD fragmentation process and results in

a charge neutralisation effect (see Eq. 21), thus only multiply-charged precursor

ions are suitable for ECD fragmentation to avoid generation of neutral species

which cannot be detected by MS.   The electron capture process can be

represented by the following equation:

$$[M + zH]^{z+} + e^- \; (1 - 10 \text{ eV}) \rightarrow [M + zH]^{(z-1)+\bullet}$$ [Eq. 21]
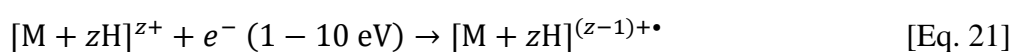
where M is the precursor ion, z is the charge of the ion, and e is the electron.

Apart from c and z ions, a and y ions may also be generated from the ECD MS/MS

as minor products (see Figure 1.22B).
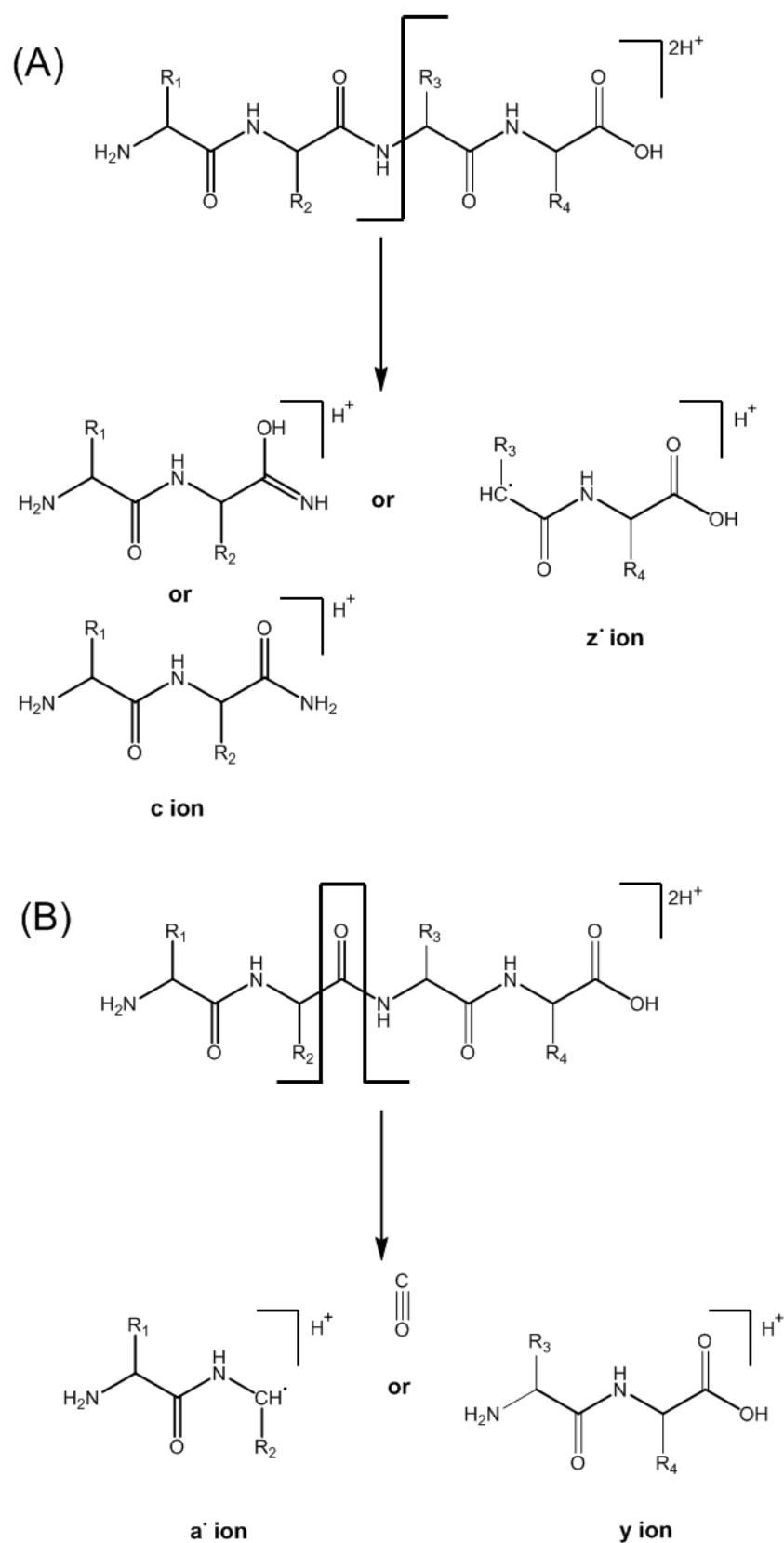
Figure 1.22. Diagram to show the generation of the (A) c and z ions and (B) a and y ions in ECD MS/MS fragmentation.[92]

35

Implementation of ECD. A hollow cathode behind the FTICR cell is used to generate electrons for ECD fragmentation. The kinetic energy difference between the electrons and precursor ion cloud is required to be as minimum as possible in order to maximise the electron capture process.[84]

ECD fragmentation mechanisms. Several fragmentation mechanisms were proposed to explain the generation of c and z$^\bullet$ ions during ECD MS/MS. Herein, the two major proposed ECD MS/MS fragmentation mechanisms are discussed.

Cornell mechanism. The Cornell mechanism was first proposed by the research group of McLafferty in Cornell University in 1998.[95] The Cornell mechanism suggests electrons are captured by protonated sites, including the amino group at N-terminus and the side chain of basic amino acids (histidine, arginine, and lysine), and form a hydrogen atom. The hydrogen atom is then attracted to the amide oxygen (at the protein/peptide backbone) and forms a carbon-centered aminoketyl radical precursor ion. The N-C$_\alpha$ bond that is adjacent to the carbonyl group is cleaved and forms c and z ions (see Figure 1.23)
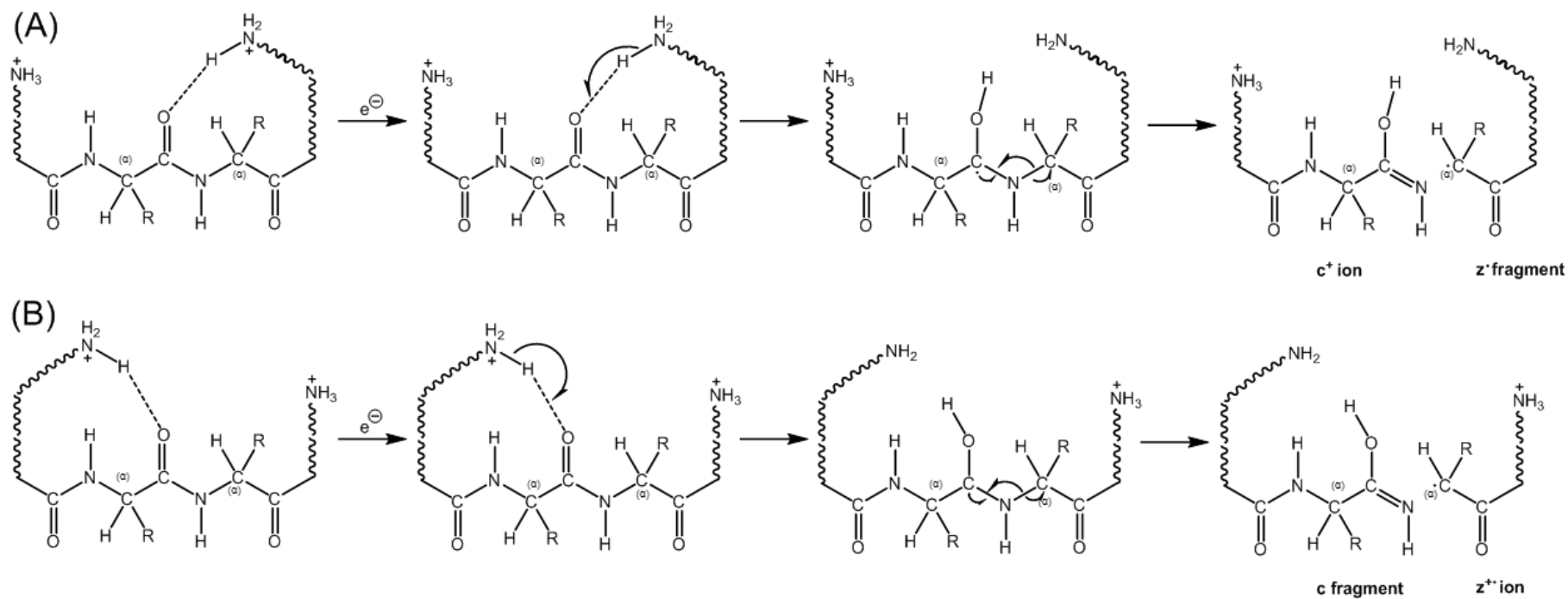
Figure 1.23. The Cornell mechanisms of ECD MS/MS with electron captured at the (A) C-terminus and (B) N-terminus to generate the c and z ions.[83]

Utah-Washington mechanism. The Utah-Washington mechanism was first proposed by the research group of Simons from University of Utah in 2003.[99] The mechanism was then further explained by the research group of Turecek from University of Washington in 2005.[100] In the Utah-Washington mechanism, electrons are directly captured at the amide group (the protein/peptide backbone) and formed a carbon-centered aminoketyl radical precursor ion. The $N-C_\alpha$ bond which is adjacent to the carbonyl group is cleaved and generates c and z ions (see Figure 1.24). The carbonyl anion generated during the electron capture process is then neutralised by a proton that is transferred from a charged site.
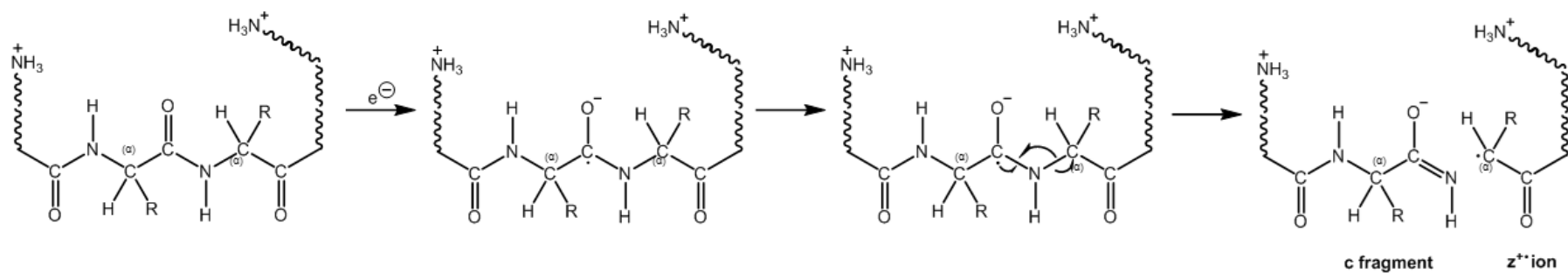
Figure 1.24. The Utah-Washington mechanism of ECD MS/MS to generate the c and z ions.[83]

Fragmentation efficiency of ECD MS/MS. The fragmentation efficiency can be measured by comparing the total abundance of fragment ions against the total abundance of the precursor ion. When well-tuned, the fragmentation efficiency of an ECD MS/MS efficient is ~ 20% for a doubly charged peptide ion, however, > 80 % fragmentation efficiency can be achieved with some highly charged protein ions.[84] The fragmentation efficiency of ECD in a highly ordered structural protein can also be enhanced by combining the fragmentation techniques of IRMPD and ECD MS/MS as the energy contributed by IRMPD can destabilise the internal non-covalent bonds which results in obtaining more c and z ion fragments.[101, 102]

Applications of ECD MS/MS. Being radical-based rather than slow-heating based, ECD MS/MS has a very different fragmentation mechanism compared to CAD and IMRPD MS/MS, thus it has several advantages in the following applications.

De Novo Sequencing. In the studies of ECD MS/MS, a high sequence coverage was achieved in the top-down spectra of ubiquitin (~ 8.6 kDa) and cytochrome c (~ 12.4 kDa),[84, 95] which indicated ECD has extensive cleavage and is useful for de novo sequencing analysis. Previous studies also showed that complementary fragments were also observed when both CAD and ECD MS/MS were applied in studying biomolecules, so that a larger sequence coverage was obtained which contributed to a higher successful rate in de novo sequencing of the biomolecules.[103]

Preservation of fragile PTMs and non-covalent interaction. Phosphorylation and glycosylation are two important protein modifications which relate to protein functionality.[104] The covalent linkages between the protein backbone and phospho-/glycan group are weaker than amide bonds, which results

in the PTM being lost rapidly with CAD and IRMPD MS/MS. Thus, it is challenging to locate the site of phosphorylation and glycosylation in the protein backbone using CAD and IRMPD MS/MS. ECD MS/MS, on the other hand, preserves these fragile linkages; while still fragmenting the protein backbone through radical mechanisms and results in providing sufficient information to localise the fragile PTMs on the biomolecules.[4, 105]

Non-covalent interactions between proteins and protein-ligands are commonly observed; however, it is challenging to localise the interaction sites between these molecules as the non-covalent bonds are weaker than the covalent bonds in the protein backbone, which results in fragmentation at the non-covalent interaction site when using CAD and IRMPD MS/MS, losing structural information. However, ECD MS/MS can localise the interaction site of proteins and protein-ligand complexes by fragmenting the protein backbone while preserving the non-covalent bonding. Using ECD MS/MS, the non-covalent interaction sites between proteins and protein-ligands can be located quite accurately, often to several amino acid residues.[37, 106, 107]

Disulfide Bond Cleavage. In CAD and IRMPD MS/MS, the presence of a disulfide bond significantly hinders protein fragmentation.[108] Thus, reduction and alkylation are usually required prior to CAD and IRMPD MS/MS in order to yield a good sequence coverage of the protein. In ECD MS/MS, the electron is preferentially captured at the disulfide bond site which results in a reduction of the S-S bond, and allows fragmentation on the protein backbone (see Figure 1.25).[92, 109]
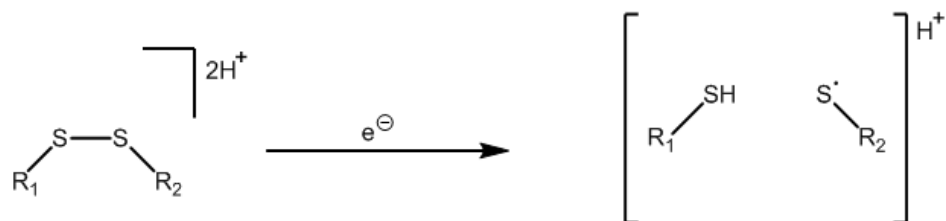
Figure 1.25. Diagram to show the electron is captured by the disulfide bond and results in the S-S bond cleavage.[92]

Differentiation of isomeric products from deamidation. Iso-aspartic acid/ aspartic acid and γ-glutamic acid/ glutamic acid are the common products obtained during deamidation of proteins (detail discussion in section 1.2.2).[110, 111] The mass of iso-aspartic acid and γ-glutamic acid are exactly the same as aspartic acid and glutamic acid respectively, the only difference is the methylene group ($-CH_2-$) from the side chain of aspartic acid and glutamic acid is added into the backbone of iso-aspartic acid and γ-glutamic acid (see Figure 1.26). It is challenging to differentiate these isomeric products using CAD and IRMPD MS/MS. ECD MS/MS, however, can differentiate the iso-aspartic acid and γ-glutamic acid from aspartic acid and glutamic acid respectively.[112, 113] In iso-aspartic acid and γ-glutamic acid, there is a possibility that the bond cleavage takes place at the C_α-C bond rather than the N-C_α bond at the iso-aspartic acid and γ-glutamic acid deamidation site, which results in a formation of a c ion with an addition of 58 Da ($C_2H_2O_2$) at the deamidation site and a complementary z ion with a loss of 57 Da (a hydrogen ion is used to convert $z^\bullet$ to z, thus a loss of 57 Da instead of 58 Da) (see Figure 1.27).[112, 114, 115]
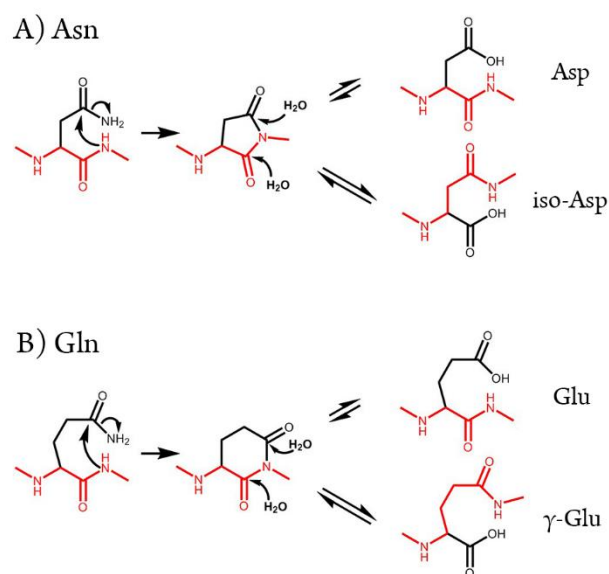
Figure 1.26. Deamidation mechanisms of (A) asparagine and (B) glutamine. The protein backbone is coloured in red; while the amino acid side chain groups are coloured in black.
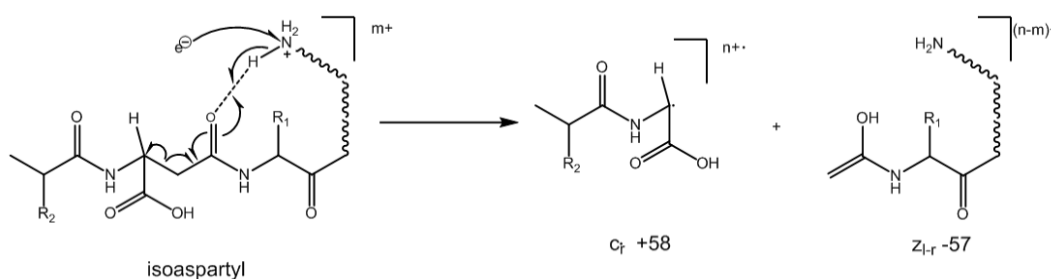
Figure 1.27. Diagram to show the ECD fragmentation mechanism at the iso-aspartic acid deamidation site to generate a (c+58 Da) ion and a (z-57 Da) ion.[114]

Other electron-based fragmentation techniques. Since ECD MS/MS is usually applied on FTICR MS, electron transfer dissociation (ETD) was then invented in 2004 by Syka *et al.*[38] Reactive radical anions are generated by ionising the reagent molecules, i.e. azulene and fluoranthene, through a negative chemical ionisation process. The precursor ions and radical anions are then transferred to the ion trap MS where ion-ion interaction takes place. During the

ion-ion interaction, electrons from the radical anions are transferred onto the precursor ions, which results in the formation of radical cations and leads to further fragmentation.

Hot ECD MS/MS is an interaction between precursor ions and more energetic electrons (~ 6-10 eV). In hot ECD, the chance of secondary fragmentation increases which results in extensive fragmentation, indicating the generation of w-ions.[92] Electrons with higher energy (~ 20 – 100 eV) are used in in electron induced dissociation (EID) which results in the most extensive protein/peptide fragmentation as all bonds in the backbone are possible to be cleaved.[116]

For negative precursor ions, electron detachment dissociation (EDD) with energetic electrons (~ 10-25 eV) is used and results in C_α-C backbone cleavage and generates a-, x-, c-, and z-ions.

1.1.4 Bruker 12 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer

In a Bruker 12 T FTICR-MS (see Figure 1.28), ions can be generated by a nESI, ESI, or MALDI source. Ions generated by ESI or nESI are then passed through the glass capillary with temperature ranged from 40ºC to 220ºC which helps to evaporate the solvent surrounding the analytes. The orthogonal design of the ion source and ion funnels allows the separation of neutral molecules from the ionised analytes because the electric field from ion funnels only affects the path of charged molecules, as a result only ions are transmitted to the later parts of the instrument. The dual ion funnels with a source octopole are used to focus the ions before passing through to the Qh front end which contains a quadrupole and a hexapole collision cell. The quadrupole in the Qh front end is used for *m/z* filtering if mass isolation is required. Ions are then passed to the hexapole for

accumulation and CAD/ETD fragmentation can be performed if necessary. The radical anions are generated in the nCI source for ETD fragmentation. Ions are then transferred through a hexapole (ion guide) and reach the infinity cell which is placed within a superconducting magnet (12 T).
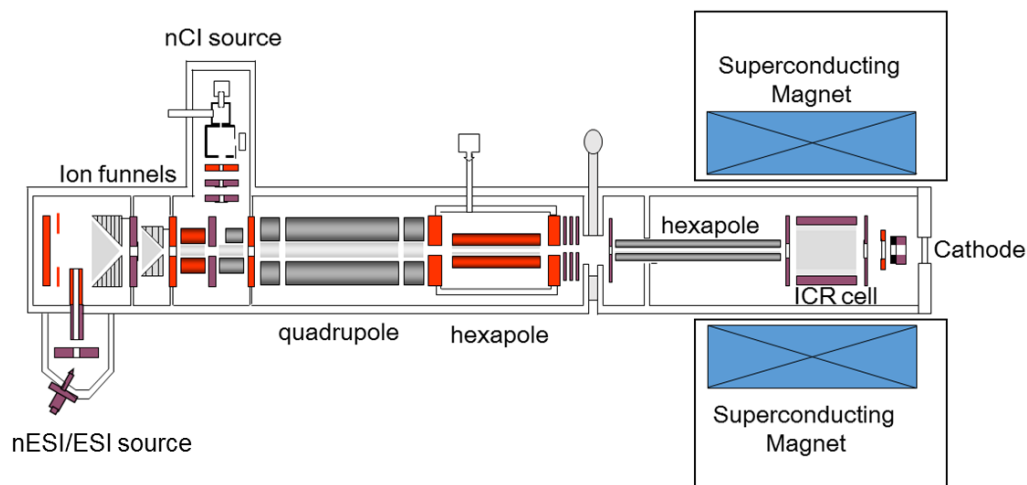
Figure 1.28. Diagram to show the instrument design of the Bruker 12 T solariX FTICR MS used in this work (Reproduced from the tutorial material provided by Bruker Daltonics, Bremen, Germany).

Ions in the infinity cell can be fragmented by IRMPD or ECD if necessary. The electron cathode is placed at the opposite side of the ion entrance in the ICR cell with lenses placed between the cathode and ICR cell to extract electrons for ECD MS/MS. IR laser is placed at the same side of the cathode. The on-axis laser beam (generated from a CO_2 laser) passes through the hollow cathode and reaches the ICR cell for IRMPD MS/MS.

The resolving power of the Bruker 12 T FTICR MS is around 600,000 at m/z 400 acquiring in broadband mode with the low mass detection at m/z 147 and four mega-word data points (32 bits). Mass accuracy is usually less than 0.5 ppm with internal mass calibration and less than three ppm with external mass calibration.

1.2 Applications of Mass Spectrometry

There are a wide range of applications that are based on MS analysis.[9, 21, 47, 81, 114, 117, 118] In this thesis, the applications of MS are mainly focused on amyloid proteins studies, including the underlying aggregation and inhibition mechanisms of amyloid proteins, the effects of deamidation on amyloid proteins aggregation and inhibition, as well as amyloid-related proteomics studies.

1.2.1 Introduction of Amyloid Proteins

Abnormal aggregation of proteins and peptides into mature amyloid fibrils is a hallmark in approximately 50 human disorder diseases, including Alzheimer's disease (AD), Parkinson's disease (PD), type II diabetes (T2D), and Huntington's diseases (HD) (see Table 1.1).[119, 120] The disorder diseases can be further classified into three different groups: 1) neurodegenerative conditions, in which aggregation happens in brain; 2) non-neuropathic localised amyloidosis, in which aggregation happens in a single organ/tissue rather than brain; and 3) nonneuropathic systemic amyloidosis, in which aggregations happens in several organs/tissues simultaneously.

Disease	Aggregation protein/peptide	Number of residues
Neurodegenerative diseases		
Alzheimer's disease	Amyloid β peptide	40/ 42
Amyotrophic lateral sclerosis	Superoxide dismutase 1	153
Dementia with Lewy bodies	α-synclein	140
Familial British dementia	Abri	23
Familial Danish dementia	Adan	23
Frontotemporal dementia with Parkinsonism	Tau	352-441
Hereditary dentatorubral-pallidoluysian atrophy	Atrophin-1 with polyQ expansion	1185
Huntington's disease	Huntingtin with polyQ expansion	3144
Parkinson's disease	α-synclein	140
Songiform encephalopathies	Prion protein	253
Spinal and bulbar muscular atrophy	Androgen receptor with polyQ expansion	919
Spinocerebellar ataxia 17	TATA box-binding protein with polyQ expansion	339
Spinocerebellar ataxias	Ataxins with polyQ expansion	816

(Table 1.1 continue)

Nonneuropathic systemic amyloidoses		
AA amyloidosis	Fragments of serum amyloid A protein	76-104
AL amyloidosis	Immunoglobulin light chains	90
ApoAI amyloidosis	N-terminal fragments of apolipoprotein AI	80-93
ApoAII amyloidosis	N-terminal fragments of apolipoprotein AII	98
ApoAIV amyloidosis	N-terminal fragment of apolipoprotein AIV	70
Familial amyloidotic polyneuropathy	Mutants of transthyretin	127
Familial Mediterranean fever	Serum amyloid A protein	76-104
Fibrinogen amyloidosis	Variants of fibrinogen α-chain	27-81
Finnish hereditary amyloidosis	Fragments of gelsolin mutants	71
Hemodialysis-related amyloidosis	β2-microglobulin	99
Icelandic hereditary cerebral amyloid angiopathy	Mutant of cystatin C	120
Lysozyme amyloidosis	Mutants of lysozyme	130
Senile systemic amyloidosis	Wild-type transthyretin	127

(Table 1.1 continue)

Nonneuropathic localized diseases		
Aortic medial amyloidosis	Medin	50
Atrial amyloidosis	Atrial natriuretic factor	28
Cataract	γ-Crystallins	variable
Corneal amylodosis associated with trichiasis	Lactoferrin	692
Cutaneous lichen amyloidosis	Keratins	variable
Hereditary cerebral haemorrhage with amyloidosis	Mutants of amyloid β peptide	40/ 42
Hereditary lattice corneal dystrophy	kerato-epithelin	50-200
Inclusion-body myositis	Amyloid β peptide	40/ 42
Injection-localized amyloidosis	Insulin	21+30
Medullary carcinoma of the thyroid	Calcitonin	32
Pituitary prolactinoma	Prolactin	199
Pulmonary alveolar proteinosis	Lung surfactant protein C	35
Type II diabetes	Amylin	37

Table 1.1. Table to summarise the most common human amyloid diseases.[120]

The overall structures of amyloid fibrils are often determined by transmission electron microscopy (TEM) and atomic force microscopy (AFM).[121, 122]   An amyloid fibril is composed of insoluble mature peptide fibrils originating from smaller, soluble monomer peptides, which are believed to refold into crossed β-strands perpendicular to the fibril axis.[123, 124]   In order to determine the molecular level details of amyloid fibrils, high-resolution structural studies using solid-state nuclear magnetic resonance (SSNMR) spectrometry have been used and the molecular level interactions between the aggregates of amyloid beta 1-40 ($A\beta_{1-40}$) were observed.[125, 126]

The formation of amyloid fibrils can be explained by a "nucleated growth" mechanism, which includes a lag phase, followed by an exponential growth phase, and eventually reach a final plateau (see Figure 1.29A).[120, 127]   In the lag phase, nuclei with the smallest amount of aggregates that are stable for the further growth rather than dissociation back into the monomer units are formed.[127]   Thus, the nuclei of the aggregates contain the highest Gibbs free energy in the aggregation process (see Figure 1.29B).   Once the nuclei are formed, the aggregation process changes to the growth phase, in which the association of monomer/ early oligomer units onto the nuclei is rapid.   When most of the monomer/ early oligomers units are associated with the nuclei, the aggregation process then reaches the final plateau stage.
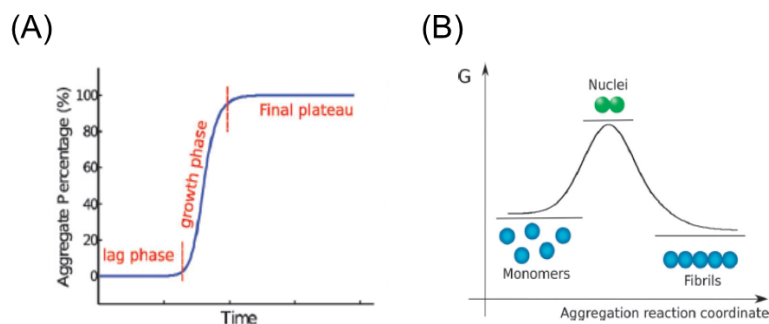
Figure 1.29.   Diagram to show (A) the various aggregation phases of amyloid

proteins and (B) the free energy of aggregates in various states during aggregation process (Reproduce from Arosio *et al.* 2015).[127]

The nucleated growth mechanism of amyloid fibril has been demonstrated by thioflavin T (ThT) fluorescence spectrometry experiments.   When ThT molecules bind to the mature β-sheet formed by the amyloid fibrils, a significant fluorescence emission is given at wavelength of 490 nm when the mixture is excited with a wavelength at 450 nm (see Figure 1.30B).[127, 128]   ThT molecule has been proposed to be stabilised at the surface side chain of the mature β-sheet fibril axis in order to generate fluorescence emission (see Figure 1.30C).[128]
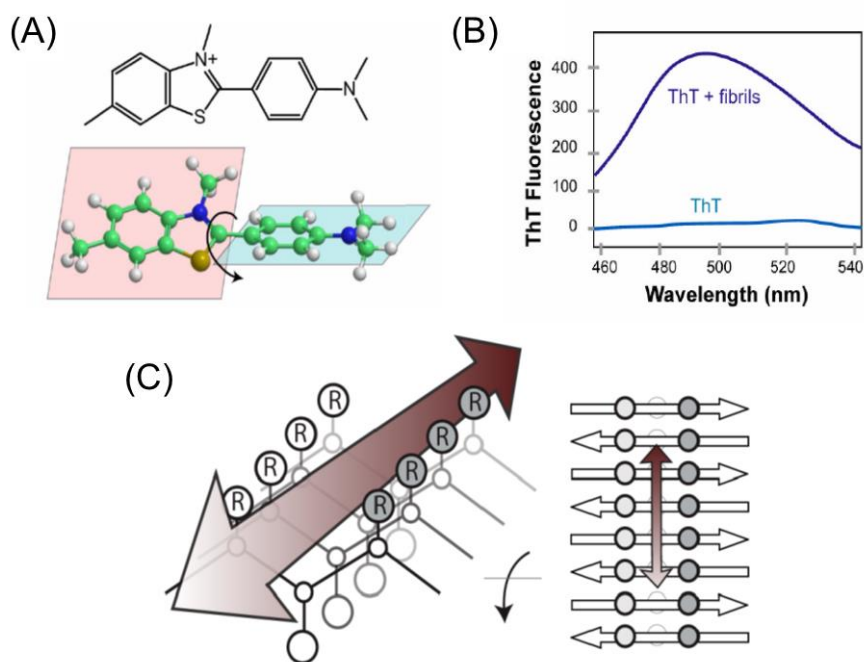
Figure 1.30.   Diagram showing (A) the molecular structure of ThT, (B) the emission wavelength (nm) of the mixture solution when ThT is bound onto amyloid fibrils, and (C) the proposed interaction mechanism between ThT molecule and mature amyloid fibrils (Reproduced from Biancalana *et al.* 2010).[128]

1.2.1.1 Amyloid Proteins Studies using Mass Spectrometry

The formation of amyloid fibrils is a hallmark in the human amyloid diseases; however, the early oligomers generated from amyloid proteins have been shown to be more toxic than the fibrils *in vitro* as the early oligomers can diffuse through cell membranes and lead to cell dysfunction.[129, 130] Experiments using TEM, AFM, SSNMR, and ThT fluorescence spectrometry have provided critical and valuable information on the structure and the time required for the formation of amyloid fibril, but the molecular level interactions between the early oligomers are difficult to observe by these methods as it is challenging to separate and detect a particular oligomer, i.e. dimer, from the monomer units and the orders of oligomers. Recently, MS has been demonstrated as an appropriate method to study the aggregation and inhibition mechanisms of amyloid proteins by observing the interaction between the early oligomers.[117, 131-137]

Aggregation mechanisms of amyloid proteins studied by MS. Previous studies have shown that early oligomers were able to be detected in micro-molar (μM) concentration in MS experiments; while different orders of oligomers were detected as each of the oligomer contains different *m/z* ratio.[117] CAD MS/MS was used to demonstrate the type of interaction between the early oligomers; in general, the bonding between the early oligomers is non-covalent.[135] Thus, CAD MS/MS cannot further demonstrate the interaction site between the early oligomers.

Recently, ion mobility-mass spectrometry (IM-MS) was developed to study the aggregation mechanism of amyloid proteins. A gradient of electric field is applied to further separate the early oligomers based on their secondary and tertiary structures, in which the compact oligomers are travelling faster than the open-structured oligomers.[138] The collision-cross section (CCS) of the oligomers

is obtained through the experiment, and computational modeling can simulate the interaction between the oligomers in order to match with the CCS provided in the experiment.[138, 139]   Thus, IM-MS is widely applied in the study of the amyloid proteins aggregation mechanisms.[117, 131, 139]

Inhibition mechanisms of amyloid proteins studied by MS.   Understanding the underlying inhibition mechanisms of the early oligomers of amyloid proteins is critical for drug development, thus MS has also been applied in determining the inhibition mechanisms of amyloid proteins.

Similar to the studies of aggregation mechanisms, MS, CAD MS/MS, and IM-MS experiments are used to observe the presence, the type, and the prediction of the interaction between the early oligomers and potential inhibitors respectively.[134, 136, 137]   MS was also promoted as a rapid screening method to predict the potential usage of the inhibitors by observing the change in the binding ratio between the early oligomers and the potential inhibitors.[137]

ECD MS/MS has also been applied in the previous studies to localise the interaction site between amyloid protein (α-synuclein) and potential inhibitor (spermine) down to the molecular level which is useful for further drug development.[37]

1.2.2. Introduction of Deamidation

Deamidation is a non-enzymatic PTM which has been proposed as an important step in many of time-related biological processes, including homeostatic control, protein turnover, as well as cell development and aging.[140-142] Deamidation occurs spontaneously in solution at asparagine (Asn) and glutamine (Gln) residues, and results in iso-aspartic acid/ aspartic acid and γ-glutamic acid/ glutamic acid formation respectively (see Figure 1.26), in which a negative charge

is induced at the deamidation site in a neutral pH solution (physiological buffer). [111, 112, 143] The deamidation rate of Asn is faster than Gln due to the shorter distance between the main chain amido group (-NH-) and the side chain amide group (-NH$_2$-).[110, 144] Upon deamidation of Asn, an additional methylene group (-CH$_2$-) is added into the backbone of iso-aspartic acid and γ-glutamic acid, which can result in a dramatic change in the tertiary structure, stability, folding, and function of proteins.[145, 146]

The deamidation rate of protein depends on the primary sequence, secondary and tertiary structure, temperature, pH, ionic strength, and buffer ions.[143, 147-150] Previous studies showed that the time frames for human proteins to be completely deamidated are ranged from 2 to >2000 days.[111, 140] Recently, the deamidation rates of proteins at pH 7 with temperature at 37°C can be predicted by the computational simulation based on the deamidation rate library generated by the pentapeptide models of Asn and Gln residues, and the reliable 3D structures of the proteins.[111, 140, 149, 150] The deamidation prediction accuracy is believed to be ~ 95% for exposed, flexible peptides.[150]

Deamidation has been shown to have significant effects in enhancing the aggregation rate of amyloid proteins. Previous studies have shown that deamidation can accelerate certain amyloid fibril formation including mutated Aβ (D7N and D23N) observed in Alzheimer's disease.[113, 151, 152] Aggregation of amylin associated with T2D,[153, 154] β$_2$-microglobulin from dialysis-associated amyloidosis disease, [5, 155] and αβ-crystallin from human eye lens have also been shown to be affected by deamidation.[156]

1.2.2.1 Deamidation Studies using Mass Spectrometry

Liquid chromatography, [157] NMR,[158] Raman spectroscopy,[159] and infrared

spectrometry[153] have been applied in studying protein modifications, but deamidation remains a challenge as a subtle change in the modification. MS is a powerful tool for examining deamidation in proteins as the deamidation reaction produces products with a mass shift of +0.984 Da (-NH$_2$ to -OH) at the deamidated residue (see Figure 1.26), which is readily detected by MS.[114, 160-162] Previous studies also suggested that deamidation can be directly quantified by measuring the peak area of the non-deamidated and deamidated peaks, thus the percentage of deamidation in the sample can be estimated.[163, 164]

MS is useful in detecting deamidation and provides deamidation percentage in the sample, tandem MS/MS can further locate the deamidation site in proteins, as well as differentiate the isomeric deamidation products.[112, 115, 160-164] CAD MS/MS helps to indicate the deamidation site with a mass shift (+0.984 Da) appears in the b and y ion fragments containing the deamidated residue, while normal fragments are obtained in the region that does not contain deamidated residue.[164] ECD MS/MS not only localises the deamidation site in protein, but also uses to differentiate the isomeric products of deamidation by the mass shift of +58 and -57 Da at the c and z ion fragments of iso-/γ- deamidation products respectively (see section 1.1.3.3). The peak intensity of iso-/γ- deamidation products can also compare with the peak intensity of aspartic acid/glutamic acid in order to determine the ratio of the isomeric products.[163]

1.2.3. Introduction to Proteomics

The terminology of "Proteomics" was first introduced in 1997 and proteome was introduced by Mark Wilkins in 1994, which is a combination of the words "protein" and "genome".[165] Proteomics is a proteome study that includes the large-scale of proteins generated by a living organism. In proteomics studies, the

structures and functions of proteins are analysed in detail.

In proteomics, MS is used to measure the mass of proteins and peptides which provides information on the identity, the chemical modifications, as well as the structures of the proteins. Proteomics research is involved significantly in searching for biomarkers for diseases,[166] quantifying the critical proteins and their relative PTMs changes,[167, 168] observing the structural changes of proteins,[58, 102] and monitoring the interactions of potential drugs with proteins in the living organism.[3, 169] These studies contribute to the prevention and treatment of diseases, such as cancer,[170] neurodegenerative diseases,[118] and bacteria diseases, etc.[171]

1.2.3.1. Proteomics Studies using Liquid Chromatography-Mass Spectrometry

Many thousands of proteins are involved in a single proteomics study. As a result there is no suitable analytical instrument to handle such a huge number of analytes in a single run other than MS.

Types of proteomics studies. In general, there are three different types of MS proteomics studies, including top-down proteomics, middle-down proteomics, and bottom-up proteomics.

Top-down proteomics. Top-down proteomics is the study of the whole size proteins without cutting the proteins into pieces. Top-down proteomics preserves the highest sequential and chemical modification information of the proteins as less treatments and chemicals are required for the samples, which helps to preserve the structure, PTMs, and full sequence of the proteins.[172, 173] Top-down proteomics, therefore, is suitable for the structural, as well as de novo sequencing studies of proteins.[58, 102, 105] However, MS requirements for top-down proteomics

are the highest due to the presence of a huge number of fragments in a single MS spectrum. Thus, ultra-high resolving power MS, i.e. FTICR MS or at least orbitrap MS, is required for top-down proteomics; as a result, top-down proteomics is not as common as bottom-up proteomics.

Bottom-up proteomics. Bottom-up proteomics is the study of peptides which are digested from proteins. Enzymatic digestion (i.e. chymotrypsin, pepsin, and trypsin digestion) and chemical digestion (i.e. cyanogen bromide) are the two common methods to cut proteins into peptides.[174, 175] In protein digestion, chemicals, like reducing and alkylating reagents, are commonly mixed with the samples in order to achieve a better result of digestion;[175, 176] however, artificial modifications are also introduced to the samples. Furthermore, incubation is normally required for the enzymatic digestion,[175] thus modifications that are highly correlated with temperature, i.e. oxidation and deamidation, are also induced into the sample.[177] At the same time, fragile chemical modifications on the proteins are easily removed during the digestion process due to the high temperature and long period of incubation.

Middle-down proteomics. Middle-down proteomics is the study of peptides that has not been fully digested from proteins, forming a middle ground between top-down and bottom-up proteomics. Middle-down proteomics minimises the digestion time of proteins from ~ 12-16 hours to ~ 2-4 hours which helps to reduce the loss of the fragile modifications, as well as decrease the artificial modifications induced in the samples.[178] Furthermore, proteins generated by middle-down proteomics contain more protonated sites which favours the fragmentation using MS/MS.[179] The proteins are that generated from middle-down proteomics are not fully digested, thus it is difficult to predict the number of missed cleavages in a single peptide/protein which results in challenging data analysis.

Chapter 1 Introduction

Separation methods applied in proteomics studies. Even though MS is capable of detecting various analytes simultaneously, it is still challenging to detect hundreds, or even thousands of proteins/peptides in a single scan. Gel and liquid chromatography (LC) separation are the two common methods applied in proteomics separation. Sodium dodcyl sulfate polyacrylamide (SDS-PAGE) gel electrophoresis is commonly applied in gel separation, however, it is only suitable for large biomolecules, i.e. proteins. LC, on the other hand, is suitable to apply to both protein and peptide separation. Herein, the various types of LC generally applied in proteomics separation are going to be discussed.

Stationary and mobile phase are the two keys requirements in LC. Stationary phase is a material (solid/liquid) that retained in the LC column which is used to adsorb the analyte molecules based on their hydrophobicity/charge/structure. Mobile phase is a material (liquid/gas) that passes through the stationary phase of LC and reduces the interaction between adsorbed analytes and stationary, which results in elution of the analyte.

Separation based on hydrophobicity. Different amino acids have different hydrophobicity value which was defined by Fauchere in 1983, and the hydrophobicity values of amino acid are ranged from -1 (most hydrophilic) to +2 (most hydrophobic). [180] Proteins and peptides are made up with different amino acids, which results in containing various hydrophobicity values.

Reversed-phase (RP) chromatography is the most common LC method to separate proteins and peptides in a mixture based on their hydrophobicity. The stationary phase of RP is made up of alkyl chains (C_4/ C_8/ C_{18}), different length of alkyl chains are attached onto the silica gel based on the separation purpose.[181] The short alkyl chain (i.e. C_4) is more suitable for the large analytes, i.e. proteins, separation; while the long alkyl chain (C_{18}) is more suitable for the small analytes,

i.e. peptide, separation.[181]   The alkyl chains in RP column provide a hydrophobic environment which adsorbs the hydrophobic analytes strongly to the stationary phase, thus the hydrophobic analytes are retained on the RP column longer than the hydrophilic analytes.   The stronger the interaction between the analytes and the stationary phase, the higher concentration of organic solvent (non-polar mobile phase) is required in order to elute the analytes out from the column.   Water is a polar solvent which can minimise the elution of analytes during loading and cleaning process.   The common non-polar organic solvent used for proteomics elution is acetonitrile as it contains a lower polarity than water; at the same time, it is miscible with water and does not damage the stationary phase.   RP is the most suitable LC column to co-operate directly with MS as the buffers used in RP are salt-free and do not aggregate or adhere to the biomolecule analytes, thus it is generally applied in MS proteomics studies.

Hydrophilic interaction liquid chromatography (HILIC) is an alternative LC method to separate proteins and peptides based on their hydrophobicity.   The stationary phases of HILIC is commonly made up of amino bonded phase, alkylamide, or zwitterionic sulfobetaine ligands attaching onto the silica gel.[182] In contrast to RP column, the stationary phase of HILIC provides a hydrophilic environment, thus hydrophilic analytes are adsorbed strongly onto the stationary phase, in which hydrophilic analytes are retained on the column longer than the hydrophobic analytes.   Acetonitrile is the most common loading buffer for HILIC column; while water/methanol is applied as the elution buffer due to its high polarity.[182]   Ammonium formate is commonly added into the buffer solutions in order to maintain the analytes at a single ionic form .[183]   HILIC has been demonstrated as a better separation on protein/peptide containing PTMs, i.e. phosphopeptide and glycopeptide.[182, 184]

Separation based on charge. Proteins and peptides contain different charge state in a particular pH solution, which can be used to enable separation by LC.

Strong cation exchange chromatography (SCX) and strong anion exchange chromatography (SAX) are two common ion-exchange chromatography applied in LC for the separation of positive and negative charged proteins/peptides respectively. The stationary phases of SCX and SAX normally contains a negatively charged functional group (sulfonate, $-SO_3^-$) or a positively charged functional group (quaternary amine, $-N(CH_3)_3^+$) respectively. The interactions between the ion-exchange chromatography columns and analytes are mainly based on electrostatic attraction as opposite charges are attracted to each other. The elution buffers of ion exchange columns are always contains high salt content, e.g. 350 mM NaCl,[185] to replace the interactions between the stationary phase functional groups and the charged analytes. Ion exchange columns have been demonstrated as an effective tools to enrich lots of modifications,[185, 186] however, extra clean-up of the samples is required prior to MS analysis.

1.2.3.2 Proteomics Studies using Two Dimensional Mass Spectrometry

LC-MS is a powerful technique for separating proteins and peptides prior to MS analysis, however, this method limits the study of fragile interaction between proteins or protein-ligand as the solvents applied in LC-MS may not preserve the weak bonding/structure of biomolecules. Thus, two dimensional mass spectrometry (2DMS) is an alternative method to proteomics study while preserving the weak interaction between proteins, and removes the requirement of LC prior to MS and MS/MS .

Principles of 2DMS. 2DMS is an MS based technique which modulates ions in-and-out of a fragmentation zone in an ion trapping device, such as an

FTICR cell.   The fragmentation technique (e.g. IRMPD) fragments the precursor ions in the FTICR cell, and 2DMS correlates the fragments with the corresponding precursor ion according to their modulated frequency.   In traditional MS/MS, precursor ion is first isolated by a mass analyser, i.e. quadrupole, according to its *m/z* ratio; in 2DMS, however, mass isolation is not required.

2DMS was first introduced by Pfandler *et al.* in 1987 and further developed and applied by Ross *et al.* in 1993 using an FTICR mass analyser.[187, 188]   In 2DMS, a modulating excitation or de-excitation pulse is used to alter the positions of ions inside the FTICR cell, ions that are modulated to overlap with the position of the fragmentation zone are dissociated with the applied radius-dependent in-cell MS/MS technique, i.e. IRMPD or ECD.   Bensimon *et al.* showed that by increasing the delay period (t_1) between the two excitation pulses (P_1 and P_2) one can modulate the ions inside a FTICR cell.[189]   In 2DMS, an excitation pulse (P_1) is first applied to all precursor ions so that the ions no longer stay in the centre of the FTICR cell.   Then, a delayed period (t_1) is applied.   Ions with a cyclotron frequency that is a multiple of t_1 with an extra half period are de-excited by the second pulse (P_2), the ions then move back to the centre of the cell; while ions with a cyclotron frequency that is an integer multiple of t_1 are further excited to a higher radius position in the FTICR cell (see Figure 1.31).   In IRMPD MS/MS, the fragmentation zone is at the centre of the cell, thus only ions which are de-excited back to the centre are fragmented, and no fragmentation is obtained for the ions that are outside the fragmentation zone.   In 2DMS, t_1 is increased proportional with the number of scans, and the phases of the ions that are excited by P_1 will change according, which results in modulating the radius of precursor ion as a function of t_1.   Since the fragment ions are modulated in the same way as the precursor ion, which can be used as an information to correlate back to the

precursor ion.
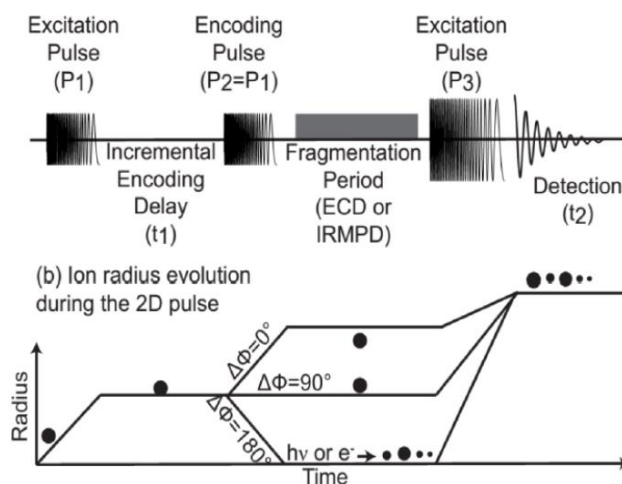
Figure 1.31. (A) A summarised diagram to show the pulse sequence applied in a 2DMS experiment. (B) The modulation of ions in 2DMS experiment (Reproduced from Van Agthoven *et al*).[190]

Applications of 2DMS. 2DMS is newly applied to proteomics studies, including both top-down and bottom-up proteomics. Previous studies showed that similar top-down and bottom-up fragments were obtained from both quadrupole precursor isolated MS/MS and 2DMS.[191] Peptides from digested collagen showed that 2DMS can differentiate and fragment the two precursor ions with a similar *m/z*, which, in general, cannot be isolated and fragmented individually by quadrupole mass isolation.[192] Peptide assignment from the 2DMS spectra of digested proteins, such as cytochrome C, calmodulin, and collagen, also showed that 2DMS has the potential for the complex proteomics studies without LC separation.[191-194]

1.3 Overview of the thesis

This thesis focuses on the studies of one of a highly amyloidogenic protein –

human islet amyloid polypeptide (hIAPP, also known as amylin).   In Chapter 2, MS is used to determine the aggregation pathway and deamidation sites of hIAPP; the rate of aggregation and the structures of amyloid fibrils formation are also determined by ThT fluorescence spectroscopy and TEM experiments.   In Chapter 3, MS is further used to determine the inhibition pathways of wild-type and deamidated hIAPP fibrillisation with several potential inhibitors; the inhibition effects of the potential inhibitors are measured by MS quantification, ThT fluorescence, and TEM experiments.   In Chapter 4, a newly developed analytical method – two dimensional mass spectrometry (2DMS) is evaluated for its potential to be applied in studying the fragile non-covalent interaction in a proteomics sample.   Conclusions of this thesis and future works are summarised in Chapter 5.

1.4 References

1.   Mann, M., Origins of mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* **2016**.

2.   Aebersold, R.; Mann, M., Mass-spectrometric exploration of proteome structure and function. *Nature* **2016,** *537* (7620), 347-355.

3.   Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.

4.   Shi, S. D.-H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W., Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. *Anal. Chem.* **2001,** *73* (1), 19-22.

5.   Soulby, A. J.; Heal, J. W.; Barrow, M. P.; Roemer, R. A.; O'Connor, P. B., Does deamidation cause protein unfolding? A top-down tandem mass spectrometry study. *Protein Sci.* **2015,** *24* (5), 850-860.

6.   Breuker, K.; Oh, H. B.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W., Detailed unfolding and folding of gaseous ubiquitin ions characterized by electron capture dissociation. *J. Am. Chem. Soc.* **2002,** *124* (22), 6407-6420.

7.   Thomson, J. J., *Rays of positive electricity and their application to chemical analyses*. Longmans, Green and Company: London, 1921; Vol. 1.

8.   De Hoffmann, E.; Stroobant, V., *Mass spectrometry: principles and applications*. John Wiley & Sons: London, 2007.

9.   Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* **2007,** *26* (1), 51-78.

10.   Gstaiger, M.; Aebersold, R., Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat. Rev. Genet.* **2009,** *10* (9), 617-627.

11.   Rodgers, R. P.; Marshall, A. G., *Petroleomics: Advanced characterization of petroleum-derived materials by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS)*. Springer: New York, 2007; p 63-93.

12.   Dempster, A., A new method of positive ray analysis. *Phys. Rev.* **1918,** *11* (4), 316.

13.   Dongre, V. G.; Karmuse, P. P.; Nimbalkar, M.; Singh, D.; Kumar, A., Application of GC–EI-MS for the identification and investigation of positional isomer in primaquine, an antimalarial drug. *J. Pharm. Biomed. Anal.* **2005,** *39* (1), 111-116.

14.   Ji, A. J.; Lawson, G. M.; Anderson, R.; Dale, L. C.; Croghan, I. T.; Hurt, R. D., A new gas chromatography – mass spectrometry method for simultaneous determination of total and free trans-3΄-hydroxycotinine and cotinine in the urine of subjects receiving transdermal nicotine. *Clin. Chem.* **1999,** *45* (1), 85-91.

15.   Harrison, A. G., *Chemical ionization mass spectrometry*. CRC press: New

York, 1992.

16. Munson, M. S.; Field, F.-H., Chemical ionization mass spectrometry. I. General introduction. *J. Am. Chem. Soc.* **1966,** *88* (12), 2621-2630.

17. Karas, M.; Bachmann, D.; Bahr, U. e.; Hillenkamp, F., Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int. J. Mass Spectrom. Ion Processes* **1987,** *78*, 53-68.

18. Karas, M.; Bachmann, D.; Hillenkamp, F., Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem.* **1985,** *57* (14), 2935-2939.

19. Lattimer, R. P.; Harris, R. E., Mass spectrometry for analysis of additives in polymers. *Mass Spectrom. Rev.* **1985,** *4* (3), 369-390.

20. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1988,** *2* (8), 151-153.

21. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **1997,** *69* (23), 4751-4760.

22. Griffin, T. J.; Smith, L. M., Single-nucleotide polymorphism analysis by MALDI–TOF mass spectrometry. *Trends Biotechnol.* **2000,** *18* (2), 77-84.

23. Knochenmuss, R., A quantitative model of ultraviolet matrix-assisted laser desorption/ionization including analyte ion generation. *Anal. Chem.* **2003,** *75* (10), 2199-2207.

24. Karas, M.; Krüger, R., Ion formation in MALDI: the cluster ionization mechanism. *Chem. Rev.* **2003,** *103* (2), 427-440.

25. Zenobi, R.; Knochenmuss, R., Ion formation in MALDI mass spectrometry. *Mass Spectrom. Rev.* **1998,** *17* (5), 337-366.

26. Mank, M.; Stahl, B.; Boehm, G., 2, 5-Dihydroxybenzoic acid butylamine and other ionic liquid matrixes for enhanced MALDI-MS analysis of biomolecules. *Anal. Chem.* **2004,** *76* (10), 2938-2950.

27. Woods, A.; Buchsbaum, J.; Worrall, T.; Berg, J.; Cotter, R., Matrix-assisted laser desorption/ionization of noncovalently bound compounds. *Anal. Chem.* **1995,** *67* (24), 4462-4465.

28. Dreisewerd, K., The desorption process in MALDI. *Chem. Rev.* **2003,** *103* (2), 395-426.

29. El-Aneed, A.; Cohen, A.; Banoub, J., Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Appl. Spectrosc. Rev.* **2009,** *44* (3), 210-230.

30. Yamashita, M.; Fenn, J. B., Electrospray ion source. Another variation on the

free-jet theme. *J. Phys. Chem.* **1984,** *88* (20), 4451-4459.

31. Loo, J. A.; Udseth, H. R.; Smith, R. D.; Futrell, J., Collisional effects on the charge distribution of ions from large molecules, formed by electrospray-ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **1988,** *2* (10), 207-210.

32. Meng, C.; Mann, M.; Fenn, J., Of protons or proteins. *Z. Phys. D: At., Mol. Clusters* **1988,** *10* (2), 361-368.

33. Banerjee, S.; Mazumdar, S., Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int. J. Anal. Chem.* **2012,** *2012*.

34. Gomez, A.; Tang, K., Charge and fission of droplets in electrostatic sprays. *Phys. Fluids* **1994,** *6* (1), 404-414.

35. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013,** *85* (1), 2-9.

36. Lomeli, S. H.; Yin, S.; Loo, R. R. O.; Loo, J. A., Increasing charge while preserving noncovalent protein complexes for ESI-MS. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (4), 593-596.

37. Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A., Top-down ESI-ECD-FT-ICR mass spectrometry localizes noncovalent protein-ligand binding sites. *J. Am. Chem. Soc.* **2006,** *128* (45), 14432-14433.

38. Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Nat. Acad. Sci. U.S.A.* **2004,** *101* (26), 9528-9533.

39. Yates, J. R., Mass spectrometry and the age of the proteome. *J. Mass Spectrom.* **1998,** *33* (1), 1-19.

40. Loo, J. A., Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **1997,** *16* (1), 1-23.

41. Chen, R.; Cheng, X.; Mitchell, D. W.; Hofstadler, S. A.; Wu, Q.; Rockwood, A. L.; Sherman, M. G.; Smith, R. D., Trapping, detection, and mass determination of coliphage T4 DNA ions by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **1995,** *67* (7), 1159-1163.

42. Perez Hurtado, P.; Lam, P. Y.; Kilgour, D.; Bristow, A.; McBride, E.; O'Connor, P. B., Use of high resolution mass spectrometry for analysis of polymeric excipients in drug delivery formulations. *Anal. Chem.* **2012,** *84* (20), 8579-8586.

43. Henriksen, T.; Juhler, R. K.; Svensmark, B.; Cech, N. B., The relative influences of acidity and polarity on responsiveness of small organic molecules to analysis with negative ion electrospray ionization mass spectrometry (ESI-MS). *J.*

Am. Soc. Mass. Spectrom. **2005,** *16* (4), 446-455.

44.   Wilm, M.; Mann, M., Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **1996,** *68* (1), 1-8.

45.   Mandal, M. K.; Yoshimura, K.; Saha, S.; Ninomiya, S.; Rahman, M. O.; Yu, Z.; Chen, L. C.; Shida, Y.; Takeda, S.; Nonami, H., Solid probe assisted nanoelectrospray ionization mass spectrometry for biological tissue diagnostics. *Analyst* **2012,** *137* (20), 4658-4661.

46.   Benesch, J. L.; Sobott, F.; Robinson, C. V., Thermal dissociation of multimeric protein complexes by using nanoelectrospray mass spectrometry. *Anal. Chem.* **2003,** *75* (10), 2208-2214.

47.   Spengler, B., De novo sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2004,** *15* (5), 703-714.

48.   Qian, K.; Robbins, W. K.; Hughey, C. A.; Cooper, H. J.; Rodgers, R. P.; Marshall, A. G., Resolution and identification of elemental compositions for more than 3000 crude acids in heavy petroleum by negative-ion microelectrospray high-field Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* **2001,** *15* (6), 1505-1511.

49.   Kujawinski, E. B.; Hatcher, P. G.; Freitas, M. A., High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: improvements and comparisons. *Anal. Chem.* **2002,** *74* (2), 413-419.

50.   Paul, W.; Steinwedel, H., Ein neues massenspektrometer ohne magnetfeld. *Z. Naturforsch., A: Phys. Sci.* **1953,** *8* (7), 448-450.

51.   Finnigan, R. E., Quadrupole mass spectrometers. *Anal. Chem.* **1994,** *66* (19), 969A-975A.

52.   Steel, C.; Henchman, M., Understanding the quadrupole mass filter through computer simulation. *J. Chem. Educ.* **1998,** *75* (8), 1049.

53.   Miller, P. E.; Denton, M. B., The quadrupole mass filter: basic operating concepts. *J. Chem. Educ.* **1986,** *63* (7), 617.

54.   March, R. E.; McMahon, A. W.; Londry, F. A.; Alfred, R. L.; Todd, J. F.; Vedel, F., Resonance excitation of ions stored in a quadrupole ion trap. Part 1. A simulation study. *Int. J. Mass Spectrom. Ion Processes* **1989,** *95* (2), 119-156.

55.   Hopfgartner, G.; Varesio, E.; Tschäppät, V.; Grivet, C.; Bourgogne, E.; Leuthold, L. A., Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. *J. Mass Spectrom.* **2004,** *39* (8), 845-855.

56.   Gobey, J.; Cole, M.; Janiszewski, J.; Covey, T.; Chau, T.; Kovarik, P.; Corr,

J., Characterization and performance of MALDI on a triple quadrupole mass spectrometer for analysis and quantification of small molecules. *Anal. Chem.* **2005,** *77* (17), 5643-5654.

57.   Patrie, S. M.; Robinson, D. E.; Meng, F.; Du, Y.; Kelleher, N. L., Strategies for automating top-down protein analysis with Q-FTICR MS. *Int. J. Mass spectrom.* **2004,** *234* (1), 175-184.

58.   Kellie, J. F.; Tran, J. C.; Lee, J. E.; Ahlf, D. R.; Thomas, H. M.; Ntai, I.; Catherman, A. D.; Durbin, K. R.; Zamdborg, L.; Vellaichamy, A., The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Mol. Biosyst.* **2010,** *6* (9), 1532-1539.

59.   Stephens, W., A Pulsed Mass Spectrometer with Time Disaersion. *Phys. Rev.* **1946,** *69*, 691.

60.   Wiley, W.; McLaren, I. H., Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* **1955,** *26* (12), 1150-1157.

61.   Opsal, R. B.; Owens, K. G.; Reilly, J. P., Resolution in the linear time-of-flight mass spectrometer. *Anal. Chem.* **1985,** *57* (9), 1884-1889.

62.   Lange, W.; Greifendorf, D.; Van Leyen, D.; Niehuis, E.; Benninghoven, A., Analytical applications of high-performance TOF-SIMS. In *Ion Formation from Organic Solids (IFOS III)*, Springer: Berlin, Germany 1986; pp 67-73.

63.   Önnerfjord, P.; Nilsson, J.; Wallman, L.; Laurell, T.; Marko-Varga, G., Picoliter sample preparation in MALDI-TOF MS using a micromachined silicon flow-through dispenser. *Anal. Chem.* **1998,** *70* (22), 4755-4760.

64.   Imrie, D.; Pentney, J.; Cottrell, J., A Faraday cup detector for high-mass ions in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1995,** *9* (13), 1293-1296.

65.   Moniatte, M.; Van der Goot, F.; Buckley, J.; Pattus, F.; Van Dorsselaer, A., Characterisation of the heptameric pore-forming complex of the Aeromonas toxin aerolysin using MALDI-TOF mass spectrometry. *FEBS Lett.* **1996,** *384* (3), 269-272.

66.   Mamyrin, B.; Karataev, V.; Shmikk, D.; Zagulin, V., The massreflect ron, a new non-magnetic time-of-flight mass spectrometer with high resolution. *Zh. Eksp. Teor. Fiz* **1973,** *64*, 82-89.

67.   Comisarow, M. B.; Marshall, A. G., Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.* **1974,** *25* (2), 282-283.

68.   Smith, D. F.; Kiss, A.; Leach, F. E.; Robinson, E. W.; Paša-Tolić, L.; Heeren, R. M., High mass accuracy and high mass resolving power FT-ICR secondary ion mass spectrometry for biological tissue imaging. *Anal. Bioanal. Chem.* **2013,** *405* (18), 6069-6076.

69. Marshall, A. G.; Hendrickson, C. L., Fourier transform ion cyclotron resonance detection: principles and experimental configurations. *Int. J. Mass spectrom.* **2002,** *215* (1-3), 59-75.

70. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom. Rev.* **1998,** *17* (1), 1-35.

71. Barrow, M. P.; Burkitt, W. I.; Derrick, P. J., Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. *Analyst* **2005,** *130* (1), 18-28.

72. Wang, M.; Marshall, A. G., Laboratory-frame and rotating-frame ion trajectories in ion cyclotron resonance mass spectrometry. *Int. J. Mass Spectrom. Ion Processes* **1990,** *100*, 323-346.

73. Comisarow, M. B.; Marshall, A. G., Frequency-sweep Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.* **1974,** *26* (4), 489-490.

74. Marshall, A. G.; Roe, D. C., Theory of Fourier transform ion cyclotron resonance mass spectroscopy: Response to frequency-sweep excitation. *J. Chem. Phys.* **1980,** *73* (4), 1581-1590.

75. Marshall, A. G.; Wang, T. C. L.; Ricca, T. L., Tailored excitation for Fourier transform ion cyclotron mass spectrometry. *J. Am. Chem. Soc.* **1985,** *107* (26), 7893-7897.

76. Comisarow, M. B., Signal modeling for ion cyclotron resonance. *J. Chem. Phys.* **1978,** *69* (9), 4097-4104.

77. Caravatti, P.; Allemann, M., The 'Infinity Cell': a New Trapped-ion Cell With Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991,** *26* (5), 514-518.

78. Beu, S. C.; Laude Jr, D. A., Open trapped ion cell geometries for Fourier transform ion cyclotron resonance mass spectrometry. *Int. J. Mass Spectrom. Ion Processes* **1992,** *112* (2-3), 215-230.

79. Marshall, A. G., Theoretical signal-to-noise ratio and mass resolution in Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **1979,** *51* (11), 1710-1714.

80. Scigelova, M.; Hornshaw, M.; Giannakopulos, A.; Makarov, A., Fourier transform mass spectrometry. *Mol. Cell. Proteomics* **2011,** *10* (7), M111. 009431.

81. Smith, D. F.; Podgorski, D. C.; Rodgers, R. P.; Blakney, G. T.; Hendrickson, C. L., 21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures. *Anal. Chem.* **2018**.

82. Wells, J. M.; McLuckey, S. A., Collision-induced dissociation (CID) of

peptides and proteins. In *Biol. Mass Spectrom.*, Burlingame, A. L., Ed. 2005; Vol. 402, pp 148-185.

83. Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, U. A.; Tsybin, Y. O., Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.* **2013,** *42* (12), 5014-5030.

84. Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W., Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* **2000,** *72* (3), 563-573.

85. Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W., Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal. Chem.* **1994,** *66* (18), 2809-2815.

86. Ly, T.; Julian, R. R., Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics. *Angew. Chem. Int. Ed.* **2009,** *48* (39), 7130-7137.

87. Zubarev, R. A., Electron-capture dissociation tandem mass spectrometry. *Curr. Opin. Biotechnol.* **2004,** *15* (1), 12-16.

88. Kalli, A.; Grigorean, G.; Håkansson, K., Electron induced dissociation of singly deprotonated peptides. *J. Am. Soc. Mass. Spectrom.* **2011,** *22* (12), 2209-2221.

89. Yang, J.; Mo, J.; Adamson, J. T.; Håkansson, K., Characterization of oligodeoxynucleotides by electron detachment dissociation Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2005,** *77* (6), 1876-1882.

90. Roepstorff, P.; Fohlman, J., Proposal For a Common Nomenclature For Sequence Ions In Mass-Spectra Of Peptides. *Biomed. Mass Spectrom.* **1984,** *11* (11), 601-601.

91. Jennings, K. R., Collision-induced decompositions of aromatic molecular ions. *Int J Mass Spectrom Ion Physics* **1968,** *1* (3), 227-235.

92. Cooper, H. J.; Håkansson, K.; Marshall, A. G., The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.* **2005,** *24* (2), 201-222.

93. Stefan, S. E.; Eyler, J. R., Differentiation of glucose-containing disaccharides by infrared multiple photon dissociation with a tunable CO_2 laser and Fourier transform ion cyclotron resonance mass spectrometry. *Int. J. Mass spectrom.* **2010,** *297* (1-3), 96-101.

94. Brodbelt, J. S.; Wilson, J. J., Infrared multiphoton dissociation in quadrupole ion traps. *Mass Spectrom. Rev.* **2009,** *28* (3), 390-424.

95. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W., Electron capture

dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **1998,** *120* (13), 3265-3266.

96.   Deguchi, K.; Ito, H.; Baba, T.; Hirabayashi, A.; Nakagawa, H.; Fumoto, M.; Hinou, H.; Nishimura, S. I., Structural analysis of O-glycopeptides employing negative-and positive-ion multi-stage mass spectra obtained by collision-induced and electron-capture dissociations in linear ion trap time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2007,** *21* (5), 691-698.

97.   Voinov, V. G.; Beckman, J. S.; Deinzer, M. L.; Barofsky, D. F., Electron-capture dissociation (ECD), collision-induced dissociation (CID) and ECD/CID in a linear radio-frequency-free magnetic cell. *Rapid Commun. Mass Spectrom.* **2009,** *23* (18), 3028-3030.

98.   Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W., Nonergodic and conformational control of the electron capture dissociation of protein cations. *Proc. Natl. Acad. Sci. USA* **2004,** *101* (39), 14011-14016.

99.   Sawicka, A.; Skurski, P.; Hudgins, R. R.; Simons, J., Model calculations relevant to disulfide bond cleavage via electron capture influenced by positively charged groups. *J. Phys. Chem. B* **2003,** *107* (48), 13505-13511.

100. Syrstad, E. A.; Tureček, F., Toward a general mechanism of electron capture dissociation. *J. Am. Soc. Mass. Spectrom.* **2005,** *16* (2), 208-224.

101. Mikhailov, V. A.; Cooper, H. J., Activated ion electron capture dissociation (AI ECD) of proteins: synchronization of infrared and electron irradiation with ion magnetron motion. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (5), 763-771.

102. Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.-H.; Begley, T. P.; McLafferty, F. W., Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. *J. Am. Chem. Soc.* **2002,** *124* (4), 672-678.

103. Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Zubarev, R. A., Proteomics-grade de novo sequencing approach. *J. Proteome Res.* **2005,** *4* (6), 2348-2354.

104. Ahmad, I.; Hoessli, D. C.; Walker-Nasir, E.; Choudhary, M. I.; Rafik, S. M.; Shakoori, A. R., Phosphorylation and glycosylation interplay: Protein modifications at hydroxy amino acids and prediction of signaling functions of the human β3 integrin family. *J. Cell. Biochem.* **2006,** *99* (3), 706-718.

105. Breuker, K.; Jin, M.; Han, X.; Jiang, H.; McLafferty, F. W., Top-down identification and characterization of biomolecules by mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2008,** *19* (8), 1045-1053.

106. Breuker, K.; McLafferty, F. W., Native electron capture dissociation for the structural characterization of noncovalent interactions in native cytochrome c. *Angew. Chem. Int. Ed.* **2003,** *42* (40), 4900-4904.

107. Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M., Probing the nature of noncovalent interactions by mass spectrometry. A study of protein− CoA ligand binding and assembly. *J. Am. Chem. Soc.* **1996,** *118* (36), 8646-8653.

108. Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D., Effect of reducing disulfide-containing proteins on electrospray ionization mass spectra. *Anal. Chem.* **1990,** *62* (7), 693-698.

109. Li, H. L.; O'Connor, P. B., Electron Capture Dissociation of Disulfide, Sulfur-Selenium, and Diselenide Bound Peptides. *J. Am. Soc. Mass. Spectrom.* **2012,** *23* (11), 2001-2010.

110. Robinson, A. B.; Rudd, C. J., Deamidation of glutaminyl and asparaginyl residues in peptides and proteins. *Curr. Top. Cell. Regul.* **1974,** *8* (0), 247-95.

111. Robinson, N. E.; Robinson, A. B., Molecular clocks. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (3), 944-949.

112. Cournoyer, J. J.; Pittman, J. L.; Ivleva, V. B.; Fallows, E.; Waskell, L.; Costello, C. E.; O'Connor, P. B., Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Sci.* **2005,** *14* (2), 452-463.

113. Sargaeva, N. P.; Lin, C.; O'Connor, P. B., Identification of Aspartic and Isoaspartic Acid Residues in Amyloid beta Peptides, Including A beta 1-42, Using Electron-Ion Reactions. *Anal. Chem.* **2009,** *81* (23), 9778-9786.

114. Cournoyer, J. J.; Lin, C.; O'Connor, P. B., Detecting deamidation products in proteins by electron capture dissociation. *Anal. Chem.* **2006,** *78* (4), 1264-1271.

115. Hurtado, P. P.; O'Connor, P. B., Differentiation of isomeric amino acid residues in proteins and peptides using mass spectrometry. *Mass Spectrom. Rev.* **2012,** *31* (6), 609-625.

116. Zubarev, R. A.; Yang, H., Multiple Soft Ionization of Gas-Phase Proteins and Swift Backbone Dissociation in Collisions with≤ 99 eV Electrons. *Angew. Chem. Int. Ed.* **2010,** *49* (8), 1439-1441.

117. Ashcroft, A. E., Mass Spectrometry and the Amyloid Problem-How Far Can We Go in the Gas Phase? *J. Am. Soc. Mass. Spectrom.* **2010,** *21* (7), 1087-1096.

118. Adav, S. S.; Qian, J. R.; Ang, Y. L.; Kalaria, R. N.; Lai, M. K. P.; Chen, C. P.; Sze, S. K., iTRAQ Quantitative Clinical Proteomics Revealed Role of Na+K+-ATPase and Its Correlation with Deamidation in Vascular Dementia. *J. Proteome Res.* **2014,** *13* (11), 4635-4646.

119. Ross, C. A.; Poirier, M. A., Protein aggregation and neurodegenerative disease. *Nat. Med.* **2004,** *10*, S10-S17.

120. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human

disease. In *Annu. Rev. Biochem*, Annual Reviews: Palo Alto, 2006; Vol. 75, pp 333-366.

121. Gras, S. L.; Waddington, L. J.; Goldie, K. N., Transmission electron microscopy of amyloid fibrils. In *Protein Folding, Misfolding, and Disease*, Springer: New York, 2011; pp 197-214.

122. Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W.-M.; Mattson, M. P.; Tycko, R., Self-propagating, molecular-level polymorphism in Alzheimer's ß-amyloid fibrils. *Science* **2005,** *307* (5707), 262-265.

123. Sipe, J. D.; Cohen, A. S., Review: History of the amyloid fibril. *J. Struct. Biol.* **2000,** *130* (2-3), 88-98.

124. Ross, C. A.; Poirier, M. A., Protein aggregation and neurodegenerative disease. *Nat. Med.* **2004,** *10*, S10.

125. Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R., A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. USA* **2002,** *99* (26), 16742-16747.

126. Antzutkin, O. N.; Balbach, J. J.; Leapman, R. D.; Rizzo, N. W.; Reed, J.; Tycko, R., Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of β-sheets in Alzheimer's β-amyloid fibrils. *Proc. Natl. Acad. Sci. USA* **2000,** *97* (24), 13045-13050.

127. Arosio, P.; Knowles, T. P.; Linse, S., On the lag phase in amyloid fibril formation. *PCCP* **2015,** *17* (12), 7606-7618.

128. Biancalana, M.; Koide, S., Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta, Proteins Proteomics* **2010,** *1804* (7), 1405-1412.

129. Haass, C.; Selkoe, D. J., Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. *Nat. Rev. Mol. Cell Biol.* **2007,** *8* (2), 101.

130. Kayed, R.; Sokolov, Y.; Edmonds, B.; McIntire, T. M.; Milton, S. C.; Hall, J. E.; Glabe, C. G., Permeabilization of lipid bilayers is a common conformation-dependent activity of soluble amyloid oligomers in protein misfolding diseases. *J. Biol. Chem.* **2004,** *279* (45), 46363-46366.

131. Bleiholder, C.; Do, T. D.; Wu, C.; Economou, N. J.; Bernstein, S. S.; Buratto, S. K.; Shea, J.-E.; Bowers, M. T., Ion Mobility Spectrometry Reveals the Mechanism of Amyloid Formation of A beta(25-35) and Its Modulation by Inhibitors at the Molecular Level: Epigallocatechin Gallate and Scyllo-inositol. *J. Am. Chem. Soc.* **2013,** *135* (45), 16926-16937.

132. Gessel, M. M.; Bernstein, S.; Kemper, M.; Teplow, D. B.; Bowers, M. T.,

Familial Alzheimer's Disease Mutations Differentially Alter Amyloid beta-Protein Oligomerization. *ACS Chem. Neurosci.* **2012,** *3* (11), 909-918.

133. Dupuis, N. F.; Wu, C.; Shea, J. E.; Bowers, M. T., Human Islet Amyloid Polypeptide Monomers Form Ordered beta-hairpins: A Possible Direct Amyloidogenic Precursor. *J. Am. Chem. Soc.* **2009,** *131* (51), 18283-18292.

134. Young, L. M.; Ashcroft, A. E.; Radford, S. E., Small molecule probes of protein aggregation. *Curr. Opin. Chem. Biol.* **2017,** *39*, 90-99.

135. Young, L. M.; Cao, P.; Raleigh, D. P.; Ashcroft, A. E.; Radford, S. E., Ion Mobility Spectrometry-Mass Spectrometry Defines the Oligomeric Intermediates in Amylin Amyloid Formation and the Mode of Action of Inhibitors. *J. Am. Chem. Soc.* **2014,** *136* (2), 660-670.

136. Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Ashcroft, A. E.; Radford, S. E., ESI-IMS-MS: A method for rapid analysis of protein aggregation and its inhibition by small molecules. *Methods* **2016,** *95*, 62-69.

137. Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E., Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nat. Chem.* **2015,** *7* (1), 73-81.

138. Ruotolo, B. T.; Benesch, J. L.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V., Ion mobility–mass spectrometry analysis of large protein complexes. *Nat. Protoc.* **2008,** *3* (7), 1139.

139. Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T., Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to beta-sheet in amyloid fibril formation. *Nat. Chem.* **2011,** *3* (2), 172-177.

140. Robinson, N. E.; Robinson, A. B., Deamidation of human proteins. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (22), 12409-12413.

141. Solstad, T.; Flatmark, T., Microheterogeneity of recombinant human phenylalanine hydroxylase as a result of nonenzymatic deamidations of labile amide containing amino acids. *The FEBS Journal* **2000,** *267* (20), 6302-6310.

142. Takemoto, L.; Boyle, D., Specific glutamine and asparagine residues of γ-S crystallin are resistant to in vivo deamidation. *J. Biol. Chem.* **2000,** *275* (34), 26109-26112.

143. Geiger, T.; Clarke, S., Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides - Succinimide-Linked Reactions that Contribute to Protein-Degradation. *J. Biol. Chem.* **1987,** *262* (2), 785-794.

144. Wright, H. T., Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. *Crit. Rev. Biochem. Mol. Biol.* **1990,** *26* (1), 1-52.

145. Mamula, M. J.; Gee, R. J.; Elliott, J. I.; Sette, A.; Southwood, S.; Jones, P. J.; Blier, P. R., Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J. Biol. Chem.* **1999,** *274* (32), 22321-22327.

146. Doyle, H. A.; Mamula, M. J., Post-translational protein modifications in antigen recognition and autoimmunity. *Trends Immunol.* **2001,** *22* (8), 443-449.

147. Robinson, A. B.; McKerrow, J. H.; Cary, P., Controlled deamidation of peptides and proteins: an experimental hazard and a possible biological timer. *Proc. Natl. Acad. Sci. USA* **1970,** *66* (3), 753-757.

148. Robinson, A. B.; Scotchler, J. W.; McKerrow, J. H., Rates of nonenzymic deamidation of glutaminyl and asparaginyl residues in pentapeptides. *J. Am. Chem. Soc.* **1973,** *95* (24), 8156-8159.

149. Robinson, N. E.; Robinson, A. B., Prediction of protein deamidation rates from primary and three-dimensional structure. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (8), 4367-4372.

150. Robinson, N. E., Protein deamidation. *Proc. Natl. Acad. Sci. USA* **2002,** *99* (8), 5283-5288.

151. Tomidokoro, Y.; Rostagno, A.; Neubert, T. A.; Lu, Y.; Rebeck, G. W.; Frangione, B.; Greenberg, S. M.; Ghiso, J., Iowa Variant of Familial Alzheimer's Disease Accumulation of Posttranslationally Modified A beta D23N in Parenchymal and Cerebrovascular Amyloid Deposits. *Am. J. Pathol.* **2010,** *176* (4), 1841-1854.

152. Shimizu, T.; Fukuda, H.; Murayama, S.; Izumiyama, N.; Shirasawa, T., Isoaspartate formation at position 23 of amyloid beta peptide enhanced fibril formation and deposited onto senile plaques and vascular amyloids in Alzheimer's disease. *J. Neurosci. Res.* **2002,** *70* (3), 451-461.

153. Dunkelberger, E. B.; Buchanan, L. E.; Marek, P.; Cao, P.; Raleigh, D. P.; Zanni, M. T., Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. *J. Am. Chem. Soc.* **2012,** *134* (30), 12658-12667.

154. Nilsson, M. R.; Driscoll, M.; Raleigh, D. P., Low levels of asparagine deamidation can have a dramatic effect on aggregation of amyloidogenic peptides: Implications for the study of amyloid formation. *Protein Sci.* **2002,** *11* (2), 342-349.

155. Kad, N. M.; Thomson, N. H.; Smith, D. P.; Smith, D. A.; Radford, S. E., beta(2)-microglobulin and its deamidated variant, N17D form amyloid fibrils with a range of morphologies in vitro. *J. Mol. Biol.* **2001,** *313* (3), 559-571.

156. Gupta, R.; Srivastava, O. P., Effect of deamidation of asparagine 146 on functional and structural properties of human lens alpha B-crystallin. *Invest. Ophthalmol. Vis. Sci.* **2004,** *45* (1), 206-214.

157. Karlsson, G.; Gellerfors, P.; Persson, A.; Noren, B.; Edlund, P. O.; Sandberg, C.; Birnbaum, S., Separation of oxidized and deamidated human growth hormone variants by isocratic reversed-phase high-performance liquid chromatography. *J. Chromatogr. A* **1999,** *855* (1), 147-155.

158. Chazin, W. J.; Kordel, J.; Thulin, E.; Hofmann, T.; Drakenberg, T.; Forsen, S., Identification of an Isoaspartyl Linkage formed upon Deamidation of Bovine Calbindin-D9K and Strcutural Characterization by 2D H-1-NMR. *J. Biochem.* **1989,** *28* (21), 8646-8653.

159. Wong, H.-W.; Choi, S.-M.; Phillips, D. L.; Ma, C.-Y., Raman spectroscopic study of deamidated food proteins. *Food Chem.* **2009,** *113* (2), 363-370.

160. Chan, W. Y. K.; Chan, T. D.; O'Connor, P. B., Electron transfer dissociation with supplemental activation to differentiate aspartic and isoaspartic residues in doubly charged peptide cations. *J. Am. Soc. Mass. Spectrom.* **2010,** *21* (6), 1012-1015.

161. Hurtado, P. P.; O'Connor, P. B., Deamidation of Collagen. *Anal. Chem.* **2012,** *84* (6), 3017-3025.

162. Li, X.; Yu, X.; Costello, C. E.; Lin, C.; O'Connor, P. B., Top-Down Study of beta(2)-Microglobulin Deamidation. *Anal. Chem.* **2012,** *84* (14), 6150-6157.

163. Cournoyer, J. J.; Lin, C.; Bowman, M. J.; O'Connor, P. B., Quantitating the relative abundance of isoaspartyl residues in deamidated proteins by electron capture dissociation. *J. Am. Soc. Mass. Spectrom.* **2007,** *18* (1), 48-56.

164. Yang, H.; Zubarev, R. A., Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. *Electrophoresis* **2010,** *31* (11), 1764-1772.

165. Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I., Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **1995,** *16* (1), 1090-1094.

166. McDonald, W. H.; Yates, J. R., Shotgun proteomics and biomarker discovery. *Dis. Markers* **2002,** *18* (2), 99-105.

167. Bennett, E. J.; Shaler, T. A.; Woodman, B.; Ryu, K.-Y.; Zaitseva, T. S.; Becker, C. H.; Bates, G. P.; Schulman, H.; Kopito, R. R., Global changes to the ubiquitin system in Huntington's disease. *Nature* **2007,** *448* (7154), 704.

168. Gillette, M. A.; Carr, S. A., Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. *Nat. Methods* **2013,** *10* (1), 28.

169. Wootton, C. A.; Lam, Y. P.; Willetts, M.; van Agthoven, M. A.; Barrow, M. P.; Sadler, P. J.; Peter, B., Automatic assignment of metal-containing peptides in proteomic LC-MS and MS/MS data sets. *Analyst* **2017,** *142* (11), 2029-2037.

170. Zhang, P.; Chiu, C. K.; Huang, H.; Lam, Y. P.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; Sadler, P. J., Organoiridium photosensitizers induce specific oxidative attack on proteins within cancer cells. *Angew. Chem. Int. Ed.* **2017,** *56* (47), 14898-14902.

171. Gravett, M. G.; Novy, M. J.; Rosenfeld, R. G.; Reddy, A. P.; Jacob, T.; Turner, M.; McCormack, A.; Lapidus, J. A.; Hitti, J.; Eschenbach, D. A., Diagnosis of intra-amniotic infection by proteomic profiling and identification of novel biomarkers. *JAMA* **2004,** *292* (4), 462-469.

172. Siuti, N.; Kelleher, N. L., Decoding protein modifications using top-down mass spectrometry. *Nat. Methods* **2007,** *4* (10), 817.

173. Kelleher, N. L., Peer Reviewed: Top-Down Proteomics. *Anal. Chem.* **2004,** *76* (11), 196 A-203 A.

174. Switzar, L.; Giera, M.; Niessen, W. M., Protein digestion: an overview of the available techniques and recent developments. *J. Proteome Res.* **2013,** *12* (3), 1067-1077.

175. Gundry, R. L.; White, M. Y.; Murray, C. I.; Kane, L. A.; Fu, Q.; Stanley, B. A.; Van Eyk, J. E., Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow. *Curr. Protoc. Mol. Biol.* **2009**, 10.25. 1-10.25. 23.

176. Suttapitugsakul, S.; Xiao, H.; Smeekens, J.; Wu, R., Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics. *Mol. Biosyst.* **2017,** *13* (12), 2574-2582.

177. Formolo, T.; Heckert, A.; Phinney, K. W., Analysis of deamidation artifacts induced by microwave-assisted tryptic digestion of a monoclonal antibody. *Anal. Bioanal. Chem.* **2014,** *406* (26), 6587-6598.

178. Sidoli, S.; Schwämmle, V.; Ruminowicz, C.; Hansen, T. A.; Wu, X.; Helin, K.; Jensen, O. N., Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *Proteomics* **2014,** *14* (19), 2200-2211.

179. Cristobal, A.; Marino, F.; Post, H.; van den Toorn, H. W.; Mohammed, S.; Heck, A. J., Toward an Optimized Workflow for Middle-Down Proteomics. *Anal. Chem.* **2017,** *89* (6), 3318-3325.

180. Fauchere, J.-L.; Pliska, V., Hydrophobic parameters pi of amino-acid side chains from the partitioning of N-acetyl-amino-acid amides. *Eur. J. Med. Chem* **1983,** *18* (3), 369-375.

181. Catherman, A. D.; Skinner, O. S.; Kelleher, N. L., Top down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* **2014,** *445* (4), 683-693.

182. Boersema, P. J.; Mohammed, S.; Heck, A. J. R., Hydrophilic interaction liquid

chromatography (HILIC) in proteomics. *Anal. Bioanal. Chem.* **2008,** *391* (1), 151-159.

183. Buszewski, B.; Noga, S., Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Anal. Bioanal. Chem.* **2012,** *402* (1), 231-247.

184. Ahn, Y. H.; Kim, J. Y.; Yoo, J. S., Quantitative mass spectrometric analysis of glycoproteins combined with enrichment methods. *Mass Spectrom. Rev.* **2015,** *34* (2), 148-165.

185. Villén, J.; Gygi, S. P., The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat. Protoc.* **2008,** *3* (10), 1630.

186. Fíla, J.; Honys, D., Enrichment techniques employed in phosphoproteomics. *Amino Acids* **2012,** *43* (3), 1025-1047.

187. Pfändler, P.; Bodenhausen, G.; Rapin, J.; Houriet, R.; Gäumann, T., Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *Chem. Phys. Lett.* **1987,** *138* (2-3), 195-200.

188. Ross III, C. W.; Guan, S.; Grosshans, P. B.; Ricca, T. L.; Marshall, A. G., Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry/mass spectrometry with stored-waveform ion radius modulation. *J. Am. Chem. Soc.* **1993,** *115* (17), 7854-7861.

189. Bensimon, M.; Zhao, G.; Gäumann, T., A method to generate phase continuity in two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *Chem. Phys. Lett.* **1989,** *157* (1-2), 97-100.

190. van Agthoven, M. A.; Barrow, M. P.; Chiron, L.; Coutouly, M.-A.; Kilgour, D.; Wootton, C. A.; Wei, J.; Soulby, A.; Delsuc, M.-A.; Rolando, C., Differentiating fragmentation pathways of cholesterol by two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (12), 2105-2114.

191. Floris, F.; van Agthoven, M. A.; Chiron, L.; Wootton, C. A.; Lam, P. Y. Y.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Bottom-Up Two-Dimensional Electron-Capture Dissociation Mass Spectrometry of Calmodulin. *J. Am. Soc. Mass. Spectrom.* **2018,** *29* (1), 207-210.

192. Simon, H.; van Agthoven, M.; Lam, P.; Floris, F.; Chiron, L.; Delsuc, M.-A.; Rolando, C.; Barrow, M.; O'Connor, P., Uncoiling collagen: a multidimensional mass spectrometry study. *Analyst* **2016,** *141* (1), 157-165.

193. Maria A. van Agthoven, C. A. W., Lionel Chiron, Marie-Aude Coutouly, Andrew Soulby, Juan Wei, Mark. P. Barrow, Marc-Andre Delsuc, Christian Rolando, Peter B. O'Connor, Two-Dimensional Mass Spectrometry for Proteomics, a Comparative Study with Cytochrome c. *J. Anal. Chem.* **2016.**

194. Floris, F.; van Agthoven, M.; Chiron, L.; Soulby, A. J.; Wootton, C. A.; Lam, Y. P.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., 2D FT-ICR MS of calmodulin: a top-down and bottom-up approach. *J. Am. Soc. Mass. Spectrom.* **2016,** *27* (9), 1531-1538.

Chapter 2 Exploring the Aggregation and Deamidation Mechanisms of Human Islet Amyloid Polypeptide Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

This chapter demonstrates the applications of Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) together with fluorescence spectroscopy and transmission electron microscopy (TEM) to study the aggregation and deamidation mechanisms of a highly amyloidogenic protein − human islet amyloid protein (hIAPP).

The results presented in this chapter were predominantly carried out by the thesis Author, aside from acquisition of transmission electron microscopy (TEM) for which data were obtained by Mr. Ian-Hands Portman from the Department of Life Science, University of Warwick.

Two manuscripts entitled "Exploring the Aggregation Mechanisms of Human Islet Amyloid Polypeptide Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry" and "Human Islet Amyloid Polypeptide as a Molecular Clock: Does Deamidation cause Dimerization?" by Yuko P. Y. Lam, Christopher A. Wootton, Ian Hands-Portman; Juan Wei, Cookson K. C. Chiu, I. Romero-Canelon, Frederik Lermyte, Mark P. Barrow, and Peter B. O'Connor, have been prepared for submission to Analytical Chemistry based on the results presented in this chapter.

Chapter 2 – Aggregation and Deamidation Mechanisms of hIAPP

2.1 Abstract

Human islet amyloid polypeptide (hIAPP) aggregation is a hallmark in type II diabetes (T2D) which results in the death of pancreatic beta-cells and the failure of islet cell transplants. The aggregation mechanism of hIAPP has been studied for more than ten years, but the aggregation sites have not yet been conclusively determined. Herein, ultrahigh resolution mass spectrometry (MS) has been applied in monitoring the formation of early oligomers, especially the dimers and trimers of hIAPP. Tandem MS using electron-based dissociation reveals the most important aggregation sites of hIAPP to be the region near Gly-33 and Ser-34 residues, which contributes to the understanding of the aggregation mechanism. Furthermore, the results from MS, fluorescence spectroscopy, and transmission electron microscopy (TEM) clearly show that deamidation is not essential for the early oligomer aggregation of hIAPP in an aqueous solution since the rate of deamidation is much slower than the rate of aggregation. However, the rate of aggregation is accelerated by seeding deamidated hIAPP into solutions before incubation. The Asn-21, Asn-22, and Asn-35 amino acid residues in the incubated hIAPP solutions were found to be deamidated within seven days. Residues Asn-21, Asn-22, and Asn-35 all showed similar degrees of deamidation in the incubated solutions, while the degree of deamidation found in the fibril pellet is higher than in the incubated solution. Additionally, aggregation of mutant hIAPP with iso-Asp residue mutations at possible deamidation sites showed very different fibril formation behaviour compared to the wild-type hIAPP, which may help to explain the deamidation-induced acceleration of hIAPP aggregation.

2.2 Introduction

Amyloid fibril formation is a hallmark in a range of human diseases, including Alzheimer's disease,[1] Parkinson's disease,[2] and type II diabetes (T2D).[3] An amyloid fibril is composed of insoluble, mature peptide fibrils originating from smaller, soluble monomer peptides, which are believed to refold into crossed β-strands perpendicular to the fibril axis.[4, 5]   The overall structures of amyloid fibrils have been studied in detail for more than a half-century;[6-8] the molecular details of aggregation, however, are still limited by the difficulties in solubilising, separating, and identifying the aggregated biomolecules.

Human islet amyloid polypeptide (hIAPP; also referred to as amylin) is a 37-residue hormone peptide co-secreted with insulin and involved in regulating blood glucose levels.[9, 10]   The concentration of hIAPP is around 1% - 2% relative to the level of insulin in secretory granules.[11]   hIAPP contains an intramolecular disulfide bond between Cys-2 and Cys-7, as well as, an amidated C-terminus (see Figure 2.1A and 2.1B),[12] and is an intrinsically disordered protein with a low level of persistent helical structure between residues Asn-3 and Leu-27 in solution.[13] Early oligomers of hIAPP have been shown to be a key factor in causing the decline of pancreatic β-cell mass and the failure of islet cell transplants used for the treatment of T2D.[14, 15]   In the past decade, electron microscopy, circular dichroism (CD), and computational modelling simulations have been the common techniques for studying hIAPP aggregates, and a solid β-sheet structure was observed in the mature hIAPP fibrils.[16-19]   However, the aggregation mechanism remains poorly understood.[20]

Computational docking simulations with residue replacement techniques have been applied to explain the aggregation mechanism of hIAPP based on its molecular structure.[17]   Computational models suggested Phe-15, in the α-helix

region of hIAPP, was a critical residue for the aggregation since the rate of fibril formation was significantly reduced in the mutants where the Phe-15 residue was replaced *in silico* (F15D and F15K mutant hIAPP). Nuclear magnetic resonance (NMR) spectroscopy experiments led to the interpretation that the N-terminal segment of hIAPP (residues 1 – 17) was essential for the formation of amyloid fibrils.[21] These models contrast with the hypothesis that the region of hIAPP between residues 20-29 contains amyloidogenic tendency, which was demonstrated by comparing the rate of fibril formation between the rat IAPP (rIAPP) and hIAPP.[16] Rat and human IAPP sequences differ by six amino acid residues (of the 37 present), five of which are located within residues 20-29; however, experimental data have shown rIAPP did not aggregate over a 5-day incubation period while hIAPP was shown to aggregate readily within 20 hours.[16, 22] This hypothesis has been further studied on multiple occasions using short sections of hIAPP to identify the active aggregation region, where several studies using peptides containing residues 20-29 were shown to aggregate effectively,[22-29] indicating the active aggregation region could be within this range. Following this, solid state NMR also revealed that hIAPP units stack to form β-strands and then further aggregate into layers of β-sheets (see Figure 2.1C).[12] Ion mobility mass spectrometry (IM-MS) was recently used for the study of hIAPP aggregation.[4, 30-34] The structures of monomer and oligomers were predicted by molecular dynamics modeling based on the collision cross-sectional area obtained from IM-MS.[33, 35, 36] β-strand interaction was determined as the most stable and favourable binding motif in the hIAPP dimer.[32] A fragment of hIAPP (residues 28-33) was proposed by Bleiholder *et al.* as an amyloidogenic region via a 'steric zipper' assembly pathway using IM-MS.[4] Early oligomers of hIAPP are toxic to the β-cells;[37] however, the conflicting reports on the binding motif hinder drug

development.   Improved understanding of the binding motif of hIAPP is therefore required.
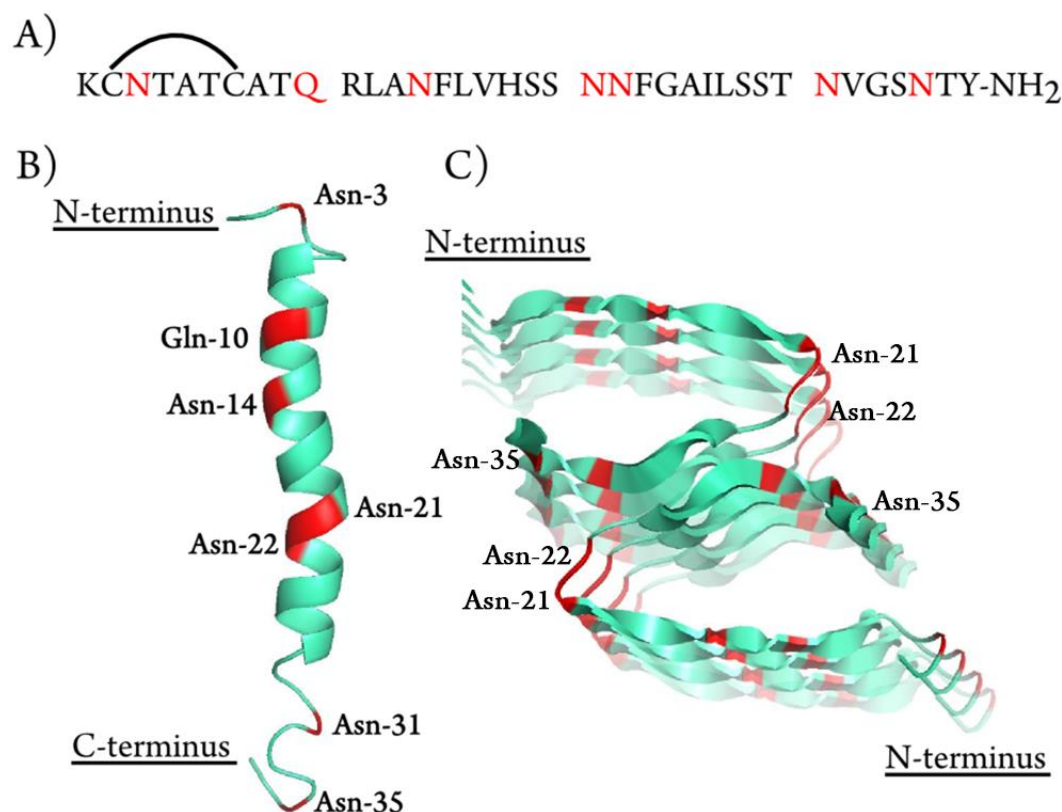
Figure 2.1.   (A) The amino acid sequence of hIAPP.   An intramolecular disulfide bond is located between Cys-2 and Cys-7, and hIAPP has an amidated C-terminus.   (B) The structure of hIAPP monomer (PDB 2KB8) as reported by Patil *et al*.[38] using ^{1}H NMR, which contains an α-helical structure from residues 5-28.   (C) The oligomer structure of hIAPP as demonstrated by Luca *et al*.[12] indicates each monomer of hIAPP interacts at the β-strand and forms β-sheet layers. The potential deamidation sites are coloured in red.

Deamidation is a non-enzymatic post-translational modification (PTM) which contributes to protein ageing.[39, 40]   Deamidation occurs spontaneously in solution at asparagine (Asn) and glutamine (Gln) residues, and results in iso-aspartic acid/aspartic acid and γ-glutamic acid/ glutamic acid formation

respectively.[41-43]   The deamidation rate of Asn is faster than Gln due to the shorter distance between the main chain amido group (-NH-) and the side chain amide group (-NH$_2$-).[44, 45]   With an additional methylene group (-CH$_2$-) into the backbone of iso-aspartic acid and γ-glutamic acid; the tertiary structure, stability, folding, and function of proteins can be dramatically altered by deamidation.[46, 47] Previous studies showed that deamidation can accelerate certain amyloid fibril formation, including mutated amyloid-β (D7N and D23N) observed in Alzheimer's disease.[48-50]   Aggregation of β$_2$-microglobulin from dialysis-associated amyloidosis disease,[51, 52] and αB-crystallin from human eye lens have also been shown to be affected by deamidation.[53]

hIAPP is an amyloid protein which contains six Asn and one Gln potential deamidation sites (see Figure 2.1).   A previous study from Nilsson *et al.* used a variant of hIAPP segment, residues 20-29 (SNNF<u>P</u>AILSS) which is non-amyloidogenic and showed that addition of less than 5% of deamidated monomer was required to accelerate the fibril formation.[54]   Results from Dunkelberger *et al.* also showed that deamidation of hIAPP disrupted the β-sheet structure around Leu-27, resulting in a morphology change of hIAPP fibrils, along with accelerated fibrils formation.[55]   Even though deamidation shows a significant impact on fibril formation, the effects on the deamidation rate, the deamidation site(s), and the effects of isomeric deamidation products of hIAPP have not yet been fully addressed.

Herein, Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) is applied in monitoring the aggregation and deamidation of hIAPP. Electron capture dissociation tandem MS (ECD MS/MS) is used for determining the aggregation sites between the monomer subunits within the dimer and the trimer of hIAPP and localising the binding area to the amino acid residue level

without disrupting remaining non-covalent bonds.[56-59] The concentration of soluble hIAPP and the deamidation rate are recorded periodically for 28 days in order to study the correlation between aggregation and deamidation in an aqueous solution. The deamidation sites and the extent of deamidation of hIAPP are further determined by collisionally activated dissociation (CAD) MS/MS. Transmission electron microscopy (TEM) and Thioflavin T (ThT) fluorescence spectroscopy are also employed to monitor the rate of fibril formation. The experimental results reveal the aggregation motif of hIAPP, the effects of deamidation on hIAPP aggregation, and the rate of deamidation of hIAPP in aqueous solution.

2.3 Experimental Section

Sample preparation for early oligomers in wild-type hIAPP, the synthetic mutant hIAPPs, the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide, and the mixture of hIAPP & ^{30}TNVGSNTY^{37}-NH$_2$. Wild-type hIAPP lyophilised powder (D2162 with purity >97%; Sigma Aldrich Company Ltd, Dorset, England; see Supplementary Information, Figure S2.1 – S2.2), the synthetic mutant hIAPPs with purity >95% (Pepscan Company Ltd, The Netherlands; see Supplementary Information, Figure S2.3), and the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide with purity >95% (Eurogentec, England) were initially dissolved in 100% dimethyl sulfoxide (DMSO) at a concentration of 500 µM for storage (at -80°C) in order to minimise the aggregation rate, as used by Abedini *et al.*.[60] The 500 µM stock solutions were further diluted with Milli-Q (Direct-Q® 3 UV System, Millipore Corporation, US) H$_2$O (~pH 7.5) to 10 µM solutions, and the final concentrations of DMSO in each solution were 2% (v/v).

Wild-type hIAPP lyophilised powder was also dissolved in Milli-Q H$_2$O at a concentration of 500 µM and further diluted into 10 µM solution. The mixture of hIAPP & ^{30}TNVGSNTY^{37}-NH$_2$ was prepared by mixing 10 µM hIAPP (in H$_2$O) with 10 µM ^{30}TNVGSNTY^{37}-NH$_2$ (in 100% DMSO) in 1:1 ratio in aqueous solution, and the final concentration of DMSO in hIAPP aqueous solutions was 2% (v/v).

Sample preparation for deamidated hIAPP and fibrils. The 10 µM hIAPP aqueous solution (2000 µL) was incubated at 37°C for 28 days; 50 µL solution was collected each day. Each collected aliquot solution was then centrifuged at 14,000 rpm for one hour to separate the soluble hIAPP (supernatant) from the insoluble hIAPP fibril (pellet). The supernatant solution containing soluble hIAPP was then diluted 20-fold with 49.5:49.5:1 water/acetonitrile/formic

acid prior to MS analysis. hIAPP fibrillary pellets from 7-, 14-, 21-, and 28-day solutions were rinsed with Milli-Q H_2O three times and then re-dissolved with 20 µL of 47.5:47.5:5 water/acetonitrile/formic acid, and sonicated in water bath at 37 $^{\circ}$C for one hour. Re-dissolved samples were further diluted with 80 µL of 50:50 water/acetonitrile. The final concentration of formic acid in solutions was 1% (v/v).

Sample preparation for seeding mutant peptides. Wild-type hIAPP lyophilised powder and synthetic mutant hIAPPs were dissolved in Milli-Q H_2O to a concentration of 500 µM. The seeding experiments were performed by mixing unmodified hIAPP stock solution with 5%, 10%, 25%, or 50% mutant $((D)_3hIAPP)$ or $((isoD)_3hIAPP)$ solutions. Samples were then diluted to a final concentration of 10 µM wild-type hIAPP plus mutant $((D)_3hIAPP)$ or $((isoD)_3hIAPP)$, which were incubated for one week at 37°C. The supernatants and fibrillary pellets were separated and prepared as mentioned above.

FTICR MS analysis. Mass spectra were acquired on a 12 tesla solariX FTICR MS (Bruker Daltonik GmbH, Bremen, Germany). All samples were analysed in positive ionisation mode. For the early oligomers study, nano-electrospray ionisation (nESI) with a capillary voltage of 0.6 – 1 kV was applied and the source temperature was set to 80 $^{\circ}$C. The nESI glass capillaries were purchased from World Precision Instruments and pulled on a Sutter P-97 capillary puller instrument (One Digital Drive Novato, California, USA). For the detection of deamidated hIAPP and dissociated fibrils, an Apollo II electrospray ionisation (ESI) source (Bruker Daltonik GmbH, Bremen, Germany) was used with a capillary voltage of 4-4.5 kV. The ESI flow rate was optimised to 100-150 µL/h and the source temperature was set to 200°C. Ions were externally accumulated in a hexapole collision cell before transferred to an infinity cell (ICR

cell) for excitation and detection.[61]   Data obtained from FTICR-MS were analysed using Bruker DataAnalysis 4.2 software (Bruker Daltonics, Bremen, Germany).

MS/MS fragmentation in the FTICR MS.   For the collisionally activated dissociation (CAD) experiments, precursor ions were first isolated using the quadrupole mass filter, then collided with argon gas and accumulated in the collision cell.   The collision energy was optimised to 2-18 V and the ion accumulation time to 1-3 seconds.   Fragments were then transferred to the infinity cell for detection.   For the ECD experiments, precursor ions were first isolated using the quadrupole with a wide mass-to-charge isolation window (~100 Da) and accumulated in the collision cell for 1-3 seconds.   Ions were then transferred to the infinity cell, further isolated using Multi-CHEF[62] isolation with an excitation power of 45-55%, and then irradiated with 1.2 eV electrons from a 1.5 A heated hollow cathode.   For the infrared-ECD (IR-ECD) experiments, a 25 W continuous-wave CO_2 infrared laser (Synrad Inc, US) was used to activate the protein during ECD in order to increase the fragmentation efficiency.[63]   The pulse length and laser power were optimised to 20-50 ms and 12.5 W (50% power) respectively, which is just below the influence threshold for obtaining IRMPD fragment ions.   The most intense isotopic peak from each fragment with signal-to-noise ratio (S/N) over five was manually matched with the theoretical *m/z*.   All of the fragments were internally calibrated and then assigned with an uncertainty less than one part-per-million (ppm) (see Supplementary Information, Table S2.1-S2.12).

Studying intermolecular binding strength within oligomers via FTICR MS/MS.   The area of the highest isotope peak of each oligomer was measured

using Bruker DataAnalysis 4.2 software. The percentage of dimer ion (%) was calculated as follows:

$$\text{Dimer } (\%) = \frac{\text{Peak area of dimer}}{\text{Total peak area of monomer} + \text{dimer}} \times 100\%$$

A dissociation curve for dimer ions was obtained by plotting the percentage of dimer remaining against the CAD energy used. The same calculation method was applied to obtain the dissociation curve of other oligomer ions.

Transmission electron microscopy (TEM). The TEM images of the incubated solutions, including 10 µM hIAPP solutions, 10 µM synthetic mutant hIAPPs solutions, 10 µM synthetic ^{30}TNVGSNTY^{37}-NH$_2$ solution, and mixture of hIAPP & ^{30}TNVGSNTY^{37}-NH$_2$ solution, were acquired on a Jeol 2010F TEM operated at 200 kV. 10 µL of incubated solution was transferred onto a carbon-coated grid and settled for one minute, followed by removing the excess solution using filter paper. A 2% (w/v) uranyl acetate solution was used for the negative stain. Multiple images with magnification from x10,000 to x40,000 were acquired.

Thioflavin T (ThT) fluorescence reactivity. The fluorescence reactivity of the 10 µM hIAPP, mutant ((D)$_3$hIAPP), mutant ((isoD)$_3$hIAPP), and hIAPP seeding solutions were measured using a GloMax®-Multi Detection System (Promega; Wisconsin, USA). All samples were placed in a black 96 well-plate and mixed with 10 µM ThT aqueous solution. Fluorescence spectra of the samples were acquired every 45minutes with excitation at 405 nm and emission measurement at 490 nm, in a similar fashion to Chan *et al*.[64, 65].

Incubated solutions were diluted with water to 10 µM and mixed with 10 µM ThT solution. Fluorescence spectra of incubated hIAPP solutions were acquired under same experimental conditions as above. The intensities obtained from the

fluorescence spectrometer were normalised to the signal intensity of the most mature fibril.

2.4 Results

Early oligomer formation in hIAPP.   Aggregated hIAPP is composed of mature and insoluble polypeptide fibrils, whose overall structure has been reported previously.[33, 66]   The early, soluble oligomers are viable for MS study and can provide insights into the fundamental amino acid residue level binding between individual hIAPP monomer units and how they link together to form a higher order structure.   High resolution MS allowed observation of the early oligomers of hIAPP via direct infusion without prior separation.   Oligomers, up to the pentamer, were detected in the nESI-MS spectrum of the 10 μM hIAPP in aqueous solution with 2% DMSO residual present in the solvent (see Figure 2.2A).   Similar observations of the oligomers' distribution pattern were found in the nESI-MS spectrum of the 10 μM hIAPP in aqueous solution only (see Supplementary Information, Figure S2.4).   Despite very closely positioned oligomer peaks and overlapping isotopic envelopes with various charge states (5+ dimer/ 10+ tetramer and 2+ monomer/ 4+ dimer/ 6+ trimer), the isotopic patterns of each oligomer were still well-resolved.   Observed signal intensity of the oligomers decreased with increasing size of oligomer species.

A) MS spectrum of the 10 µM hIAPP solution
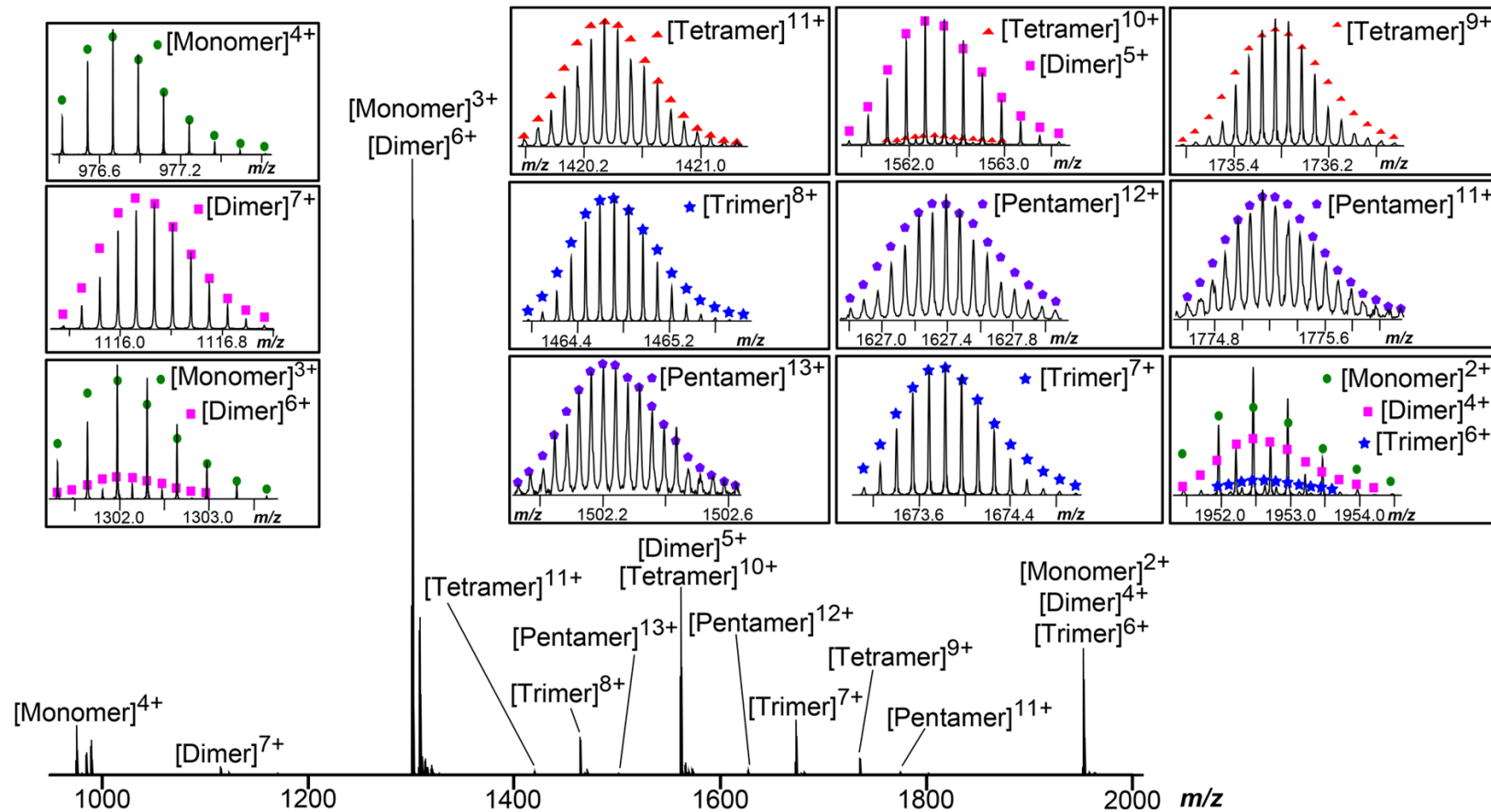
(Figure 2.2. Continue)

B) ECD spectrum of the 7+ hIAPP dimer
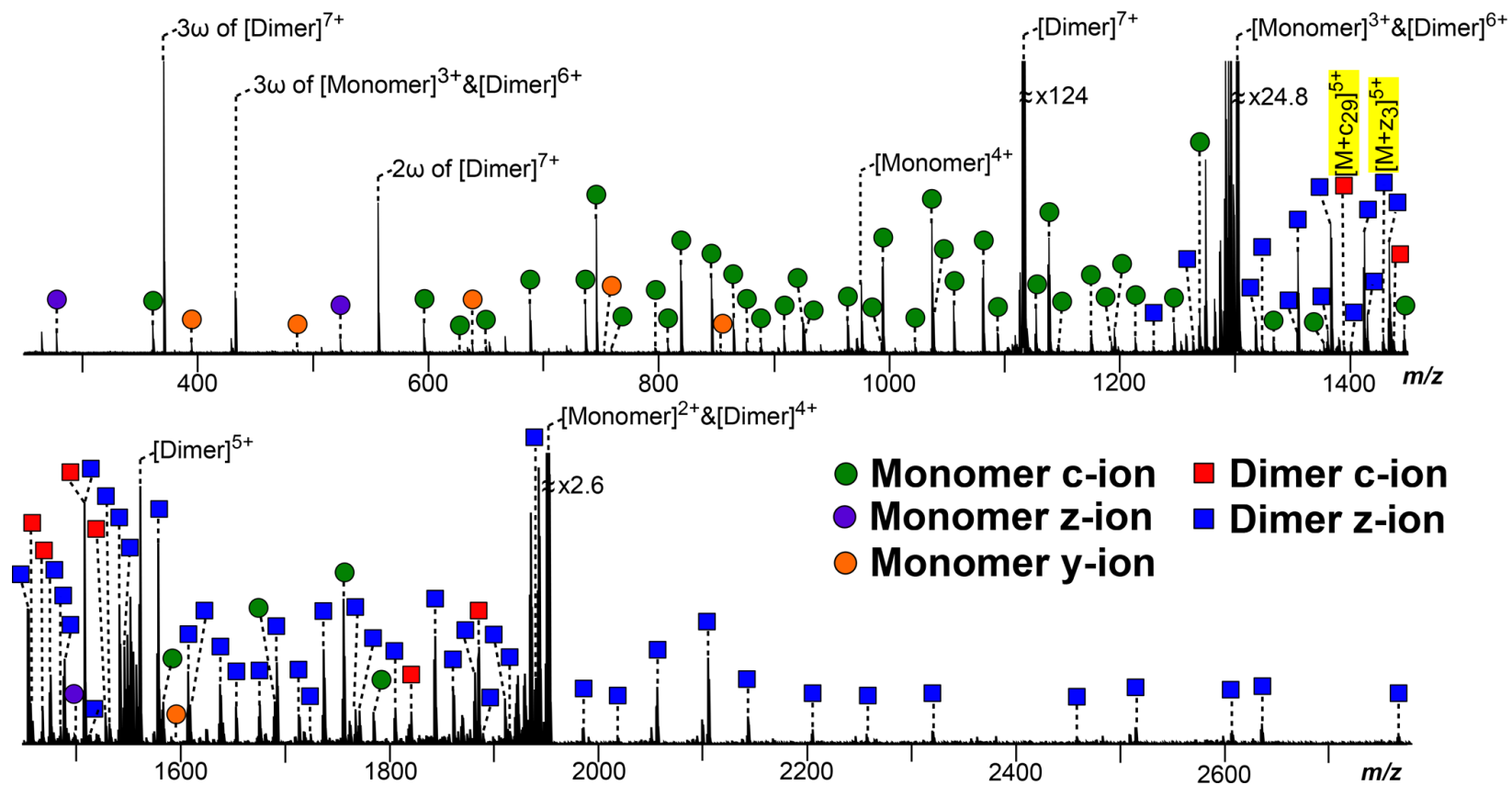
(Figure 2.2. Continue)
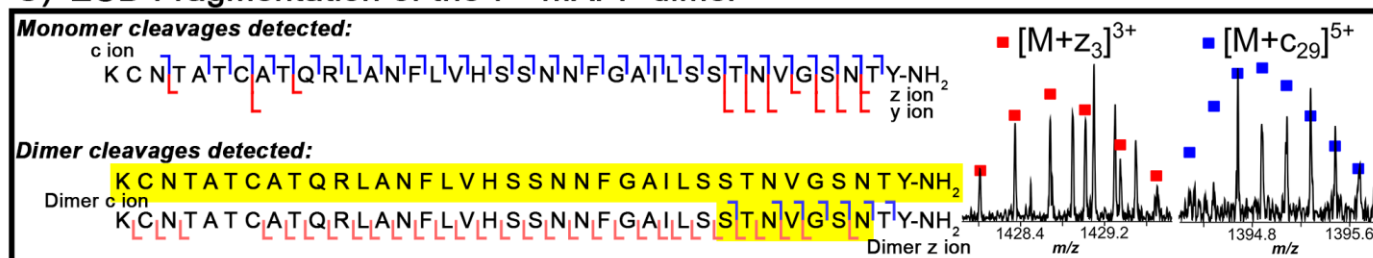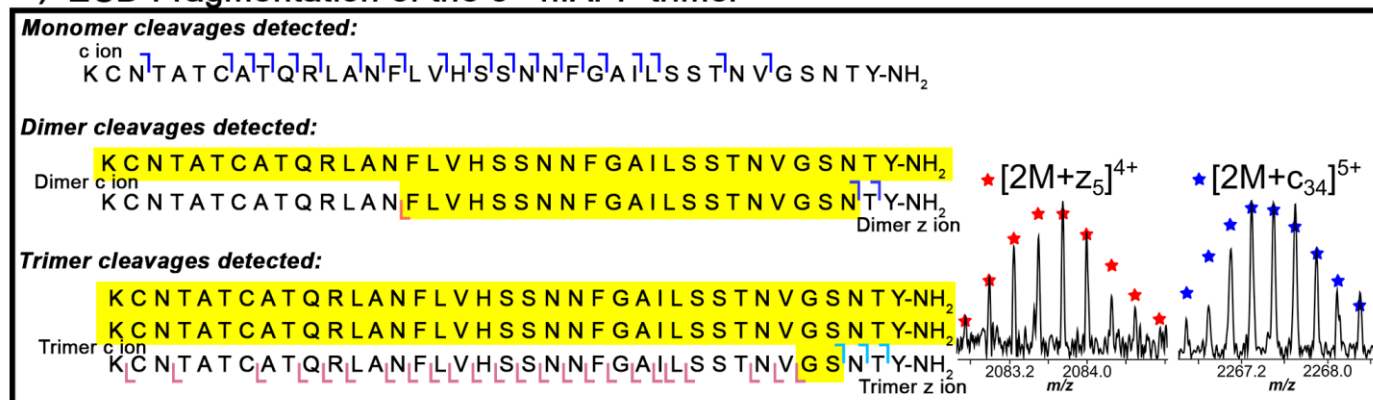
Figure 2.2. (A) An nESI-MS spectrum of the early oligomers from the 10 μM hIAPP aqueous solution with 2% (v/v) residual DMSO.

Oligomers; monomer (circle), dimer (square), trimer (star), tetramer (triangle), and pentamer (pentagon), were detected in the spectrum. Inset,

enlarged regions of spectra are shown for the labelled species. Coloured shapes represent the calculated isotope distributions for the different species overlaid onto the observed patterns. (B) ECD MS/MS spectrum of the 7+ charge state hIAPP dimer. Highlighted fragments represent the critical dimer c- and z-ion observed. The assigned ECD MS/MS fragments of the 7+ charge state hIAPP are included in Supplementary Information Figure S2.9 and Table S2.4. Summarised fragments observed in the ECD MS/MS spectrum of the (C) 7+ dimer and (D) 8+ trimer of hIAPP. Examples of the critical assigned dimer/ trimer c- and z-ions from the ECD MS/MS are inserted. Highlighted sequences represent the proposed non-covalent interaction region. Coloured shapes represent the calculated isotope distributions as in (A)

In order to gain sequence-specific hIAPP interaction information, the hIAPP dimer and trimer ions were also studied via ECD MS/MS.   The ECD MS/MS of the 7+ hIAPP dimer ion (see Figure 2.2B-2.2C; Supplementary Information, Figure S2.9 & Table S2.4) produced 33 monomer c-ions, four monomer z-ions, seven monomer y-ions, seven dimer c-ions, and 30 dimer z-ions (representing the mass of a hIAPP unit plus a c-/z-ion fragment).   The most critical dimer fragments identified from the spectrum were $M+z_3$ and $M+c_{29}$.

ECD MS/MS was further applied to the hIAPP trimer to fragment the oligomer while preserving the non-covalent interaction between hIAPP units. Compared to the previous hIAPP dimer spectrum, the ECD MS/MS spectrum of the 8+ charge state trimer was more complex due to the presence of monomer, dimer, and trimer fragments simultaneously, though sequence informative fragments were successfully observed (see Figure 2.2D; Supplementary Information, Figure S2.10 & Table S2.5).   Aside from the monomer and dimer fragments, trimer c- and z-ion (representing the mass of two hIAPP units plus a c-/z-ion fragment) were also obtained.   21 monomer c-ions, two dimer c-ions, one dimer z-ion, three trimer c-ions, and 24 trimer z-ions were identified from the ECD MS/MS spectrum.   The most critical trimer fragments found in the spectrum were $2M+z_5$ and $2M+c_{34}$.   Similar ECD MS/MS fragments were obtained from hIAPP dimer and trimer which were dissolved solely in aqueous solution (see Supplementary Information, Figure S2.11 & Table S2.6-S2.8).

In order to test whether the proposed interaction region (around Gly-33 and Ser-34) was amyloidogenic, an eight-residue segment of hIAPP (^{30}TNVGSNTY^{37}-NH$_2$) was synthesised, but the structure of the sequence was unknown.   The 3+ charge state trimer of ^{30}TNVGSNTY^{37}-NH$_2$ at *m/z* 1281 (overlapping with 6+ charge state hexamer of the segment) was isolated and fragmented with ECD

MS/MS. The most critical trimer fragments (representing mass of two ^{30}TNVGSNTY^{37}-NH$_2$ plus c-/z-ion) were 2M+z$_5$ and 2M+c$_5$ (see Supplementary Information, Figure S2.13B & Table S2.9).

From soluble oligomers to mature and insoluble fibrils. hIAPP has been shown to oligomerise during incubation,[4, 31, 55] where the soluble oligomers can combine and orientate into a β-sheet structure forming a mature and insoluble fibril. When hIAPP monomers aggregate, the concentration of free hIAPP monomer decreases, and as a result can be used to follow the extent and rate of hIAPP aggregation into oligomers. In order to determine the aggregation rate, a 10 μM solution of hIAPP was incubated at 37°C; the concentration of soluble hIAPP was then determined by comparison of relative MS peak area of the monomer in each aliquot of the incubated sample over 28 days. A reduction in the percentage of soluble hIAPP was observed after the first week of incubation (see Figure 2.3A). After 12 days incubation, ~ 50% reduction of the percentage of soluble hIAPP was observed. After 28 days incubation, only 25% of remaining soluble hIAPP monomer was observed in the 10 μM solution.

A) Quantification of soluble hIAPP monomer

$R^2 = 0.99$

$t_{50\%} = 12$ days

Percentage of Soluble hIAPP (%)

Day of Incubation

- Percentage of Soluble hIAPP
- Exponential Decay Function Fitting
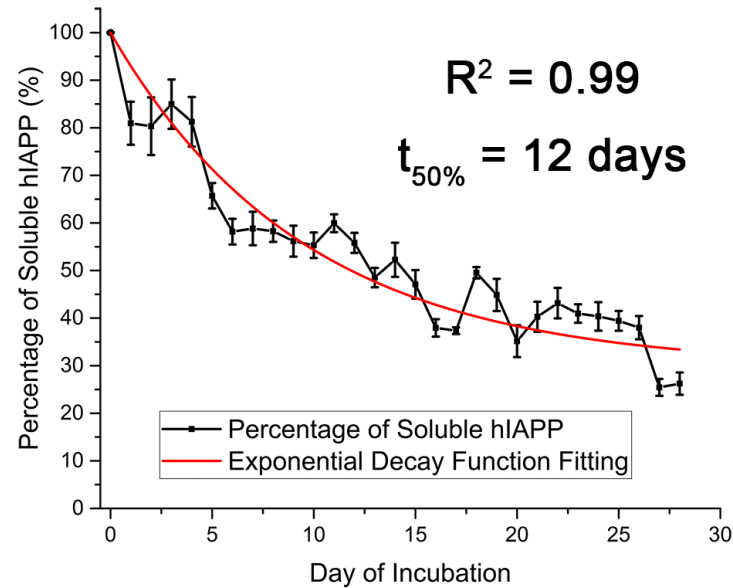
B) Measurement of fluorescence emission at 490 nm

Relative Fluorescence Intensity (at 490 nm)

Hour of Incubation

- 10 µM ThT (Control)

- 5 µM hIAPP
  $R^2 = 0.98$
  $t_{50\%} = 50$ hours
  $y_{max} = 0.36$

- 10 µM hIAPP
  $R^2 = 0.99$
  $t_{50\%} = 49$ hours
  $y_{max} = 0.97$
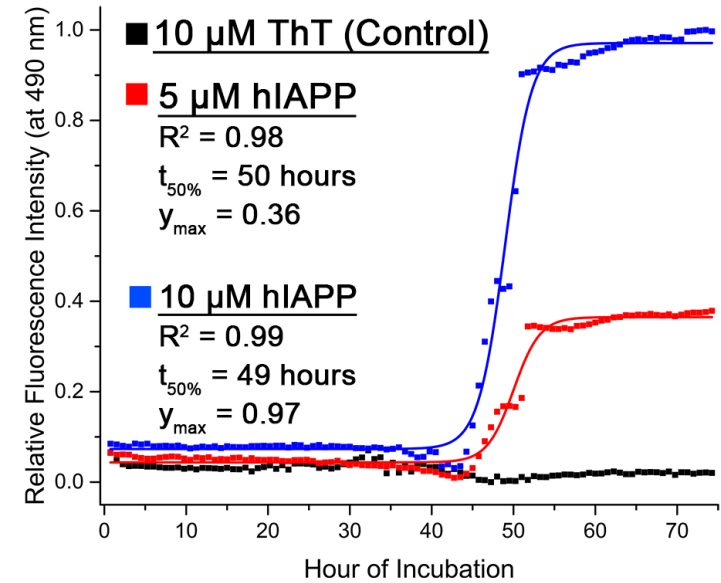
(Figure 2.3. Continue)
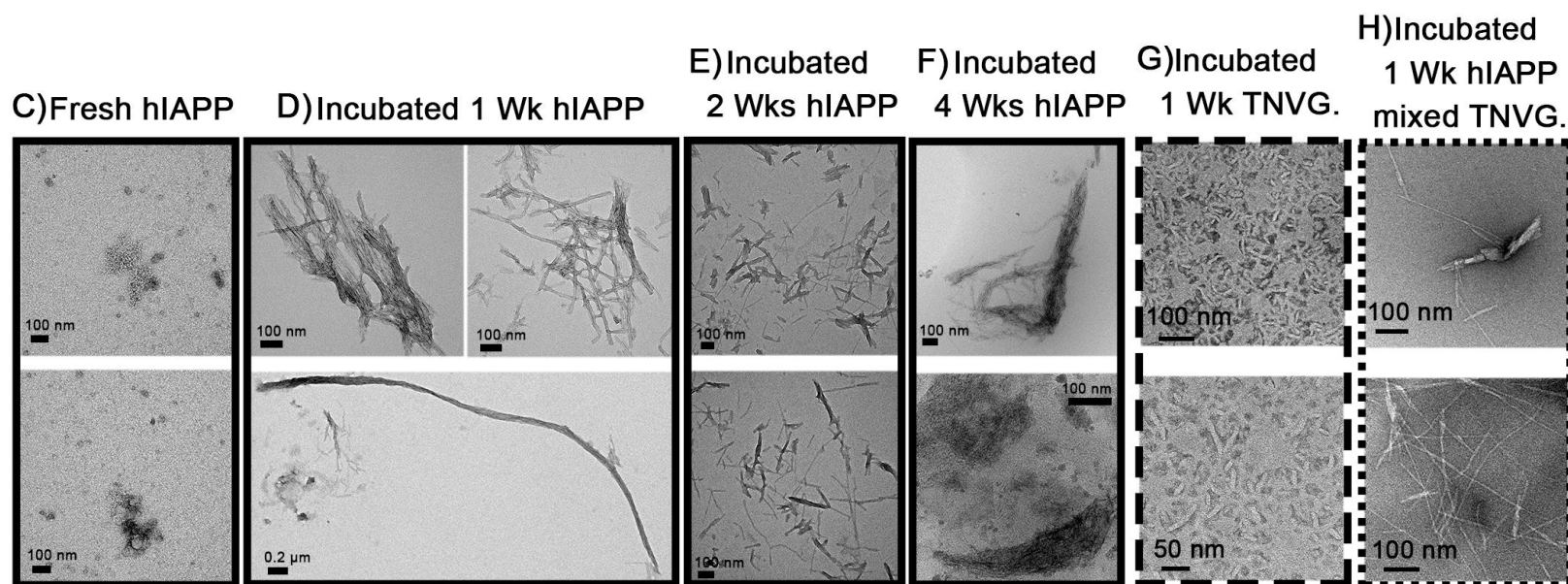
Figure 2.3. (A) Plot showing the amount of soluble hIAPP in the 10 μM solution changing over time measured using MS. (B) The relative fluorescence activity of the 10 μM ThT solution, incubated 5 μM and 10 μM hIAPP solutions measured at emission 490 nm with excitation at 405 nm. The TEM images of the (C) fresh, (D) 1-week, (E) 2-week, and (F) 4-week incubated 10 μM hIAPP solutions. The TEM images of the (G) 1-week 10 μM hIAPP segment solution and (H) 1-week 10 μM hIAPP segment solution mixed with 10 μM hIAPP solution. The scale bars for each TEM image are inset.

FTICR MS is well suited for the quantification of soluble proteins (such as hIAPP) in solution, but ESI is limited to soluble components only, and cannot effectively ionise the mature, insoluble fibrils.   Thus, the relative concentration of amyloid fibril formed in the incubated solutions was determined by the ThT fluorescence reactivity at emission 490 nm (see Figure 2.3B).[64, 67]   Both solutions showed limited detectable aggregation by ThT fluorescence for the first 40 hours of incubation (lag phase), followed by a sharp increase to a maximum value at around 55 hours.   The end time relative emission of the 10 μM hIAPP solution was 2.7 times higher than the emission recorded in the 5 μM solution.

The formation of amyloid fibrils in the 10 μM solutions was also monitored by TEM.   Freshly solubilised hIAPP molecules clump together, which can be observed as small black clusters shown in the TEM images (see Figure 2.3C). Loosely-packed and elongated fibrils were clearly observed in the TEM images of the 1-week incubated sample images (see Figure 2.3D).   The longest fibril observed within the 1-week incubated solution was 5.1 μm.   Long fibrils were still observed in the incubated 2-week sample (see Figure 2.3E), along with shorter, more densely packed fibrils, showing a gradual change in the hIAPP fibril structure at longer incubation time.   Very dense fibrils were observed at the 4-week stage (see Figure 2.3F), showing little remaining space between the fibril branches. Relatively short fibrils were observed in the 10 μM solution of ^{30}TNVGSNTY^{37}-NH$_2$ which had been incubated for one week.   When the hIAPP and segment peptides were mixed together, short fibrils were found in the 1-week incubated mixture of hIAPP and ^{30}TNVGSNTY^{37}-NH$_2$ solution; however, it was impossible to differentiate the components of the fibrils using TEM.

Deamidation site(s) and rate of hIAPP.   Deamidated hIAPP has been

proposed as an accelerating factor for the formation of amyloid fibrils,[55] thus determining the specific deamidation sites is required in order to study the effect of deamidation on the natural hIAPP aggregation mechanism.   Deamidation sites were determined using CAD MS/MS on the deamidated hIAPP species.   The 1+ charge state of the y_5 fragment clearly showed a deamidated peak (+0.984 *m/z*) in the incubated sample (see Figure 2.4A).   For the b fragments, no deamidated peaks were observed for the first 20 amino acid residues (see Figure 2.4B); however, a deamidated peak was present corresponding to the b_{21} fragment (see Figure 2.4C).   Furthermore, another deamidated peak was detected corresponding to the b_{22} fragment ion in the incubated solution (see Figure 2.4D).
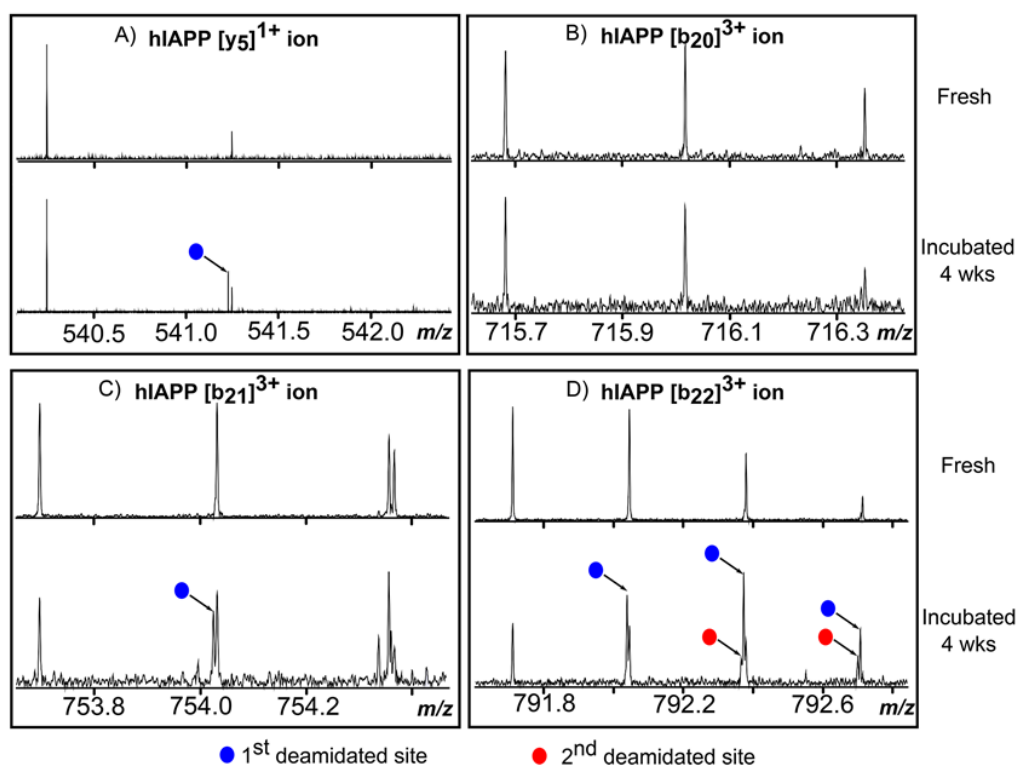
Figure 2.4.   Key CAD MS/MS fragments of hIAPP revealing deamidation at amino acid residues after four weeks incubation at 37°C (A) y_5, (B) b_{20}, (C) b_{21}, and (D) b_{22} are shown. The blue/red circles indicate the isotopic peaks of singly/doubly deamidated fragments, respectively.

Subsequently, the deamidation rates of each deamidated site were determined by calculating the percentage ratio between the non-deamidated and the deamidated fragment ion peak areas for each sample. In the incubated solution, the earliest deamidation instance was observed on day 5, and was identified in the y_5 fragment ion (see Figure 2.5A). Deamidation at the b_{21} and b_{22} fragment ions was observed on day 8. Though the extent of deamidation increased with the incubation time; a similar percentage and rate of deamidation were found in the y_5, b_{21}, and b_{22} fragment ions at the various time points for the incubated solutions. After 28 days incubation, each deamidated fragment ion (y_5, b_{21}, and b_{22}) of the soluble hIAPP showed ~18% deamidation. The fibrillary pellets showed a greater extent of deamidation than the soluble hIAPP species. Deamidation was observed in the day-7, day-14, day-21, and day-28 samples (see Figure 2.5B). On day-7, ~7.5% of the y_5 fragment ion was deamidated while no deamidation was observed for the b_{21} and b_{22} fragment ions in the pellet. On day-14, the extent of deamidation for the y_5 fragment ion increased sharply to 16% and the b21 and b22 showed ~7.5% deamidation. After 28 days incubation, the extent of deamidation of the y_5, b_{21}, and b_{22} fragment ions were at 39%, 30%, and 29% respectively, distinctly larger than for the soluble hIAPP species above.
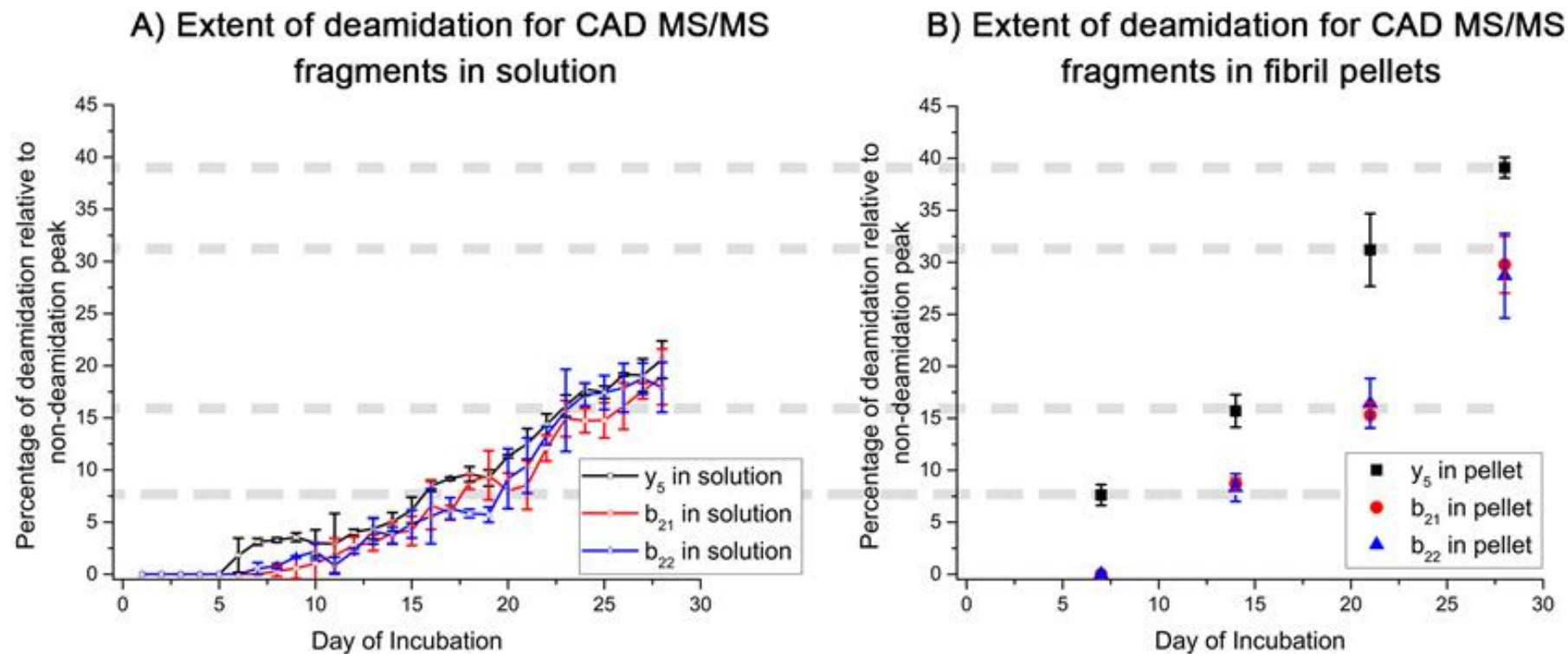
Figure 2.5. The extent of deamidation observed in CAD MS/MS fragments; the y_5 (black), b_{21} (red), and b_{22} (blue) for 10 μM aqueous hIAPP (A) solutions and (B) dissolved fibril solutions.

Aggregation of mutant ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP). In order to ascertain the effects of deamidation on the aggregation process, pure mutant hIAPP peptides were synthesised with aspartic acid ((D)$_3$hIAPP) or isoaspartic acid ((isoD)$_3$hIAPP) replacement at the detected deamidation site from above, i.e. residues 21, 22, and 35 (see Supplementary Information, Figure S2.3). The 7+ charge state dimers of ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) observed in the nESI-MS spectrum of the 10 μM aqueous solutions were fragmented by ECD MS/MS (see Supplementary Information, Figure S2.16, S2.17 & Table S2.11, 2.12), and similar fragments were obtained in both spectra as in the ECD spectra of the 7+ charge state dimer of wild-type hIAPP. The relative fluorescence activities of the 10 μM solutions of ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) were measured and the lag phases of the formation of mature fibrils were ~ 30 hours (see Supplementary Information; Figure S2.18). TEM images of the ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) solutions incubated at 37$^\text{o}$C for one week were also obtained (see Figure 2.6A & 2.6B). Branched and elongated fibrils were observed in ((D)$_3$hIAPP) incubated solution; while densely packed fibrils were found in ((isoD)$_3$hIAPP) solution. The fibril structures observed in ((isoD)$_3$hIAPP) solution were distinctly different from the fibrils found in wild-type hIAPP and ((D)$_3$hIAPP).
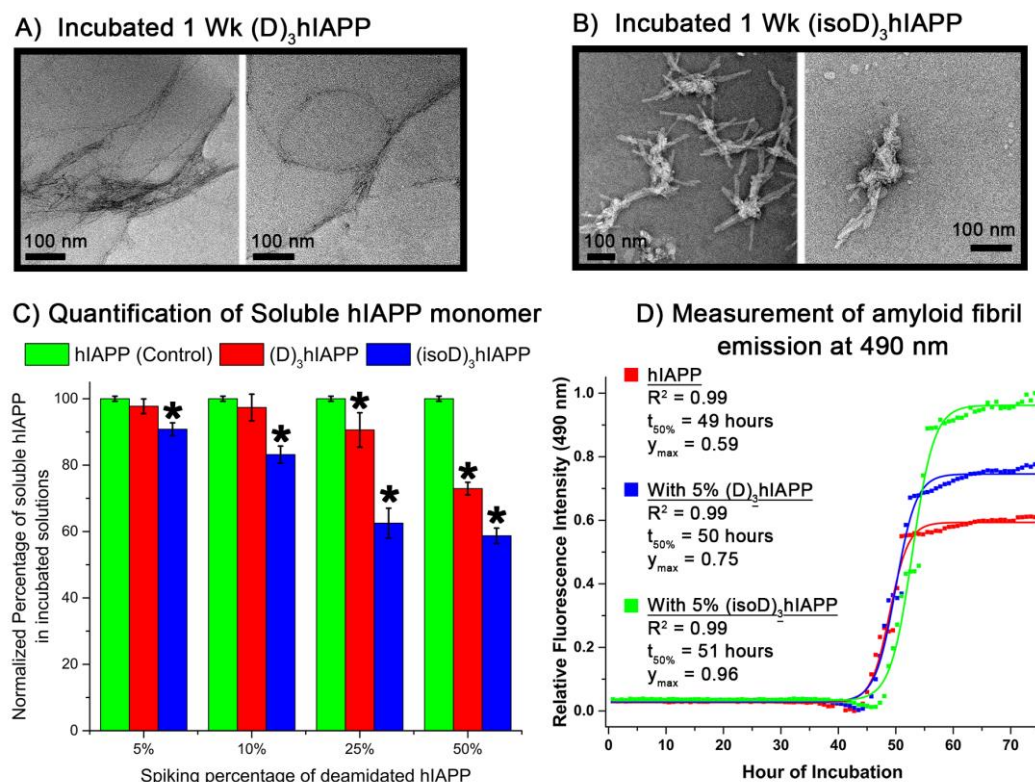
Figure 2.6. The TEM images of 10 μM (A) ((D)$_3$hIAPP) and (B) ((isoD)$_3$hIAPP) solutions after one week incubation at 37°C. The scale bar for each TEM image is shown accordingly. (C) A plot showing the normalised percentage of soluble hIAPP in the 10 μM solutions with spiking different percentage of mutant ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP) to wild-type hIAPP (control) measured using MS after one week incubation. * indicates the p-value of the sample percentages are less than 0.05 compared to the control percentage in a paired t-test. (D) The relative fluorescence activity of 10 μM hIAPP, 9.5 μM hIAPP plus 0.5 μM (5%) ((D)$_3$hIAPP), and 9.5 μM hIAPP plus 0.5 μM (5%) ((isoD)$_3$hIAPP) measured at emission 490 nm.

((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) solutions were individually spiked into wild-type hIAPP solution at 5%, 10%, 25%, and 50% amounts. After one week incubation, the percentage of remaining soluble monomeric hIAPP in the seeded

solutions was consistently less in the spiked solutions than the control hIAPP incubated solution, and less soluble hIAPP was found in the mutant ((isoD)$_3$hIAPP) spiked solutions compared to the ((D)$_3$hIAPP) spiked solutions (Figure 2.6C). The p-values of seeding 5% and 10% mutant ((D)$_3$hIAPP) into the solution were higher than 0.05, which indicated the differences between the seeded solutions and the control sample were not significant.   The p-values of the remaining seeded solutions were less than 0.05, which showed significant differences between the control and the seeded solutions.   The normalised percentage of soluble hIAPP with 5% spiking of ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP) was 4% and 9% less than the soluble hIAPP found in the control incubated solutions respectively.   The normalised percentage of soluble hIAPP consistently decreased with increased amounts of ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP) in the incubated solutions.   For the solutions with 50% spiking of ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP), the normalised percentage of soluble hIAPP was 28% and 46% less than the soluble hIAPP found in the control solutions respectively.

The measured percentage change of fluorescence reactivity in 5% mutant hIAPP seeded solutions further indicated that the amount of amyloid fibrils formed in the seeded solutions was consistently higher than the wild-type hIAPP incubated solution and the overall percentage of amyloid fibrils formed in ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) seeded solutions was 27% and 48% higher than the wild-type hIAPP incubated solution respectively (Figure 2.6D).   Both measurements showed that the deamidated hIAPP enhanced the amyloid fibril formation, and the ((isoD)$_3$hIAPP) had a ~2 fold larger effect on enhancing hIAPP fibril formation than ((D)$_3$hIAPP).

In order to examine the effects of mutant ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) in incubated solutions and fibrillary pellets, MS spectra of mixed wild-type hIAPP

with various spiking percentage of mutant ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) were acquired (see Supplementary Information, Figure S2.19). The peak intensity of mutant ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) were consistently higher in the fibrillary pellets compared to the incubated and the fresh solutions.

2.5 Discussion

Understanding the aggregation mechanism of hIAPP will help the development of appropriate drugs to inhibit the formation of amyloid fibrils.[68, 69] However, the difficulties in handling aggregated hIAPP and the conflicting reports of proposed binding sites hinder drug development. In addition, various reports showed that the rate of hIAPP fibril formation was accelerated by seeding the solution with deamidated polypeptides,[54, 55] but the role of deamidation under general hIAPP aggregation is poorly understood. Thus, high resolution mass spectrometry (FTICR MS) and tandem mass spectrometry were applied to determine the position of the non-covalent bonding interaction between hIAPPs, as well as the role of deamidation in aqueous solution. TEM and ThT fluorescence spectrometry were also utilised for monitoring the insoluble fibrils formed during the experiment.

FTICR MS and MS/MS investigation of hIAPP interaction regions. A series of oligomers (from monomer to pentamer) was found in the mass spectrum of the wild-type hIAPP solution (see Figure 2.2A). The detection of consecutive hIAPP oligomers in the mass spectrum (dimer, trimer, tetramer, and pentamer) suggests that hIAPP associates in a sequential process, with individual monomers attaching to growing oligomers. The relative signal intensity of each oligomer generally decreased with increasing number of monomer units, suggesting observed species are early oligomers from sequential aggregation. Larger hIAPP oligomers above pentamer were not observed, which is attributed to oligomer's decreasing solubility and thus viability for ESI-MS.

The dissociation of dimers and trimer into monomer hIAPP units in CAD MS/MS indicates the interactions between hIAPP units are fragile and non-covalent (see Supplementary Information, Figure S2.5 & S2.6). Comparatively,

the ECD MS/MS spectra of hIAPP dimer and trimer ions provide details of the non-covalent interaction between individual hIAPP units. Even though limited dimer fragments were obtained from the ECD MS/MS (see Supplementary Information, Figure S2.7) and the IR-ECD MS/MS (See Supplementary Information, Figure S2.8) spectra of the 5+ dimer ion, the 7+ hIAPP dimer ion showed extensive fragmentation under ECD MS/MS (see Figure 2.2B & 2.2C). With a higher charge state precursor ion, the number of identified dimer z-ions increased by 27, and four more dimer c-ions were observed (for a total of 30 dimer z ions and seven dimer c ions). The $M+c_{29}$ ion (representing mass equal to one hIAPP (M) plus a c_{29} ion from a second hIAPP unit) was the first c-ion dimer fragment observed from the 7+ hIAPP dimer, which indicates the hIAPP dimer begins to interact at Ser-29 (towards the N-terminus). Subsequent c-ion dimer fragments, from $M+c_{29}$ to $M+c_{36}$ were also observed, indicating an interaction region between Ser-29 and Thr-36. On the other hand the $M+z_3$ (representing mass equal to one hIAPP (M) plus a z_3 ion from a second hIAPP unit) was the first z-ion dimer fragments observed from the 7+ hIAPP dimer ECD spectrum, which suggests the hIAPP dimer interaction begins at Asn-35 (towards the C-terminus), and as a result, subsequent z-ion dimer fragments, from $M+z_5$ to $M+z_{36}$, were also detected, indicating an interaction region between Cys-2 and Asn-35. The overlapping regions from dimer c- and z-ions indicates that there are several interaction points in between Ser-29 to Asn-35. The proposed interaction region also agrees with the observation from the ECD MS/MS spectra of 5+ and 7+ charge states of hIAPP dimer that the hIAPP units are symmetrically interacting near the C-terminus.

In order to extend the knowledge of the hIAPP aggregation mechanism, the hIAPP trimer 8+ ion was also interrogated using ECD MS/MS. The only

monomeric fragments observed in the ECD spectrum of 8+ hIAPP trimer are c-ions, which demonstrates the consistency of interaction towards the C-terminus region of 5+ hIAPP dimer, 7+ hIAPP dimer, and 8+ hIAPP trimer species (see Figure 2.2D). The observation of 2M+z5 and 2M+c34 ions indicates the association region of the third hIAPP unit is located around Gly-33 and Ser-34. The non-covalent interactions observed from the ECD spectra of the 7+ hIAPP dimer and the 8+ hIAPP trimer are highly correlated showing that the aggregation mechanism of trimer is very similar to that of hIAPP dimerization.

The proposed interaction region of hIAPP oligomers was further investigated through studying the aggregation of a synthesised eight-residue segment of hIAPP (^{30}TNVGSNTY^{37}-NH$_2$). The sequential dissociation trend of ^{30}TNVGSNTY^{37}-NH$_2$ oligomers under CAD MS/MS (see Supplementary Information, Figure S2.13A) showed that a similar association pattern/region is adopted between ^{30}TNVGSNTY^{37}-NH$_2$ and hIAPP oligomers. This hypothesis is strongly supported by the ECD fragments obtained from the trimer of ^{30}TNVGSNTY^{37}-NH$_2$ as the same aggregation region (the Gly-33 and Ser-34 residues) was observed (see Supplementary Information, Figure S2.13B).

Rate of hIAPP aggregation. Soluble hIAPP units were shown to rapidly form mature fibrils within just one week of incubation which led to a dramatic reduction in the percentage of soluble polypeptide observed (see Figure 2.3A). Fluorescence experiments using ThT (see Figure 2.3B) indicated that the lag phases of mature hIAPP fibril formation were similar in the 5 µM and 10 µM hIAPP solutions; however, the amount of amyloid fibril formed was greater than two fold higher in the 10 µM hIAPP solution compared to the 5 µM hIAPP solution. The amyloid fibril formation was further revealed by the TEM images recorded at different time points (see Figure 2.3C-2.3F). The synthesised peptide

(^{30}TNVGSNTY^{37}-NH$_2$) also formed observable fibrils within one week which could indicate its ability to be amyloidogenic (see Figure 2.3G & 2.3H).

Deamidation site(s) and rate of hIAPP. The exact deamidation sites within monomer hIAPP were investigated using CAD MS/MS. The first deamidation site can be localised at Asn-21 and the second at Asn-22 via the observed deamidated peaks for the b$_{21}$ and b$_{22}$ fragment ions (see Figure 2.4). A y$_5$ fragment also showed deamidation in the 4-week sample indicating there was a third deamidation site at the Asn-35 residue within 4-weeks of incubation at 37°C. The deamidation rate of each site was then calculated using the peak area ratio between non-deamidated and deamidated peaks, and a similar deamidation rate was found for the Asn-21, Asn-22, and Asn-35 sites in the soluble hIAPP (Figure 2.5). The extent of deamidation at the Asn-21, Asn-22, and Asn-35 residues in fibril pellets was observed to be consistently higher at all time-points. The residue Asn-35 was the earliest deamidation site observed in the incubated solutions and contained the highest extent of deamidation in the fibril pellets, which agrees with the findings in previous deamidation studies,[42, 70] suggesting that the primary sequence of Ser-35Asn-Thr in hIAPP had the shortest deamidation time among all the other segments studied.

Aggregation of mutant ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP). The fluorescence activity of ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) solutions showed mutant hIAPPs samples had a shorter lag phase in forming mature fibrils compared to the wild type hIAPP (see Supplementary Information, Figure S2.18), which supports the hypothesis that deamidated hIAPPs accelerate the aggregation rate of non-deamidated hIAPP. The lag phase of amyloid fibril formation did not show a distinct increase upon seeding 5% mutant hIAPPs into solutions; the MS quantification results and the fluorescence data showed the rate and the amount of

amyloid fibril formation were enhanced upon seeding the solution with various percentages of mutant hIAPP (see Figure 2.6B & 2.6C). Additionally, a greater amount of mutant hIAPPs was observed in the fibrillary pellet compared to the control solutions; which also suggests the important role of ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP) in enhancing fibril formation (see Supplementary Information, Figure S2.19). The fluorescence measurement and MS quantification showed the effect of seeding ((isoD)$_3$hIAPP) is twice that of seeding the same amount of ((D)$_3$hIAPP) into non-deamidated hIAPP solutions. The difference between the TEM images of incubated ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) solutions also agrees the greater effect of ((isoD)$_3$hIAPP) compared to ((D)$_3$hIAPP) (see Figure 2.6A & 2.6B). The observed morphology of fibrils was different in ((isoD)$_3$hIAPP) compared to ((D)$_3$hIAPP), indicating differences in the aggregation behavior and rate of non-deamidated hIAPP, as observed in the MS and fluorescence.

2.6 Conclusion

The experimental results here demonstrate that the most important aggregation regions between individual hIAPP units within oligomers in the gas phase occurs toward the C-terminus of the peptides, around residues Gly-33 and Ser-34. The aggregation mechanisms of the dimer and trimer of hIAPP were shown by ECD MS/MS to be very similar, with both structures interacting primarily towards the C-terminus of the peptides. Furthermore, the aggregation of individual hIAPP units into oligomers was shown to occur much faster than deamidation. Mature fibrils were formed within two days of incubation and deamidated hIAPP was not observed until the fifth day, suggesting that, in general, aggregation and deamidation are independent processes. Asn-21, Asn-22, and Asn-35 residues are largely deamidated, and no deamidation was observed at the other Asn/Gln residues. The extent of deamidation was higher in the fibrillary pellets than in the solutions, which indicates the deamidated hIAPP tends to form fibrils faster than the wild-type hIAPP. This hypothesis is further supported by accelerated aggregation rate observed for ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) and a higher percentage of mutant hIAPPs identified in the fibrils of hIAPP solutions spiked with mutants ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP) in wild-type hIAPP solutions. The TEM images further showed the different morphology of (isoD)$_3$hIAPP fibrils compared to wild-type hIAPP and ((D)$_3$hIAPP) fibrils, which explains the acceleration of aggregation caused by the deamidated hIAPP. Moreover, the amount of amyloid fibril formed is directly proportional to the spiking percentage of mutant hIAPPs; and the effect of seeding ((isoD)$_3$hIAPP) is twice that of seeding the same amount of ((D)$_3$hIAPP) into non-deamidated hIAPP solutions.

2.7 References

1. Koo, E. H.; Lansbury, P. T.; Kelly, J. W., Amyloid diseases: abnormal protein aggregation in neurodegeneration. *Proc. Natl. Acad. Sci. USA* **1999,** *96* (18), 9989-9990.

2. Conway, K. A.; Lee, S. J.; Rochet, J. C.; Ding, T. T.; Williamson, R. E.; Lansbury, P. T., Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy. *Proc. Natl. Acad. Sci. USA* **2000,** *97* (2), 571-576.

3. Cao, P.; Abedini, A.; Wang, H.; Tu, L.-H.; Zhang, X.; Schmidt, A. M.; Raleigh, D. P., Islet amyloid polypeptide toxicity and membrane interactions. *Proc. Natl. Acad. Sci. USA* **2013,** *110* (48), 19279-19284.

4. Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T., Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to beta-sheet in amyloid fibril formation. *Nat. Chem.* **2011,** *3* (2), 172-177.

5. Makin, O. S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L. C., Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. USA* **2005,** *102* (2), 315-320.

6. Paul, W. E.; Cohen, A. S., Electron Microscopic studies of Amyloid Fibrils with Ferritinconjugated Antibody. *Am. J. Pathol.* **1963,** *43*, 721-38.

7. Shirahama, T.; Cohen, A. S., High-resolution electron microscopic analysis of the amyloid fibril. *J. Cell Biol.* **1967,** *33* (3), 679-708.

8. Volpatti, L. R.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P., A clear view of polymorphism, twist, and chirality in amyloid fibril formation. *ACS nano* **2013,** *7* (12), 10443-10448.

9. Kahn, S. E.; Dalessio, D. A.; Schwartz, M. W.; Fujimoto, W. Y.; Ensinck, J. W.; Taborsky, G. J.; Porte, D., Evidence of cosecretion of islet amyloid polypeptide and insulin by beta-cells. *Diabetes* **1990,** *39* (5), 634-638.

10. Cooper, G. J. S.; Day, A. J.; Willis, A. C.; Roberts, A. N.; Reid, K. B. M.; Leighton, B., Amylin and the amylin gene: structure, function and relationship to islet amyloid and to diabetes mellitus. *Biochim. Biophys. Acta* **1989,** *1014* (3), 247-258.

11. Jaikaran, E.; Clark, A., Islet amyloid and type 2 diabetes: from molecular misfolding to islet pathophysiology. *Biochim. Biophys. Acta-Mol. Basis Dis.* **2001,** *1537* (3), 179-203.

12. Luca, S.; Yau, W. M.; Leapman, R.; Tycko, R., Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *J. Biochem.* **2007,** *46* (47), 13505-13522.

13.   Williamson, J. A.; Miranker, A. D., Direct detection of transient alpha-helical states in islet amyloid polypeptide. *Protein Sci.* **2007,** *16* (1), 110-117.

14.   Westermark, G. T.; Westermark, P.; Berne, C.; Korsgren, O.; Nordic Network Clin Islet, T., Widespread amyloid deposition in transplanted human pancreatic islets. *N. Engl. J. Med.* **2008,** *359* (9), 977-979.

15.   Ashcroft, F. M.; Rorsman, P., Diabetes mellitus and the β cell: the last ten years. *Cell* **2012,** *148* (6), 1160-1171.

16.   Westermark, P.; Engstrom, U.; Johnson, K. H.; Westermark, G. T.; Betsholtz, C., Islet Amyloid Polypeptide - Pinpointing Amino-Acid residues Linked to Amyloid Fibril Formation. *Proc. Natl. Acad. Sci. USA* **1990,** *87* (13), 5036-5040.

17.   Wiltzius, J. J. W.; Sievers, S. A.; Sawaya, M. R.; Eisenberg, D., Atomic structures of IAPP (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process. *Protein Sci.* **2009,** *18* (7), 1521-1530.

18.   Fernandez, M. S., Human IAPP amyloidogenic properties and pancreatic beta-cell death. *Cell Calcium* **2014,** *56* (5), 416-427.

19.   Laghaei, R.; Mousseau, N.; Wei, G., Effect of the Disulfide Bond on the Monomeric Structure of Human Amylin Studied by Combined Hamiltonian and Temperature Replica Exchange Molecular Dynamics Simulations. *J. Phys. Chem. B* **2010,** *114* (20), 7071-7077.

20.   Akter, R.; Cao, P.; Noor, H.; Ridgway, Z.; Tu, L. H.; Wang, H.; Wong, A. G.; Zhang, X. X.; Abedini, A.; Schmidt, A. M.; Raleigh, D. P., Islet Amyloid Polypeptide: Structure, Function, and Pathophysiology. *J. Diabetes Res.* **2016**, 18.

21.   Mishra, R.; Geyer, M.; Winter, R., NMR Spectroscopic Investigation of Early Events in IAPP Amyloid Fibril Formation. *ChemBioChem* **2009,** *10* (11), 1769-1772.

22.   Ashburn, T. T.; Lansbury, P. T., Interspecies Sequences Variations Affect the Kinetics and Thermodynamics of Amyloid Formation: Peptide Models of Pancreatic Amyloid. *J. Am. Chem. Soc.* **1993,** *115* (23), 11012-11013.

23.   Azriel, R.; Gazit, E., Analysis of the structural and functional elements of the minimal active fragment of islet amyloid polypeptide (IAPP) - An experimental support for the key role of the phenylalanine residue in amyloid formation. *J. Biol. Chem.* **2001,** *276* (36), 34156-34161.

24.   Rhoades, E.; Agarwal, J.; Gafni, A., Aggregation of an amyloidogenic fragment of human islet amyloid polypeptide. *Biochim. Biophys. Acta-Protein Struct. Molec. Enzym.* **2000,** *1476* (2), 230-238.

25.   Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M. L.; Voelter, W.; Brunner, H.; Kapurniotu, A., Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic

and cytotoxic properties. *J. Mol. Biol.* **2000,** *295* (4), 1055-1071.

26. Zanuy, D.; Porat, Y.; Gazit, E.; Nussinov, R., Peptide sequence and amyloid formation: Molecular simulations and experimental study of a human islet amyloid polypeptide fragment and its analogs. *Structure* **2004,** *12* (3), 439-455.

27. Scrocchi, L. A.; Chen, Y.; Waschuk, S.; Wang, F.; Cheung, S.; Darabie, A. A.; McLaurin, J.; Fraser, P. E., Design of peptide-based inhibitors of human islet amyloid polypeptide fibrillogenesis. *J. Mol. Biol.* **2002,** *318* (3), 697-706.

28. Rhoades, E.; Gafni, A., Micelle formation by a fragment of human islet amyloid polypeptide. *Biophys. J.* **2003,** *84* (5), 3480-3487.

29. Buchanan, L. E.; Dunkelberger, E. B.; Tran, H. Q.; Cheng, P.-N.; Chiu, C.-C.; Cao, P.; Raleigh, D. P.; de Pablo, J. J.; Nowick, J. S.; Zanni, M. T., Mechanism of IAPP amyloid fibril formation involves an intermediate with a transient β-sheet. *Proc. Natl. Acad. Sci. USA* **2013,** *110* (48), 19285-19290.

30. Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E., Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nat. Chem.* **2015,** *7* (1), 73-81.

31. Young, L. M.; Cao, P.; Raleigh, D. P.; Ashcroft, A. E.; Radford, S. E., Ion Mobility Spectrometry-Mass Spectrometry Defines the Oligomeric Intermediates in Amylin Amyloid Formation and the Mode of Action of Inhibitors. *J. Am. Chem. Soc.* **2014,** *136* (2), 660-670.

32. Dupuis, N. F.; Wu, C.; Shea, J. E.; Bowers, M. T., The Amyloid Formation Mechanism in Human IAPP: Dimers Have beta-Strand Monomer-Monomer Interfaces. *J. Am. Chem. Soc.* **2011,** *133* (19), 7240-7243.

33. Dupuis, N. F.; Wu, C.; Shea, J. E.; Bowers, M. T., Human Islet Amyloid Polypeptide Monomers Form Ordered beta-hairpins: A Possible Direct Amyloidogenic Precursor. *J. Am. Chem. Soc.* **2009,** *131* (51), 18283-18292.

34. Young, L. M.; Tu, L.-H.; Raleigh, D. P.; Ashcroft, A. E.; Radford, S. E., Understanding co-polymerization in amyloid formation by direct observation of mixed oligomers. *Chem. Sci.* **2017,** *8* (7), 5030-5040.

35. Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H., Jr., Ion mobility-mass spectrometry. *J. Mass Spectrom.* **2008,** *43* (1), 1-22.

36. Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R., Ion mobility mass spectrometry of proteins and protein assemblies. *Chem. Soc. Rev.* **2010,** *39* (5), 1633-1655.

37. Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C., Islet amyloid in type 2 diabetes, and the toxic oligomer hypothesis. *Endocr. Rev.* **2008,** *29* (3), 303-316.

38. Patil, S. M.; Xu, S. H.; Sheftic, S. R.; Alexandrescu, A. T., Dynamic alpha-

Helix Structure of Micelle-bound Human Amylin. *J. Biol. Chem.* **2009,** *284* (18), 11982-11991.

39. Stadtman, E. R., Protein modification in aging. *J. Gerontol.* **1988,** *43* (5), B112-B120.

40. Lindner, H.; Helliger, W., Age-dependent deamidation of asparagine residues in proteins. *Exp. Gerontol.* **2001,** *36* (9), 1551-1563.

41. Geiger, T.; Clarke, S., Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides - Succinimide-Linked Reactions that Contribute to Protein-Degradation. *J. Biol. Chem.* **1987,** *262* (2), 785-794.

42. Robinson, N. E.; Robinson, A. B., Molecular clocks. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (3), 944-949.

43. Cournoyer, J. J.; Pittman, J. L.; Ivleva, V. B.; Fallows, E.; Waskell, L.; Costello, C. E.; O'Connor, P. B., Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Sci.* **2005,** *14* (2), 452-463.

44. Wright, H. T., Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. *Crit. Rev. Biochem. Mol. Biol.* **1990,** *26* (1), 1-52.

45. Robinson, A. B.; Rudd, C. J., Deamidation of glutaminyl and asparaginyl residues in peptides and proteins. *Curr. Top. Cell. Regul.* **1974,** *8* (0), 247-95.

46. Mamula, M. J.; Gee, R. J.; Elliott, J. I.; Sette, A.; Southwood, S.; Jones, P. J.; Blier, P. R., Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J. Biol. Chem.* **1999,** *274* (32), 22321-22327.

47. Doyle, H. A.; Mamula, M. J., Post-translational protein modifications in antigen recognition and autoimmunity. *Trends Immunol.* **2001,** *22* (8), 443-449.

48. Tomidokoro, Y.; Rostagno, A.; Neubert, T. A.; Lu, Y.; Rebeck, G. W.; Frangione, B.; Greenberg, S. M.; Ghiso, J., Iowa Variant of Familial Alzheimer's Disease Accumulation of Posttranslationally Modified A beta D23N in Parenchymal and Cerebrovascular Amyloid Deposits. *Am. J. Pathol.* **2010,** *176* (4), 1841-1854.

49. Shimizu, T.; Fukuda, H.; Murayama, S.; Izumiyama, N.; Shirasawa, T., Isoaspartate formation at position 23 of amyloid beta peptide enhanced fibril formation and deposited onto senile plaques and vascular amyloids in Alzheimer's disease. *J. Neurosci. Res.* **2002,** *70* (3), 451-461.

50. Sargaeva, N. P.; Lin, C.; O'Connor, P. B., Identification of Aspartic and Isoaspartic Acid Residues in Amyloid beta Peptides, Including A beta 1-42, Using Electron-Ion Reactions. *Anal. Chem.* **2009,** *81* (23), 9778-9786.

51. Kad, N. M.; Thomson, N. H.; Smith, D. P.; Smith, D. A.; Radford, S. E., beta(2)-microglobulin and its deamidated variant, N17D form amyloid fibrils with

a range of morphologies in vitro. *J. Mol. Biol.* **2001,** *313* (3), 559-571.

52.   Soulby, A. J.; Heal, J. W.; Barrow, M. P.; Roemer, R. A.; O'Connor, P. B., Does deamidation cause protein unfolding? A top-down tandem mass spectrometry study. *Protein Sci.* **2015,** *24* (5), 850-860.

53.   Gupta, R.; Srivastava, O. P., Effect of deamidation of asparagine 146 on functional and structural properties of human lens alpha B-crystallin. *Invest. Ophthalmol. Vis. Sci.* **2004,** *45* (1), 206-214.

54.   Nilsson, M. R.; Driscoll, M.; Raleigh, D. P., Low levels of asparagine deamidation can have a dramatic effect on aggregation of amyloidogenic peptides: Implications for the study of amyloid formation. *Protein Sci.* **2002,** *11* (2), 342-349.

55.   Dunkelberger, E. B.; Buchanan, L. E.; Marek, P.; Cao, P.; Raleigh, D. P.; Zanni, M. T., Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. *J. Am. Chem. Soc.* **2012,** *134* (30), 12658-12667.

56.   Li, H.; Lin, T.-Y.; Van Orden, S. L.; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B., Use of top-down and bottom-up Fourier transform ion cyclotron resonance mass spectrometry for mapping calmodulin sites modified by platinum anticancer drugs. *Anal. Chem.* **2011,** *83* (24), 9507-9515.

57.   Breuker, K.; McLafferty, F. W., Native electron capture dissociation for the structural characterization of noncovalent interactions in native cytochrome c. *Angew. Chem. Int. Ed.* **2003,** *42* (40), 4900-4904.

58.   Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, U. A.; Tsybin, Y. O., Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.* **2013,** *42* (12), 5014-5030.

59.   Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A., Top-down ESI-ECD-FT-ICR mass spectrometry localizes noncovalent protein-ligand binding sites. *J. Am. Chem. Soc.* **2006,** *128* (45), 14432-14433.

60.   Abedini, A.; Singh, G.; Raleigh, D. P., Recovery and purification of highly aggregation-prone disulfide-containing peptides: Application to islet amyloid polypeptide. *Anal. Biochem.* **2006,** *351* (2), 181-186.

61.   Caravatti, P.; Allemann, M., The 'Infinity Cell': a New Trapped-ion Cell With Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991,** *26* (5), 514-518.

62.   de Koning, L. J.; Nibbering, N. M. M.; van Orden, S. L.; Laukien, F. H., Mass selection of ions in a Fourier transform ion cyclotron resonance trap using

correlated harmonic excitation fields (CHEF). *Int. J. Mass spectrom.* **1997,** *165*, 209-219.

63. Horn, D. M.; Ge, Y.; McLafferty, F. W., Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal. Chem.* **2000,** *72* (20), 4778-4784.

64. Chan, F. T. S.; Schierle, G. S. K.; Kumita, J. R.; Bertoncini, C. W.; Dobson, C. M.; Kaminski, C. F., Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst* **2013,** *138* (7), 2156-2162.

65. Yang, T.; Wang, X. H.; Zhang, C. L.; Ma, X.; Wang, K.; Wang, Y. Q.; Luo, J.; Yang, L.; Yao, C.; Wang, X. Y., Specific self-monitoring of metal-associated amyloid-beta peptide disaggregation by a fluorescent chelator. *Chem. Commun.* **2016,** *52* (11), 2245-2248.

66. Morriss-Andrews, A.; Shea, J. E., Simulations of Protein Aggregation: Insights from Atomistic and Coarse-Grained Models. *J. Phys. Chem. Lett.* **2014,** *5* (11), 1899-1908.

67. del Mercato, L. L.; Pompa, P. P.; Maruccio, G.; Della Torre, A.; Sabella, S.; Tamburro, A. M.; Cingolani, R.; Rinaldi, R., Charge transport and intrinsic fluorescence in amyloid-like fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (46), 18019-18024.

68. Eisele, Y. S.; Monteiro, C.; Fearns, C.; Encalada, S. E.; Wiseman, R. L.; Powers, E. T.; Kelly, J. W., Targeting protein aggregation for the treatment of degenerative diseases. *Nat. Rev. Drug Discov.* **2015,** *14* (11), 759-780.

69. Sweeney, P.; Park, H.; Baumann, M.; Dunlop, J.; Frydman, J.; Kopito, R.; McCampbell, A.; Leblanc, G.; Venkateswaran, A.; Nurmi, A.; Hodgson, R., Protein misfolding in neurodegenerative diseases: implications and strategies. *Transl. Neurodegener.* **2017,** *6*, 13.

70. Robinson, N. E.; Robinson, A. B., Prediction of protein deamidation rates from primary and three-dimensional structure. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98* (8), 4367-4372.

2.8 Supplementary Information
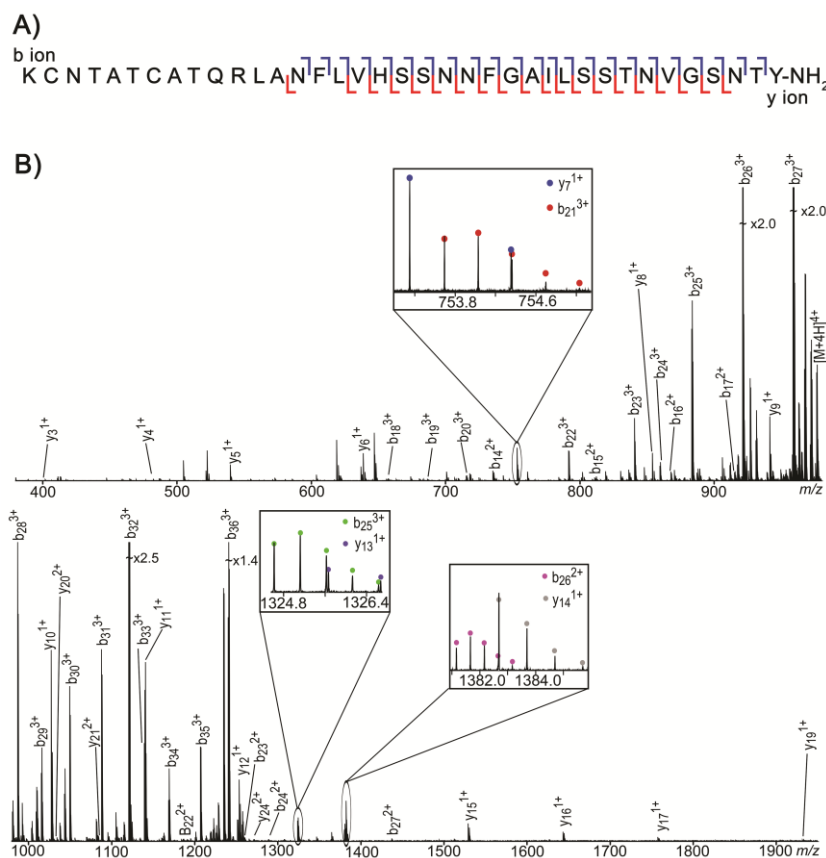
Figure S2.1. (A) Summarised b- and y-ion fragments observed in the (B) 18 V CAD MS/MS spectrum of the 4+ charge state hIAPP monomer. The hIAPP sequence coverage from the CAD MS/MS was 65% and showed effective fragmentation and production of b and y ions from Asn-14 to Try-37; while no fragments were obtained from Lys-1 to Ala-13. Water and ammonia (NH_3) loss are not labelled in the spectrum, but were common. The inserts show examples of isotopic distribution peaks, i.e. the b_{25}^{2+} and y_{13}^{1+} ions, which must be resolved with high resolution mass spectrometry in order to accurately assign fragment ions. The mass difference between the second isotopic peak of y_7 ion and the third isotopic peak of b_{21} ion was 0.0099 Da, and the resolving power at *m/z* 753 was 240,000. The assigned fragments are listed in the Supplementary Information Table S2.1.

The table of assigned CAD MS/MS fragments of the 4+ hIAPP monomer

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
MH	4+	976.223168	976.223160	0.008
b14	2+	737.848058	737.848010	0.065
b15	2+	811.382265	811.382370	-0.129
b16	2+	867.824297	867.824300	-0.003
b17	2+	917.458504	917.458500	0.004
b18	2+	985.987960	985.988030	-0.071
b18	3+	657.661065	657.661150	-0.129
b19	3+	686.671741	686.671520	0.322
b20	3+	715.682418	715.682430	-0.017
b21	3+	753.696727	753.696740	-0.017
b22	2+	1187.062917	1187.063020	-0.087
b22	3+	791.711036	791.711020	0.020
b23	2+	1260.597124	1260.597440	-0.251
b23	3+	840.733841	840.733860	-0.023
b24	2+	1289.107856	1289.107970	-0.088
b24	3+	859.740996	859.740930	0.077
b25	2+	1324.626413	1324.626300	0.085
b25	3+	883.420034	883.420040	-0.007
b26	2+	1381.168445	1381.168610	-0.119
b26	3+	921.114722	921.114690	0.035
b27	2+	1437.710477	1437.710670	-0.134
b27	3+	958.809410	958.809420	-0.010
b28	3+	987.820086	987.820110	-0.024
b29	3+	1016.830762	1016.830790	-0.028
b30	3+	1050.513322	1050.513220	0.097
b31	3+	1088.527631	1088.527500	0.120
b32	3+	1121.550436	1121.550350	0.077
b33	3+	1140.557591	1140.557563	0.025
b34	3+	1169.568267	1169.568320	-0.045
b35	3+	1207.582576	1207.582550	0.022
b36	3+	1241.265136	1241.265380	-0.197
y3	1+	396.187761	396.187760	0.003
y4	1+	483.219790	483.219820	-0.062
y5	1+	540.241254	540.241230	0.044
y6	1+	639.309668	639.309610	0.091
y7	1+	753.352596	753.352530	0.088
y8	1+	854.400275	854.400260	0.018
y9	1+	941.432304	941.432240	0.068
y10	1+	1028.464333	1028.464310	0.022
y11	1+	1141.548397	1141.548340	0.050
y12	1+	1254.632461	1254.632790	-0.262
y13	1+	1325.669575	1325.669660	-0.064
y14	1+	1382.691039	1382.690920	0.086
y15	1+	1529.759453	1529.759470	-0.011
y16	1+	1643.802381	1643.802440	-0.036
y17	1+	1757.845309	1757.844940	0.210
y19	1+	1931.909367	1931.908680	0.356
y20	2+	1034.987778	1034.987630	0.143
y21	2+	1084.521985	1084.521550	0.401
y24	2+	1271.619688	1271.619710	-0.017
			Average Error:	0.014
			Absolute Averge Error	0.087
			Standard Deviation:	0.128

Table S2.1. List of the assigned CAD MS/MS fragments of the 4+ charge state hIAPP monomer.
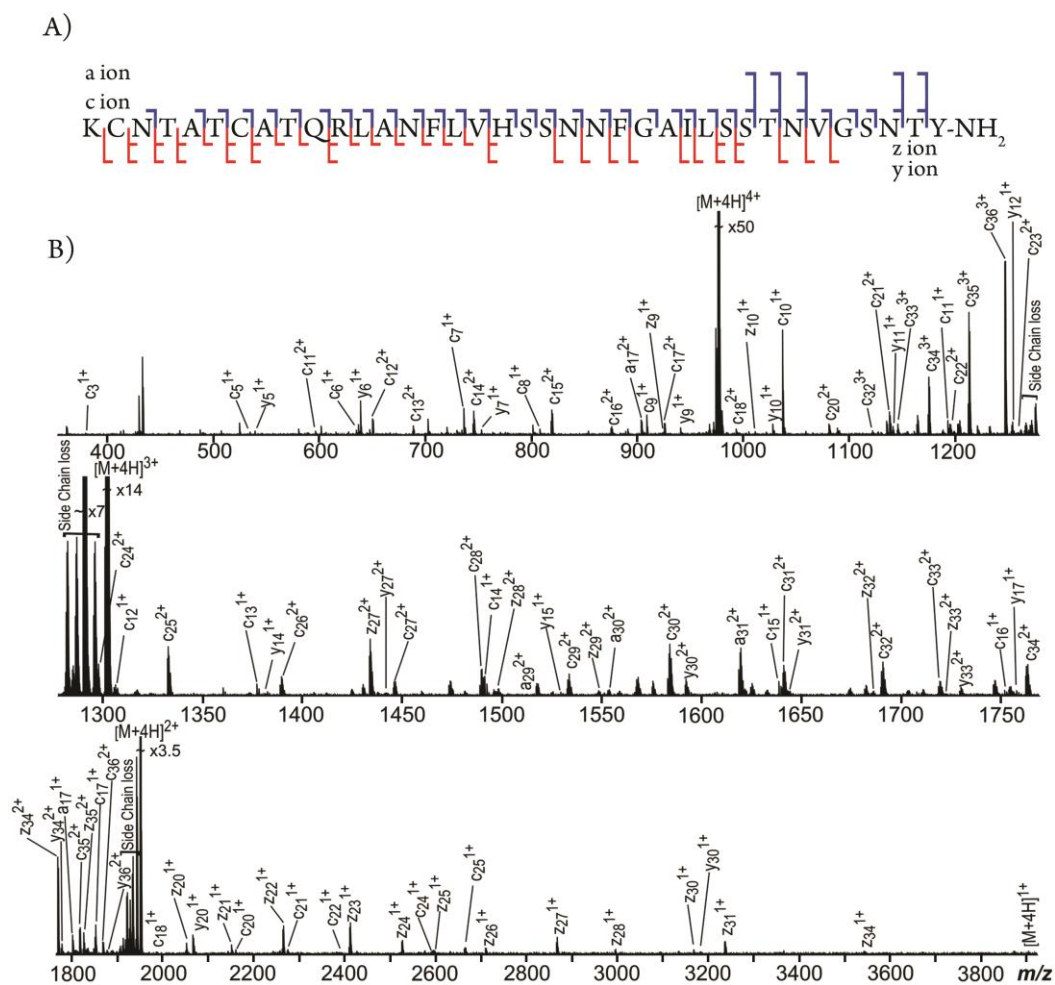
Figure S2.2. (A) Summarised a-, c-, y- and z-ion fragments observed in the (B) IR-ECD MS/MS spectrum of the 4+ charge state hIAPP monomer. The hIAPP sequence coverage from the IR-ECD MS/MS was 100%. Amino acid side chain losses are not labelled in the spectrum, but were common. The assigned fragments are listed in the Supplementary Information Table S2.2.

The table of assigned IR-ECD MS/MS fragments of the 4+ hIAPP monomer

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
MI	2+	1951.43911	1951.43866	0.231
MI	3+	1301.29518	1301.29502	0.123
MI	4+	976.22322	976.22328	-0.065
a17	1+	1807.93112	1807.93134	-0.120
a17	2+	903.96531	903.96539	-0.087
a29	2+	1511.75323	1511.75176	0.970
a30	2+	1561.77315	1561.77194	0.777
a31	2+	1618.79462	1618.79380	0.505
a35	3+	1198.58713	1198.58681	0.269
a36	3+	1232.26969	1232.26908	0.496
c3	1+	362.17313	362.17312	0.014
c5	1+	535.26574	535.26570	0.080
c6	1+	635.30568	635.30560	0.131
c7	1+	737.30696	737.30696	-0.004
c8	1+	808.34407	808.34406	0.014
c9	1+	909.39175	909.39152	0.253
c10	1+	1037.45033	1037.45041	-0.079
c11	1+	1193.55144	1193.55160	-0.135
c11	2+	597.27938	597.27939	-0.013
c12	1+	1306.63550	1306.63612	-0.472
c12	2+	653.82141	653.82139	0.037
c13	1+	1377.67262	1377.67320	-0.423
c13	2+	688.83606	688.83613	-0.104
c14	1+	1491.71555	1491.71590	-0.238
c14	2+	745.85752	745.85764	-0.158
c15	1+	1638.78396	1638.78399	-0.019
c15	2+	819.39173	819.39174	-0.013
c16	1+	1751.86802	1751.86790	0.070
c16	2+	875.93376	875.93381	-0.055
c17	1+	1850.93644	1850.93660	-0.088
c17	2+	925.46797	925.46763	0.366
c18	1+	1987.99535	1987.99514	0.105
c18	2+	993.99742	993.99771	-0.287
c19	1+	2075.02738	2075.02573	0.794
c20	1+	2162.05941	2162.05959	-0.085
c20	2+	1081.02945	1081.02974	-0.265
c21	1+	2276.10234	2276.10147	0.380
c21	2+	1138.05092	1138.05099	-0.063
c22	1+	2390.14526	2390.14589	-0.262
c22	2+	1195.07238	1195.07235	0.027
c23	2+	1268.60659	1268.60608	0.401
c24	1+	2594.23514	2594.23517	-0.011
c24	2+	1297.11732	1297.11750	-0.138
c25	1+	2665.27226	2665.27400	-0.655
c25	2+	1332.63588	1332.63575	0.096
c26	2+	1389.17791	1389.17764	0.194
c27	2+	1445.71994	1445.71981	0.091
c28	2+	1489.23596	1489.23572	0.159
c29	2+	1532.75197	1532.75136	0.399
c30	2+	1583.27581	1583.27540	0.260
c31	2+	1640.29728	1640.29653	0.454
c32	2+	1689.83148	1689.83118	0.179
c32	3+	1127.22604	1127.22587	0.148
c33	2+	1718.84613	1718.84536	0.446
c33	3+	1145.89725	1145.89698	0.236
c34	2+	1762.36214	1762.36211	0.017

(Table S2.2. Continue)

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
c34	3+	1174.90793	1174.90780	0.107
c35	2+	1819.38360	1819.38333	0.151
c35	3+	1212.92224	1212.92203	0.170
c36	2+	1869.90744	1869.90769	-0.132
c36	3+	1246.94074	1246.94060	0.110
y5	1+	540.24130	540.24125	0.098
y6	1+	639.30972	639.30975	-0.056
y7	1+	753.35265	753.35263	0.020
y9	1+	941.43235	941.43252	-0.177
y10	1+	1028.46438	1028.46443	-0.047
y11	1+	1141.54845	1141.54859	-0.126
y12	1+	1254.63251	1254.63226	0.199
y14	1+	1382.69109	1382.69111	-0.016
y15	1+	1529.75950	1529.76034	-0.548
y17	1+	1757.84536	1757.84562	-0.149
y20	1+	2067.96050	2067.96070	-0.095
y27	2+	1441.73088	1441.73073	0.105
y30	1+	3182.59781	3182.59677	0.326
y30	2+	1591.80257	1591.80241	0.099
y31	2+	1643.30716	1643.30816	-0.609
y33	2+	1729.34956	1729.34844	0.645
y34	2+	1779.87340	1779.87317	0.127
y35	2+	1836.89486	1836.89417	0.376
y36	2+	1888.39945	1888.39807	0.732
z9	1+	924.40580	924.40610	-0.320
z10	1+	1011.43783	1011.43803	-0.195
z20	1+	2053.95743	2053.95716	0.131
z21	1+	2153.02584	2153.02496	0.410
z22	1+	2266.10991	2266.10964	0.118
z23	1+	2413.17832	2413.17751	0.336
z24	1+	2527.22125	2527.22156	-0.123
z25	1+	2598.25836	2598.25951	-0.441
z26	1+	2711.34243	2711.34193	0.183
z27	1+	2867.44354	2867.44323	0.107
z27	2+	1433.72152	1433.72140	0.083
z28	1+	2995.50212	2995.50392	-0.602
z28	2+	1497.75081	1497.75092	-0.075
z29	2+	1547.77074	1547.77091	-0.112
z30	1+	3167.58691	3167.58800	-0.344
z30	2+	1583.79321	1583.79456	-0.856
z31	1+	3236.60843	3236.60698	0.447
z32	2+	1685.31772	1685.31730	0.252
z33	2+	1721.84155	1721.84026	0.749
z34	1+	3541.71292	3541.71563	-0.766
z34	2+	1770.85621	1770.85563	0.327
z35	2+	1827.87767	1827.87733	0.188
			Average Error:	0.056
			Absolute Average Error:	0.245
			Standard Deviation:	0.326

Table S2.2. List of the assigned IR-ECD MS/MS fragments of the 4+ charge state hIAPP monomer.

Deamidated (D)₃hIAPP: KCNTATCATQ RLANFLVHSS DDFGAILSST NVGSDTY-NH₂

Deamidated (isoD)₃hIAPP: KCNTATCATQ RLANFLVHSS (isoD)(isoD)FGAILSST NVGS(isoD)TY-NH₂
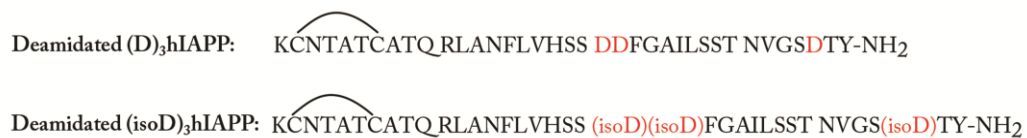
Figure S2.3. Sequences of synthetic mutant hIAPPs. Asn residues at position 21, 22, and 35 are replaced with aspartic acid – deamidated ((D)₃hIAPP) or isoaspartic acid – deamidated ((isoD)₃hIAPP) in order to act as deamidation mimics.
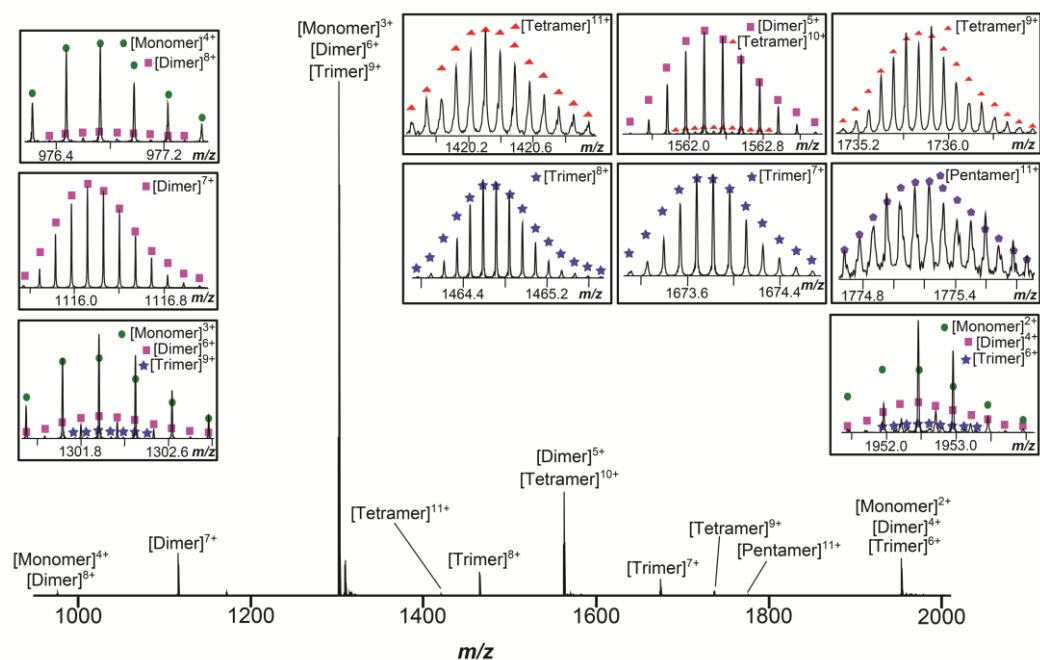
Figure S2.4. nESI-MS spectrum of the early oligomers from 10 μM hIAPP aqueous solution only. hIAPP monomer (circle), dimer (square), trimer (star), tetramer (triangle), and pentamer (pentagon), were detected in the spectrum. The peak intensity decreased gradually with increasing oligomer size. Inset, enlarge regions of spectra are shown for the labelled species. Coloured shapes represent theoretical calculated species overlaid onto observed pattern.
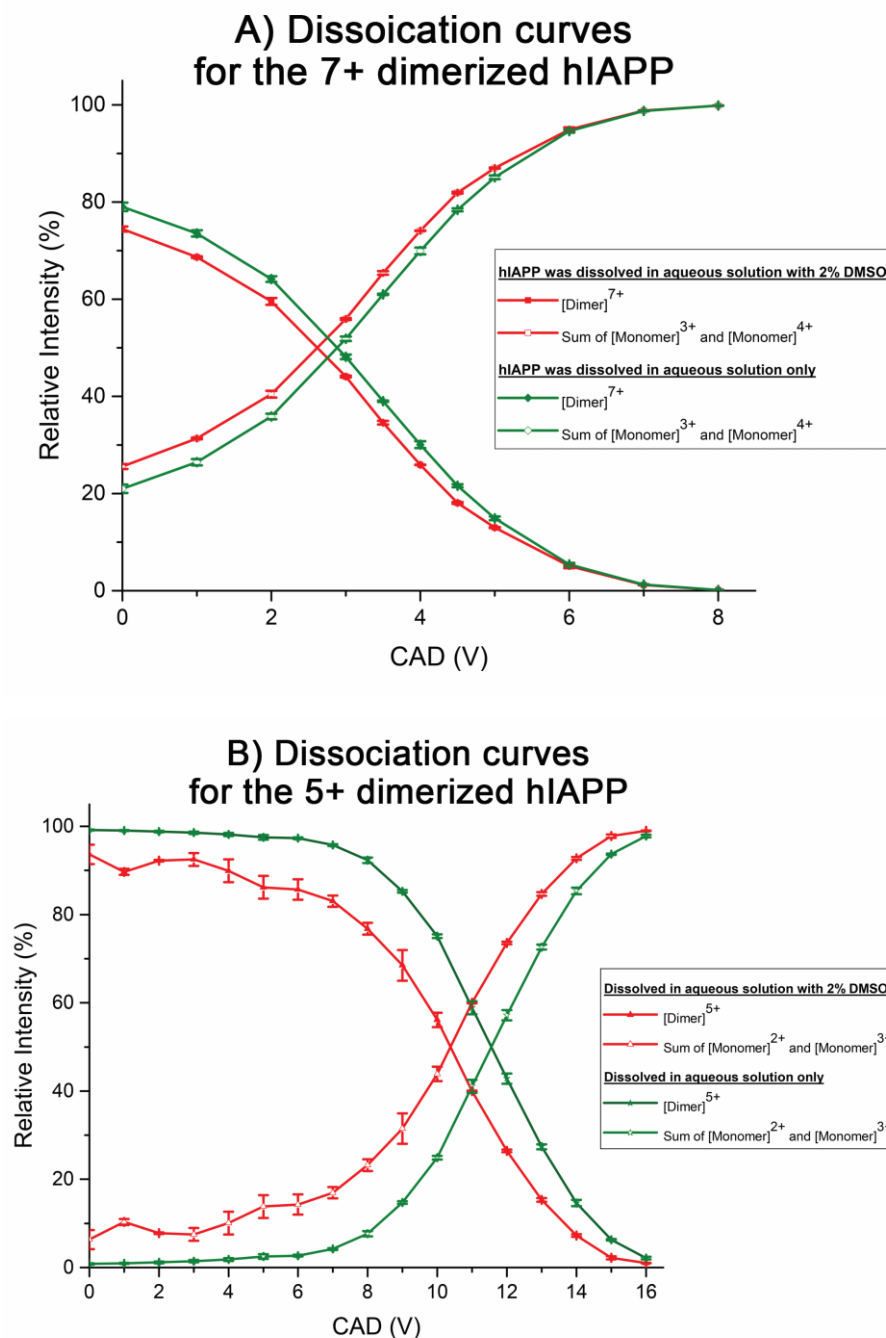
Figure S2.5. (A) The relative peak area of the monomer and the dimer ions during the dissociation of the 7+ charge state of hIAPP dimer in aqueous solution with 2% residual DMSO (red) and aqueous solution only (green) in the CAD MS/MS. (B) The relative peak area of the monomer and the dimer ions during the dissociation of the 5+ charge state of hIAPP dimer in aqueous solution with 2% residual DMSO (red) and aqueous solution only (green) in the CAD MS/MS.

The 7+ charge state of hIAPP dimer was dissociated into 3+/4+ monomer peaks, and the 5+ charge state of hIAPP dimer was dissociated into 2+/3+ monomer peaks, indicating near-symmetrical charge distribution between monomer units. The dissociation energy required to dissociate 50% of a given complex was used as a metric for comparison. The midpoint dissociation energies for the 5+ and the 7+ charge state dimers were 10.5 V and 3.4 V respectively. Excess CAD energy during dimer dissociation resulted in the monomer fragmentation which yielded b and y ions, and reduced the monomer relative peak area. Similar dissociation energies were observed from the 5+ and the 7+ charge state dimers in aqueous solution only.
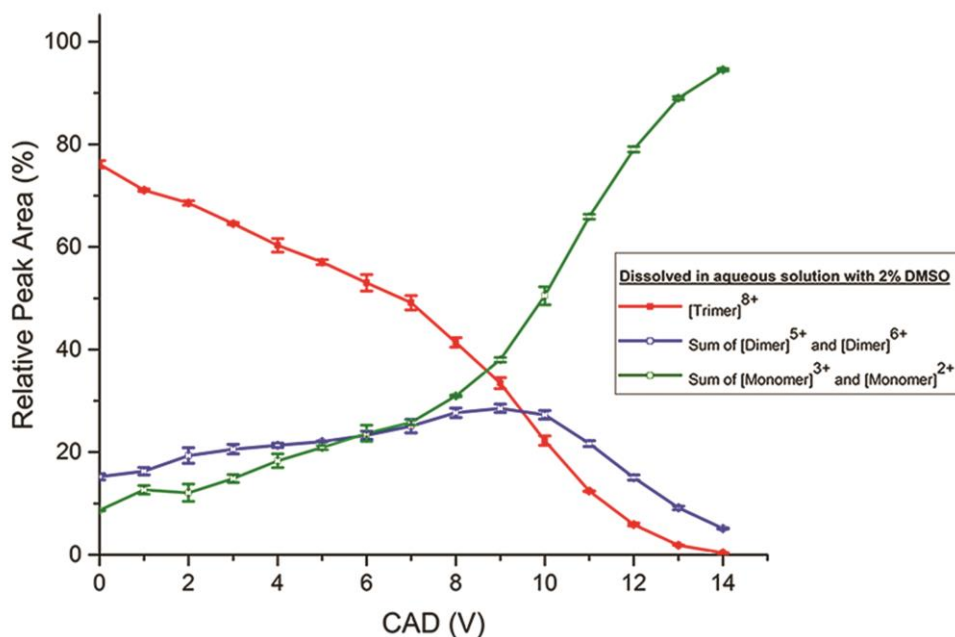
Figure S2.6. The relative peak areas of monomer (red), dimer (blue), and trimer (green) during the dissociation of the 8+ charge state of hIAPP trimer in aqueous solution with 2% residual DMSO in the CAD MS/MS. The midpoint dissociation energies of the 8+ charge state trimer is at 8.4 V. In the 8+ charge state trimer hIAPP, the dissociation of trimer hIAPP results in the increase of 5+ charge state hIAPP dimer (blue circle) from 0 V to 9 V; when the dissociation energy exceeds 10 V, the amount of dimer hIAPP decreases and results in a significant increase in the monomer hIAPP. The results demonstrate the oligomers of hIAPP are dissociating sequentially.
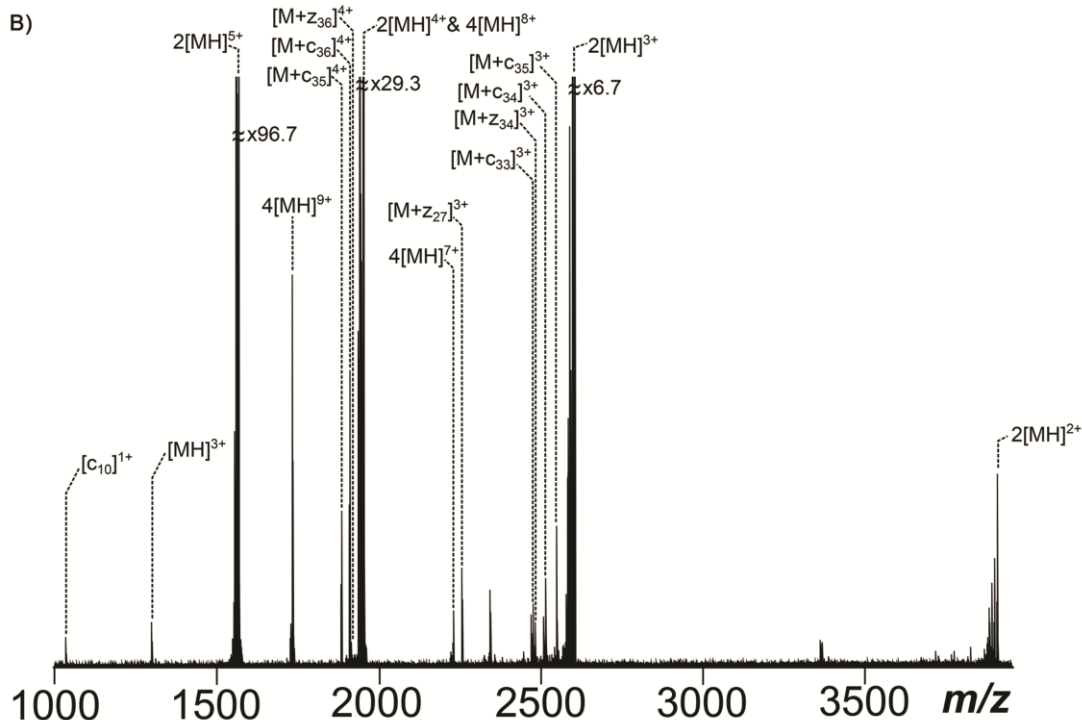
Figure S2.7. (A) Summarised c- and z-ion fragments observed in the (B) ECD MS/MS spectrum of the 5+ charge state hIAPP dimer (in aqueous solution with 2% residual DMSO). Only one monomer c-ion fragment (c10), three dimer z-ion fragments (representing the mass of a hIAPP plus a z ion fragment), and three dimer c-ion fragments (representing the mass of a hIAPP plus a c ion fragment) were detected from the ECD MS/MS spectrum of the 5+ hIAPP dimer ion which did not provide sufficient information for interrogating the interaction sites between hIAPP units. Highlighted sequence represents the proposed non-covalent interaction region. The amino acid side chain losses are not labelled in the spectrum but were common. The isotopic distribution of the 10+ charge state hIAPP tetramer was present and overlapped with the 5+ charge state hIAPP dimer during isolation, and higher order oligomers, i.e. tetramer, are observed in the

dimer ECD spectrum. The assigned fragments are listed in the Supplementary

Information Table S2.3.

The table of assigned ECD MS/MS fragments of the 5+ hIAPP dimer

Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
MI	3+	1301.963507	1301.963571	0.049
2[MI]	5+	1561.352698	1561.352462	-0.151
2[MI]	4+	1951.691010	1951.691642	0.324
2[MI]	3+	2603.591637	2603.591864	0.087
2[MI]	2+	3905.889909	3905.887841	-0.529
4[MI]	9+	1735.727472	1735.727310	-0.093
4[MI]	7+	2231.936089	2231.935654	-0.195
c10	1+	1037.450279	1037.450296	0.016
M+c33	3+	2477.204351	2477.203234	-0.451
M+c34	3+	2515.218663	2515.218503	-0.064
M+c35	4+	1886.413782	1886.414164	0.203
M+c35	3+	2549.237172	2549.237617	0.175
M+c36	4+	1911.675709	1911.675955	0.129
M+z27	3+	2257.776545	2257.776772	0.101
M+z34	3+	2482.868916	2482.869352	0.176
M+z36	4+	1916.414523	1916.414374	-0.078
			Average Error:	-0.019
			Absolute Average Error:	0.176
			Standard Deviation:	0.231

Table S2.3. List of the assigned ECD MS/MS fragments of the 5+ hIAPP dimer

ion (in aqueous solution with 2% residual DMSO). M represents one hIAPP unit.

Figure S2.8. Key IR-ECD fragments of (A) $[M+z_{22}]^{3+}$ (red), (B) $[M+z_{23}]^{3+}$ (blue), and (C) $[M+z_{24}]^{3+}$ (green) from the 5+ charge state hIAPP dimer with various IR pulse lengths. When no IR laser was applied, no dimer fragments from $M+z_{22}$ to $M+z_{24}$ were observed. When IR pulse length was set to 0.03 s, an extra dimer fragment, $M+z_{23}$, was shown (B). Dimer fragments from $M+z_{22}$ to $M+z_{24}$ were identified from the spectrum while IR pulse length was increased to 0.05 s ((A) and (C) respectively). The result from IR-ECD fragments of the 5+ hIAPP dimer demonstrates the binding region between the dimer hIAPP is relatively stronger near the C-terminus of the protein as the binding interaction region was first broken towards the center of hIAPP units (Asn-14 to Leu-16), so the binding region of individual hIAPPs units is proposed between Leu-16 and Ser-34, which contains 19 out of 37 residues

A) ECD Fragmentation of the 7+ hIAPP dimer

Monomer cleavages detected:

c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH$_2$

z ion
y ion

$[M+z_3]^{3+}$ $[M+c_{29}]^{5+}$

Dimer cleavages detected:

Dimer c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH$_2$

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH$_2$

Dimer z ion

1428.4   1429.2   1394.8   1395.6

m/z   *m/z*

B)

$[2M+7H]^{7+}$

x124

$[c_6]^{1+}$
$[y_6]^{1+}$

$[c_{22}]^{3+}$ $[c_{24}]^{3+}$
$[c_8]^{1+}$ $[c_{16}]^{2+}$

$[z_2]^{1+}$
$[c_3]^{1+}$
$[y_3]^{1+}$
$[y_4]^{1+}$
$[z_5]^{1+}$
$[c_{11}]^{2+}$
$[c_{12}]^{2+}$
$[c_{13}]^{2+}$
$[c_7]^{1+}$
$[c_{22}]^{3+}$
$[c_{15}]^{2+}$
$[c_{23}]^{3+}$
$[c_9]^{1+}$
$[c_{25}]^{3+}$
$[c_{17}]^{2+}$
$[c_{26}]^{3+}$
$[c_{27}]^{3+}$
$[c_{28}]^{3+}$
$[M+4H]^{4+}$
$[c_{18}]^{2+}$
$[c_{10}]^{1+}$
$[c_{19}]^{2+}$
$[c_{29}]^{3+}$
$[c_{30}]^{3+}$
$[c_{20}]^{2+}$
$[c_{31}]^{3+}$
$[c_{14}]^{2+}$

400   600   800   1000   *m/z*

135

(Table S2.9. Continue)

Figure S2.9. (A) Summary of c-, y-, and z-ion fragments observed in the (B) ECD MS/MS spectrum of the 7+ charge state hIAPP dimer (in aqueous solution with 2% residual DMSO). Highlighted sequence represents the proposed non-covalent interaction region. The side chain losses are not labelled for clarity. An increased number of fragments are observed in the ECD MS/MS spectrum of the 7+ charge state dimer than the 5+ charge state dimer spectrum. Ultrahigh resolution FTICR-MS provides the performance required to resolve the overlapping fragments observed, i.e. M+c_{35} and M+z_6 at *m/z* 1509.3, which are only differentiated by ~ 0.05 Da. The assigned fragments are listed in the Supplementary Information Table S2.4.

The table of assigned ECD MS/MS fragments of the 7+ hIAPP dimer

Ion	Charge	Theoretical	Experimental	Error (ppm)
MI	4+	976.223165	976.223805	0.656
MI	3+	1301.295128	1301.295174	0.035
MI	2+	1952.444365	1952.444262	-0.053
2[MI]	7+	1116.112577	1116.111471	-0.991
2[MI]	6+	1301.462261	1301.462328	0.051
2[MI]	5+	1561.956388	1561.956346	-0.027
2[MI]	4+	1953.196034	1953.195933	-0.052
c3	1+	362.173076	362.173094	0.050
c5	1+	534.257869	534.258045	0.329
c6	1+	635.305548	635.305514	-0.054
c7	1+	737.306908	737.307100	0.260
c8	1+	808.344022	808.344287	0.328
c9	1+	909.391701	909.391934	0.256
c10	1+	1037.450279	1037.450719	0.424
c11	2+	597.279333	597.279334	0.002
c11	1+	1193.551390	1193.551623	0.195
c12	2+	653.821365	653.821234	-0.201
c13	2+	689.339922	689.340001	0.115
c13	1+	1378.675984	1378.675689	-0.214
c14	2+	746.361386	746.361455	0.092
c15	2+	819.895593	819.895640	0.057
c16	2+	876.437625	876.437762	0.156
c17	2+	925.971832	925.971982	0.162
c18	2+	994.501527	994.501605	0.078
c19	2+	1038.017303	1038.017169	-0.129
c20	2+	1081.533317	1081.533376	0.055
c21	3+	759.706520	759.706974	0.598
c21	2+	1139.056142	1139.056500	0.314
c22	3+	797.386589	797.386802	0.267
c22	2+	1196.077601	1196.077549	-0.043
c23	3+	846.409393	846.409492	0.117
c23	2+	1269.110452	1269.110945	0.388
c24	3+	865.416548	865.416691	0.165
c25	3+	889.429834	889.430058	0.252
c25	2+	1333.139741	1333.139860	0.089
c26	3+	926.790272	926.790897	0.674
c26	2+	1390.183152	1390.183739	0.422
c27	3+	964.484962	964.484859	-0.107
c27	2+	1447.226592	1447.226424	-0.116
c28	3+	993.829896	993.830644	0.753
c28	2+	1490.241206	1490.241252	0.031
c29	3+	1022.840573	1022.840462	-0.109
c29	2+	1533.757222	1533.758500	0.833
c30	3+	1056.188874	1056.189084	0.199
c30	2+	1584.281062	1584.280913	-0.094
c31	3+	1094.203184	1094.203179	-0.005
c31	2+	1641.302522	1641.302019	-0.306
c32	3+	1127.225988	1127.225763	-0.200
c32	2+	1691.338138	1691.338733	0.352
c33	3+	1146.567401	1146.567371	-0.026
c33	2+	1720.350316	1720.350904	0.342
c34	3+	1176.246205	1176.246590	0.327
c34	2+	1762.362086	1762.362122	0.020
c35	3+	1213.258129	1213.258089	-0.033
c35	2+	1820.386340	1820.386545	0.113
c36	3+	1247.274945	1247.275773	0.664
c36	2+	1870.408779	1870.409286	0.271
y2	1+	282.144833	282.144822	-0.039
y3	1+	396.187761	396.187745	-0.040
y4	1+	483.219829	483.219862	0.068
y6	1+	639.309668	639.309557	-0.174
y7	1+	753.352596	753.352769	0.230
y8	1+	854.400275	854.400522	0.289
y30	2+	1592.303929	1592.304003	0.046
z2	1+	266.126109	266.126109	0.000
z5	1+	525.230353	525.230345	-0.015
z28	2+	1498.756089	1498.757430	0.895
z34	2+	1771.861482	1771.862629	0.647
M+c29	5+	1394.680928	1394.680459	-0.336
M+c31	5+	1438.300318	1438.300930	0.426

(Table S2.4. Continue)

Ion	Charge	Theoretical	Experimental	Error (ppm)
M+c32	5+	1457.511616	1457.510861	-0.518
M+c32	4+	1822.140215	1822.139354	-0.473
M+c33	5+	1469.519461	1469.519063	-0.271
M+c34	5+	1486.323445	1486.323359	-0.058
M+c35	5+	1509.332474	1509.331626	-0.562
M+c35	4+	1886.665738	1886.665069	-0.355
M+c36	5+	1529.942864	1529.942434	-0.281
M+z3	3+	1428.350791	1428.350757	-0.024
M+z5	3+	1476.704568	1476.704885	0.215
M+z6	3+	1509.058958	1509.059320	0.240
M+z7	3+	1547.741690	1547.741830	0.090
M+z7	2+	2320.607666	2320.607643	-0.010
M+z8	3+	1581.088313	1581.089390	0.681
M+z9	3+	1610.769064	1610.769547	0.300
M+z9	2+	2459.168869	2459.170923	0.835
M+z10	4+	1229.584072	1229.584641	0.463
M+z10	3+	1639.445612	1639.445934	0.196
M+z11	3+	1676.471867	1676.471173	-0.414
M+z11	2+	2515.206996	2515.206528	-0.186
M+z12	3+	1714.835007	1714.834497	-0.297
M+z13	3+	1738.514047	1738.514068	0.012
M+z13	2+	2607.771504	2607.772357	0.327
M+z14	4+	1317.890129	1317.890715	0.445
M+z14	3+	1757.521201	1757.520755	-0.254
M+z14	2+	2635.780806	2635.780280	-0.200
M+z15	4+	1355.158480	1355.158464	-0.012
M+z15	3+	1807.212338	1807.212131	-0.115
M+z16	4+	1382.917263	1382.917504	0.174
M+z16	3+	1844.892492	1844.892443	-0.027
M+z16	2+	2767.339383	2767.339020	-0.131
M+z17	4+	1412.430522	1412.430040	-0.341
M+z17	3+	1882.906802	1882.906777	-0.013
M+z18	4+	1433.436703	1433.436422	-0.196
M+z18	3+	1912.251644	1912.251179	-0.243
M+z19	4+	1455.695966	1455.695701	-0.182
M+z19	3+	1941.264102	1941.263992	-0.057
M+z20	4+	1489.960695	1489.960191	-0.338
M+z20	3+	1985.946098	1985.945390	-0.357
M+z21	4+	1514.727806	1514.727036	-0.508
M+z21	3+	2019.971451	2019.971805	0.175
M+z22	4+	1543.000167	1542.999155	-0.656
M+z22	3+	2058.000340	2058.000870	0.258
M+z23	5+	1264.014214	1264.014352	0.109
M+z23	4+	1579.013956	1579.014330	0.237
M+z23	3+	2106.354800	2106.354957	0.075
M+z24	4+	1607.776644	1607.776314	-0.205
M+z24	3+	2144.034920	2144.034484	-0.203
M+z25	4+	1626.287918	1626.287893	-0.015
M+z25	3+	2168.382333	2168.383232	0.415
M+z26	5+	1323.246008	1323.245622	-0.292
M+z26	4+	1654.308308	1654.307969	-0.205
M+z26	3+	2206.411232	2206.412492	0.571
M+z27	5+	1354.466227	1354.466407	0.133
M+z27	4+	1693.584213	1693.584379	0.098
M+z27	3+	2258.112487	2258.114584	0.929
M+z28	5+	1380.679464	1380.678457	-0.729
M+z28	4+	1725.598859	1725.599307	0.260
M+z29	5+	1401.089476	1401.088932	-0.388
M+z30	5+	1415.296902	1415.296093	-0.572
M+z30	4+	1768.870673	1768.870285	-0.219
M+z34	5+	1490.724486	1490.724988	0.337
M+z34	4+	1862.904657	1862.904749	0.049
M+z35	5+	1513.332634	1513.332401	-0.154
M+z35	4+	1891.665973	1891.665087	-0.468
M+z36	5+	1533.733951	1533.734428	0.311
M+z36	4+	1916.915649	1916.915513	-0.071
			Average Error:	0.052
			Absolute Average Error:	0.257
			Standard Deviation:	0.337

Table S2.4. List of the assigned fragments from the ECD MS/MS spectrum of the 7+ charge state hIAPP dimer.   M represents one hIAPP unit.

A) ECD Fragmentation of the 8+ hIAPP trimer

Monomer cleavages detected:

c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH₂

Dimer cleavages detected:

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH₂

Dimer c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH₂

Dimer z ion

$\star [2M+z_5]^{4+}$ $\star [2M+c_{34}]^{5+}$

Trimer cleavages detected:

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH₂

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH₂

Trimer c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y-NH₂

Trimer z ion

2083.2   2084.0   *m/z*

2267.2   2268.0   *m/z*

B)

$[M+3H]^{3+}$

$[c_{26}]^{3+}$

$[c_{19}]^{2+}$

$[c_{30}]^{3+}$

$[M+4H]^{4+}$

$[c_{10}]^{1+}$

$[c_3]^{1+}$

$[c_{11}]^{2+}$

$[c_{13}]^{2+}$

$[c_7]^{1+}$

$[c_{14}]^{2+}$

$[c_8]^{1+}$

$[c_{15}]^{2+}$

$[c_{23}]^{3+}$

$[c_9]^{1+}$

$[c_{17}]^{2+}$

$[c_{27}]^{3+}$

$[c_{18}]^{2+}$

$[c_{20}]^{2+}$

$[c_{32}]^{3+}$

$[c_{21}]^{2+}$

$[c_{11}]^{1+}$

$[c_{22}]^{2+}$

$[c_{23}]^{2+}$

$[c_{24}]^{2+}$

400   500   600   700   800   900   1000   1100   1200   *m/z*

(Figure S2.10. Continue)

Figure S2.10. (A) Summary of c-, y-, and z-ion fragments observed in the (B) ECD MS/MS spectrum of the 8+ charge state hIAPP trimer (in aqueous solution with 2% residual DMSO). Highlighted sequence represents the proposed non-covalent interaction region. The side chain losses are not labelled for clarity. The assigned fragments are listed in the Supplementary Information Table S2.5.

The table of assigned ECD MS/MS fragments of the 8+ hIAPP trimer

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
MI	4+	976.473865	976.473629	-0.242
MI	3+	1301.295128	1301.295286	0.121
[MI]	2+	1951.940453	1951.941145	0.355
2[MI]	5+	1562.356263	1562.357084	0.525
2[MI]	4+	1952.693510	1952.694205	0.356
2[MI]	3+	2603.925582	2603.925690	0.041
3[MI]	8+	1464.081791	1464.080743	-0.716
3[MI]	7+	1673.809188	1673.808501	-0.410
3[MI]	6+	1952.944506	1952.944143	-0.186
3[MI]	5+	2343.733943	2343.733157	-0.335
3[MI]	4+	2929.417034	2929.416922	-0.038
c3	1+	362.173076	362.173076	0.000
c7	1+	737.306908	737.306719	-0.256
c8	1+	808.344022	808.344125	0.127
c9	1+	909.391701	909.391441	-0.286
c10	1+	1037.450279	1037.450094	-0.178
c11	2+	597.279333	597.279295	-0.064
c11	1+	1193.551390	1193.551988	0.501
c13	2+	689.339922	689.339629	-0.425
c13	1+	1377.672568	1377.673088	0.377
c14	2+	746.361386	746.361097	-0.387
c14	1+	1491.715496	1491.715765	0.180
c15	2+	819.895593	819.895366	-0.277
c15	1+	1638.783910	1638.784035	0.076
c17	2+	925.971832	925.971525	-0.332
c18	2+	994.501288	994.501262	-0.026
c19	2+	1038.017303	1038.016936	-0.354
c20	2+	1081.533317	1081.533089	-0.211
c21	2+	1138.554781	1138.554661	-0.105
c22	2+	1195.576245	1195.576197	-0.040
c23	3+	846.409393	846.409301	-0.109
c23	2+	1269.110452	1269.110459	0.006
c24	2+	1297.621184	1297.620988	-0.151
c26	3+	926.790274	926.790027	-0.267
c26	2+	1890.183152	1890.183166	0.007
c27	3+	964.484962	964.484299	-0.687
c27	2+	1446.725192	1446.726334	0.789
c30	3+	1056.188874	1056.188926	0.049
c30	2+	1584.782465	1584.781451	-0.640
c32	3+	1127.560247	1127.560294	0.042
M+c35	4+	1886.161825	1886.160904	-0.488
M+c36	4+	1911.675709	1911.674109	-0.837
2M+c30	5+	2195.464155	2195.464715	0.255
2M+c34	5+	2267.699988	2267.699908	-0.035
2M+c35	7+	1635.791881	1635.791842	-0.024
2M+c35	6+	1908.423016	1908.422376	-0.335
2M+c35	5+	2289.906164	2289.907005	0.367
2M+c36	7+	1650.370440	1650.369669	-0.467
2M+c36	6+	1925.264301	1925.264063	-0.124
2M+c36	5+	2310.318271	2310.317485	-0.340
M+z23	3+	2106.687071	2106.687266	0.093

(Table S2.5. Continue)

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
2M+z5	4+	2083.497355	2083.498318	0.462
2M+z6	4+	2108.012511	2108.012947	0.207
2M+z7	5+	1709.019096	1709.019944	0.496
2M+z7	4+	2137.526829	2137.527322	0.231
2M+z10	5+	1764.642859	1764.643131	0.154
2M+z10	4+	2205.802305	2205.802133	-0.078
2M+z11	5+	1787.059229	1787.060986	0.983
2M+z11	4+	2235.076894	2235.077545	0.291
2M+z12	4+	2262.344371	2262.343377	-0.439
2M+z13	5+	1824.083932	1824.083442	-0.269
2M+z13	4+	2280.355611	2280.355687	0.033
2M+z14	5+	1835.287772	1835.287996	0.122
2M+z14	4+	2294.610979	2294.609671	-0.570
2M+z15	5+	1864.500986	1864.500270	-0.384
2M+z16	5+	1887.510052	1887.510169	0.062
2M+z16	4+	2359.888843	2359.889136	0.124
2M+z17	5+	1910.719554	1910.719165	-0.204
2M+z17	4+	2388.650128	2388.649501	-0.262
2M+z18	5+	1927.925508	1927.924985	-0.271
2M+z18	4+	2409.405841	2409.405482	-0.149
2M+z19	4+	2431.915600	2431.916230	0.259
2M+z20	5+	1972.944156	1972.943782	-0.190
2M+z20	4+	2466.181719	2466.182499	0.316
2M+z21	5+	1992.557390	1992.557247	-0.072
2M+z22	5+	2015.374671	2015.374524	-0.073
2M+z22	4+	2519.971551	2519.971167	-0.152
2M+z23	6+	1703.824077	1703.823394	-0.401
2M+z23	5+	2044.587908	2044.587680	-0.112
2M+z23	4+	2555.736390	2555.735772	-0.242
2M+z24	6+	1722.998291	1722.997816	-0.276
2M+z24	5+	2067.598059	2067.598981	0.446
2M+z24	4+	2584.496336	2584.495916	-0.163
2M+z26	6+	1753.685163	1753.686125	0.549
2M+z26	5+	2104.622773	2104.623063	0.138
2M+z26	4+	2630.777217	2630.776164	-0.400
2M+z27	6+	1779.367869	1779.367588	-0.158
2M+z27	5+	2135.842994	2135.842285	-0.332
2M+z27	4+	2669.803880	2669.802875	-0.376
2M+z28	6+	1801.212170	1801.210973	-0.665
2M+z28	5+	2161.254244	2161.254600	0.165
2M+z30	6+	1829.558849	1829.559634	0.429
2M+z30	4+	2745.842000	2745.843566	0.570
2M+z34	6+	1892.414312	1892.414141	-0.090
2M+z34	5+	2270.496380	2270.495936	-0.196
2M+z34	4+	2838.121979	2838.121431	-0.193
2M+z36	5+	2314.508068	2314.506366	-0.735
			Average Error:	-0.067
			Absolute Average Error:	0.280
			Standard Deviation:	0.343

Table S2.5. List of the assigned fragments from the ECD MS/MS spectrum of the 8+ charge state hIAPP trimer ion (in aqueous solution with 2% residual DMSO). M represents one hIAPP unit.
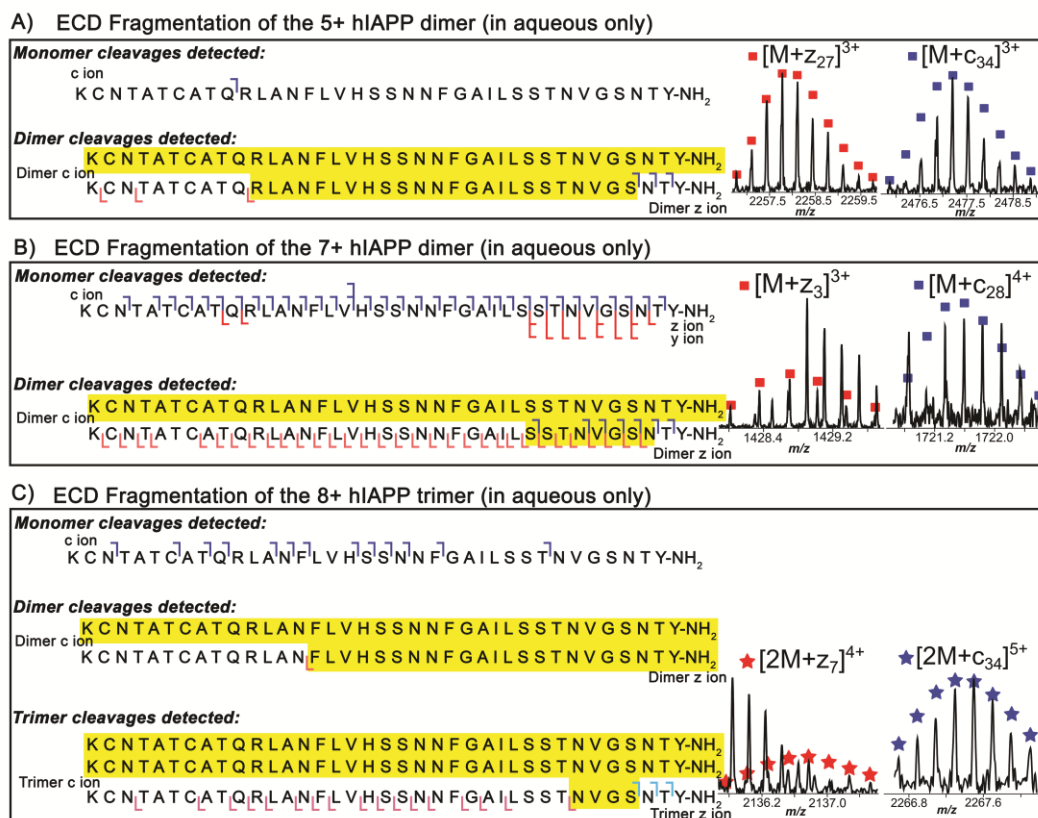
Figure S2.11. Summarised fragments observed in the ECD MS/MS spectrum of the (A) 5+ charge state dimer, (B) 7+ charge state dimer, and (C) 8+ charge state trimer of hIAPP, in aqueous solution. Examples of the smallest assigned dimer c-and z-ions from the ECD MS/MS of the 5+ and the 7+ charge state dimers are inserted, respectively. Examples of the most critical trimer c- and z-ion from the ECD MS/MS of the 8+ charge state trimer are also inserted. Highlighted sequence represents the proposed non-covalent interaction region. Coloured shapes represent theoretical calculated species overlaid onto observed pattern. The assigned fragments are listed in the Supplementary Information Tables S2.6, S2.7, and S2.8.

The table of assigned ECD MS/MS fragments of the 5+ hIAPP dimer
(in aqueous solution only)

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
[MI]	3+	1302.297541	1302.297961	0.323
2[MI]	5+	1561.553258	1561.552850	-0.261
2[MI]	4+	1952.693410	1952.695277	0.956
2[MI]	3+	2603.927471	2603.928753	0.492
2[MI]	2+	3906.890716	3906.890042	-0.173
4[MI]	9+	1735.727472	1735.727516	0.025
4[MI]	7+	2231.792930	2231.793086	0.070
c10	1+	1037.450279	1037.450277	-0.002
M+c34	3+	2477.204233	2477.204139	-0.038
M+c35	4+	1886.413773	1886.413868	0.050
M+c35	3+	2515.552587	2515.553103	0.205
M+c36	4+	1912.176753	1912.176749	-0.002
M+c36	3+	2549.237059	2549.237578	0.204
M+z27	3+	2257.442353	2257.441339	-0.449
M+z34	3+	2482.868892	2482.868192	-0.282
M+z36	4+	1916.665083	1916.666074	0.517
			Average Error:	0.102
			Absolute Average Error:	0.253
			Standard Deviation:	0.350

Table S2.6.   List of the assigned ECD MS/MS fragments of the 5+ charge state

hIAPP dimer (in aqueous solution only).   M represents one hIAPP unit.

The table of assigned ECD MS/MS fragments of the 7+ hIAPP dimer
(in aqueous solution only)

Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
[MI]	3+	1301.295128	1301.295155	0.021
[MI]	2+	1951.439053	1951.439594	0.277
2[MI]	7+	1116.112578	1116.112111	-0.418
2[MI]	6+	1302.131432	1302.130613	-0.629
2[MI]	5+	1562.557828	1562.556711	-0.715
2[MI]	4+	1953.197422	1953.197083	-0.174
2[MI]	3+	2604.597466	2604.599266	0.691
a17	2+	904.468923	904.469305	0.422
c3	1+	362.173076	362.173079	0.008
c5	1+	534.257869	534.257718	-0.283
c6	1+	635.305548	635.305497	-0.080
c7	1+	737.306908	737.306894	-0.019
c8	1+	808.344022	808.343993	-0.036
c9	1+	909.391701	909.391626	-0.082
c10	2+	519.228778	519.228684	-0.181
c10	1+	1037.450279	1037.450421	0.137
c11	2+	597.279333	597.279269	-0.107
c11	1+	1193.551390	1193.552256	0.726
c12	2+	653.821365	653.821253	-0.171
c13	2+	689.339922	689.339879	-0.062
c13	1+	1377.672568	1377.672392	-0.128
c14	2+	746.361386	746.361366	-0.027
c14	1+	1491.715496	1491.715364	-0.088
c15	2+	819.895593	819.895629	0.044
c16	2+	876.437625	876.437799	0.199
c17	2+	925.971832	925.971871	0.042
c18	2+	994.501288	994.501399	0.112
c19	2+	1038.017303	1038.017367	0.062
c20	2+	1081.533317	1081.533380	0.058
c21	3+	759.372279	759.372332	0.070
c21	2+	1138.554781	1138.554898	0.103
c22	3+	797.386589	797.386535	-0.068
c22	2+	1195.576245	1195.576726	0.402
c23	3+	846.409393	846.409348	-0.053
c23	2+	1269.110452	1269.110661	0.165
c24	3+	865.416548	865.416434	-0.132
c25	3+	889.095586	889.095288	-0.335
c26	3+	926.790274	926.790073	-0.217
c26	2+	1389.681773	1389.681581	-0.138
c27	3+	964.484962	964.484562	-0.415
c27	2+	1446.223805	1446.224334	0.366
c28	3+	993.495638	993.495502	-0.137
c28	2+	1490.241206	1490.242028	0.552
c29	3+	1023.174607	1023.174961	0.346
c29	2+	1534.759192	1534.760089	0.584
c30	3+	1056.188874	1056.188728	-0.138
c30	2+	1584.281062	1584.281747	0.432
c31	3+	1094.203184	1094.203292	0.099
c31	2+	1641.305046	1641.305402	0.217
c32	3+	1127.225988	1127.225585	-0.358
c32	2+	1691.841749	1691.841673	-0.045
c33	2+	1146.567401	1146.568161	0.663
c33	2+	1719.347463	1719.347416	-0.027
c34	2+	1175.578077	1175.578159	0.070
c34	2+	1762.362091	1762.362565	0.269
c35	3+	1213.926469	1213.926076	-0.324
c35	2+	1819.884937	1819.885330	0.216
c36	3+	1247.608904	1247.608555	-0.280
c36	2+	1871.410938	1871.411570	0.338
y3	1+	396.187761	396.187739	-0.056
y4	1+	483.219790	483.219859	0.143
y5	1+	540.241254	540.241149	-0.194
y6	1+	639.309668	639.309731	0.099
y7	1+	753.352596	753.352333	-0.349
y8	1+	854.400275	854.400411	0.159
y9	1+	941.432304	941.432472	0.178
z2	1+	266.126109	266.126109	0.000
z3	1+	381.176861	381.176838	-0.060
z5	1+	525.230353	525.230286	-0.128
z27	2+	1434.726804	1434.726864	0.042
z28	2+	1498.252176	1498.252684	0.339
M+c28	4+	1722.092874	1722.092573	-0.175
M+c31	5+	1437.899490	1437.898369	-0.780
M+c32	5+	1457.713182	1457.712567	-0.422
M+c32	4+	1822.641179	1822.641257	0.043
M+c33	5+	1469.319041	1469.319187	0.099
M+c33	4+	1835.895847	1835.896338	0.267
M+c34	5+	1486.523886	1486.523255	-0.424

(Table S2.7. Continue)

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
M+c34	4+	1857.903038	1857.903286	0.133
M+c34	3+	2477.206107	2477.206784	0.273
M+c35	6+	1257.944941	1257.944384	-0.443
M+c35	5+	1509.132031	1509.131650	-0.252
M+c35	4+	1886.413773	1886.413722	-0.027
M+c35	3+	2515.888529	2515.889693	0.463
M+c36	6+	1274.619189	1274.618504	-0.537
M+c36	5+	1530.344824	1530.344143	-0.445
M+z3	3+	1428.016652	1428.017814	0.814
M+z4	3+	1457.361471	1457.361442	-0.020
M+z5	3+	1476.370423	1476.369607	-0.553
M+z5	2+	2214.553213	2214.553546	0.150
M+z6	3+	1510.061465	1510.061814	0.231
M+z7	3+	1547.741690	1547.741300	-0.252
M+z7	2+	2321.108897	2321.107527	-0.590
M+z8	3+	1581.758350	1581.758113	-0.150
M+z9	3+	1610.434934	1610.434541	-0.244
M+z10	4+	1230.085200	1230.085218	0.015
M+z10	3+	1639.445612	1639.445189	-0.258
M+z11	3+	1676.471867	1676.471280	-0.350
M+z12	3+	1714.835007	1714.834779	-0.133
M+z13	3+	1738.514047	1738.513671	-0.216
M+z13	2+	2607.267432	2607.267132	-0.115
M+z14	4+	1318.391358	1318.390730	-0.476
M+z14	3+	1757.855327	1757.854802	-0.299
M+z14	2+	2636.279353	2636.279372	0.007
M+z15	4+	1355.158481	1355.158135	-0.255
M+z15	3+	1806.878157	1806.877692	-0.257
M+z16	4+	1383.669106	1383.669214	0.078
M+z16	3+	1844.892468	1844.892459	-0.005
M+z16	2+	2767.338976	2767.339318	0.124
M+z17	4+	1412.179946	1412.179831	-0.081
M+z17	3+	1882.906778	1882.907494	0.380
M+z17	2+	2824.861593	2824.862515	0.326
M+z18	4+	1433.937956	1433.938030	0.052
M+z18	3+	1911.917458	1911.917980	0.273
M+z19	4+	1455.695966	1455.695755	-0.145
M+z19	2+	2910.888569	2910.887702	-0.298
M+z20	4+	1489.960695	1489.961248	0.371
M+z20	3+	1986.280293	1986.280702	0.206
M+z21	4+	1514.727806	1514.727172	-0.419
M+z21	3+	2019.971374	2019.972282	0.450
M+z22	4+	1542.998831	1542.998568	-0.170
M+z22	3+	2057.331957	2057.332668	0.346
M+z23	5+	1264.214693	1264.214468	-0.178
M+z23	4+	1579.765949	1579.765670	-0.177
M+z23	3+	2105.352124	2105.352029	-0.045
M+z23	2+	3160.033644	3160.033403	-0.076
M+z24	4+	1608.276681	1608.276576	-0.065
M+z24	3+	2144.034920	2144.034716	-0.095
M+z25	4+	1626.035962	1626.035200	-0.469
M+z25	3+	2168.382264	2168.382341	0.036
M+z26	5+	1323.847526	1323.846832	-0.524
M+z26	4+	1654.306958	1654.306403	-0.335
M+z26	3+	2206.078772	2206.078781	0.004
M+z27	5+	1354.666757	1354.666205	-0.407
M+z27	4+	1693.584213	1693.583638	-0.340
M+z27	3+	2258.112467	2258.112424	-0.019
M+z28	5+	1380.679464	1380.679001	-0.335
M+z28	4+	1726.100055	1726.099418	-0.369
M+z29	5+	1400.889004	1400.888206	-0.570
M+z29	4+	1750.860782	1750.860731	-0.029
M+z30	5+	1414.895940	1414.895086	-0.604
M+z30	4+	1768.620062	1768.620698	0.360
M+z30	3+	2357.824324	2357.824847	0.222
M+z33	4+	1837.893258	1837.892451	-0.439
M+z34	5+	1490.123150	1490.122540	-0.409
M+z34	4+	1862.654075	1862.654369	0.158
M+z34	3+	2483.204833	2483.203905	-0.374
M+z35	5+	1512.931737	1512.932144	0.269
M+z35	4+	1891.164808	1891.165443	0.336
M+z36	5+	1533.331957	1533.330732	-0.799
M+z36	4+	1916.915620	1916.916114	0.258
			Average Error:	-0.044
			Absolute Average Error:	0.248
			Standard Deviation:	0.311

Table S2.7.   List of the assigned ECD MS/MS fragments of the 7+ charge state hIAPP dimer (in aqueous solution only).   M represents one hIAPP unit.

The table of assigned ECD MS/MS fragments of the 8+ hIAPP trimer
(in aqueous solution only)

Ion	Charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
[MI]	3+	1301.295128	1301.295525	0.305
[MI]	2+	1951.940453	1951.941533	0.553
2[MI]	5+	1562.356263	1562.356241	-0.014
2[MI]	4+	1952.693510	1952.693825	0.161
2[MI]	3+	2603.591528	2603.591786	0.099
3[MI]	8+	1464.582971	1464.582582	-0.266
3[MI]	7+	1673.952355	1673.951942	-0.247
3[MI]	6+	1952.944506	1952.944697	0.098
3[MI]	5+	2343.533517	2343.534255	0.315
3[MI]	4+	2929.417034	2929.416863	-0.058
c3	1+	363.180901	363.180901	0.000
c7	1+	737.306908	737.306979	0.096
c9	1+	909.391701	909.391753	0.057
c10	1+	1037.450279	1037.450632	0.340
c13	2+	689.339922	689.339981	0.086
c13	1+	1377.672568	1377.672453	-0.083
c14	2+	746.361386	746.361488	0.137
c14	1+	1491.715496	1491.716356	0.577
c15	2+	819.895593	819.895671	0.095
c15	1+	1638.783910	1638.784787	0.535
c18	2+	994.501288	994.501941	0.657
c19	2+	1038.017303	1038.017693	0.376
c20	2+	1081.533317	1081.533296	-0.019
c21	2+	1138.554781	1138.554888	0.094
c23	2+	1269.110452	1269.110367	-0.067
c30	3+	1056.857180	1056.857694	0.486
M+c14	4+	1349.651986	1349.652295	0.229
M+z23	3+	2105.685906	2105.684670	-0.587
2M+c34	5+	2267.699988	2267.699294	-0.306
2M+c35	7+	1635.791881	1635.790912	-0.592
2M+c35	6+	1908.423016	1908.423197	0.095
2M+c35	5+	2290.106592	2290.106669	0.034
2M+c36	7+	1650.370440	1650.369713	-0.441
2M+c36	6+	1925.264301	1925.265117	0.424
2M+c36	5+	2310.517700	2310.518213	0.222
2M+z7	4+	2136.773804	2136.773786	-0.008
2M+z11	4+	2234.575828	2234.573911	-0.858
2M+z13	4+	2280.103655	2280.102968	-0.301
2M+z14	5+	1835.287772	1835.287864	0.050
2M+z14	4+	2294.109849	2294.110242	0.171
2M+z16	5+	1887.510052	1887.510259	0.110
2M+z16	4+	2359.387700	2359.387809	0.046
2M+z17	5+	1910.719554	1910.718763	-0.414
2M+z17	4+	1388.149011	1388.149681	0.483
2M+z18	5+	1927.925508	1927.924537	-0.504
2M+z18	4+	2410.157588	2410.156797	-0.328
2M+z19	4+	2431.414510	2431.414302	-0.086
2M+z20	5+	1972.944156	1972.943100	-0.535
2M+z22	5+	2015.374671	2015.375993	0.656
2M+z22	4+	2519.469038	2519.470008	0.385
2M+z23	6+	1703.824077	1703.824082	0.003
2M+z23	5+	2044.587908	2044.588728	0.401
2M+z23	4+	2556.488124	2556.488923	0.313
2M+z24	6+	1722.998291	1722.999253	0.558
2M+z24	5+	2067.598059	2067.597329	-0.353
2M+z24	4+	2583.995165	2583.994963	-0.078
2M+z26	5+	2104.822114	2104.821212	-0.429
2M+z27	6+	1779.367869	1779.367687	-0.102
2M+z27	5+	2135.842994	2135.842633	-0.169
2M+z27	4+	2670.556970	2670.556856	-0.043
2M+z28	5+	2161.453606	2161.454306	0.324
2M+z30	6+	1829.558849	1829.558150	-0.382
2M+z34	6+	1892.414312	1892.413093	-0.644
2M+z34	5+	2271.499727	2271.499425	-0.133
			Average Error:	0.024
			Absolute Average Error:	0.275
			Standard Deviation:	0.348

Table S2.8.   List of the assigned fragments from the ECD MS/MS spectrum of the 8+ charge state hIAPP trimer ion (in aqueous solution only).   M represents one hIAPP unit.
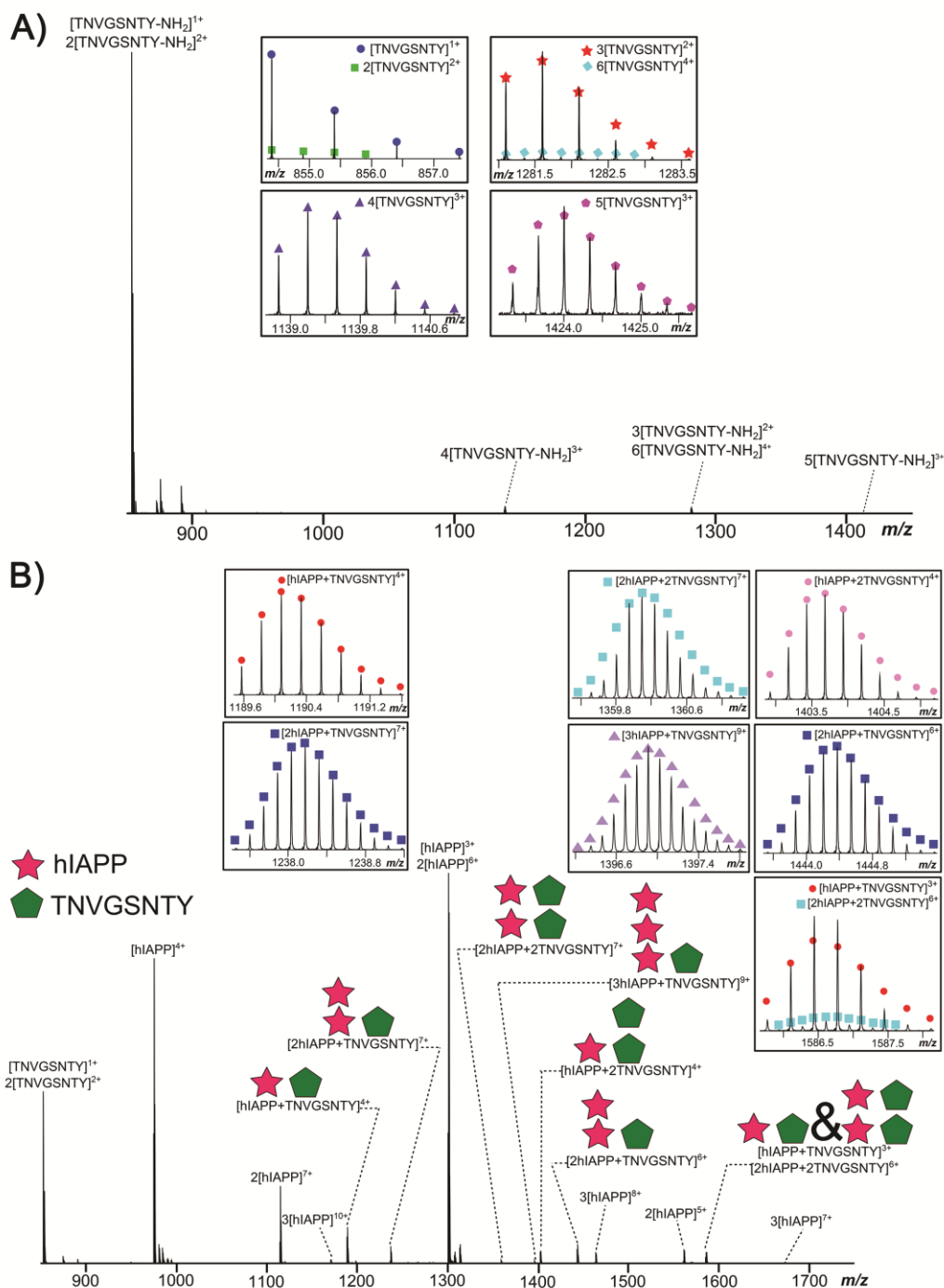
Figure S2.12. (A) MS spectrum of the early oligomers of an eight-residue segment of hIAPP (^{30}TNVGSNTY^{37}-NH$_2$) in aqueous solution. Oligomers; monomer (circle), dimer (square), trimer (star), tetramer (triangle), pentamer (pentagon), and hexamer (diamond), were detected in the MS spectrum. (B) MS spectrum of the mixture of the eight-residue segment of hIAPP (^{30}TNVGSNTY^{37}-NH$_2$) and hIAPP in 1:1 ratio. Apart from the expected early oligomers of hIAPP

and ^{30}TNVGSNTY^{37}-NH$_2$, mixtures of hIAPP and ^{30}TNVGSNTY^{37}-NH$_2$ were also detected in 1:1, 1:2, 1:3, 2:1, and 2:2 ratio respectively. Inset, enlarge regions of spectra are shown for the labelled species.   Coloured shapes represent theoretical calculated species overlaid onto observed pattern.
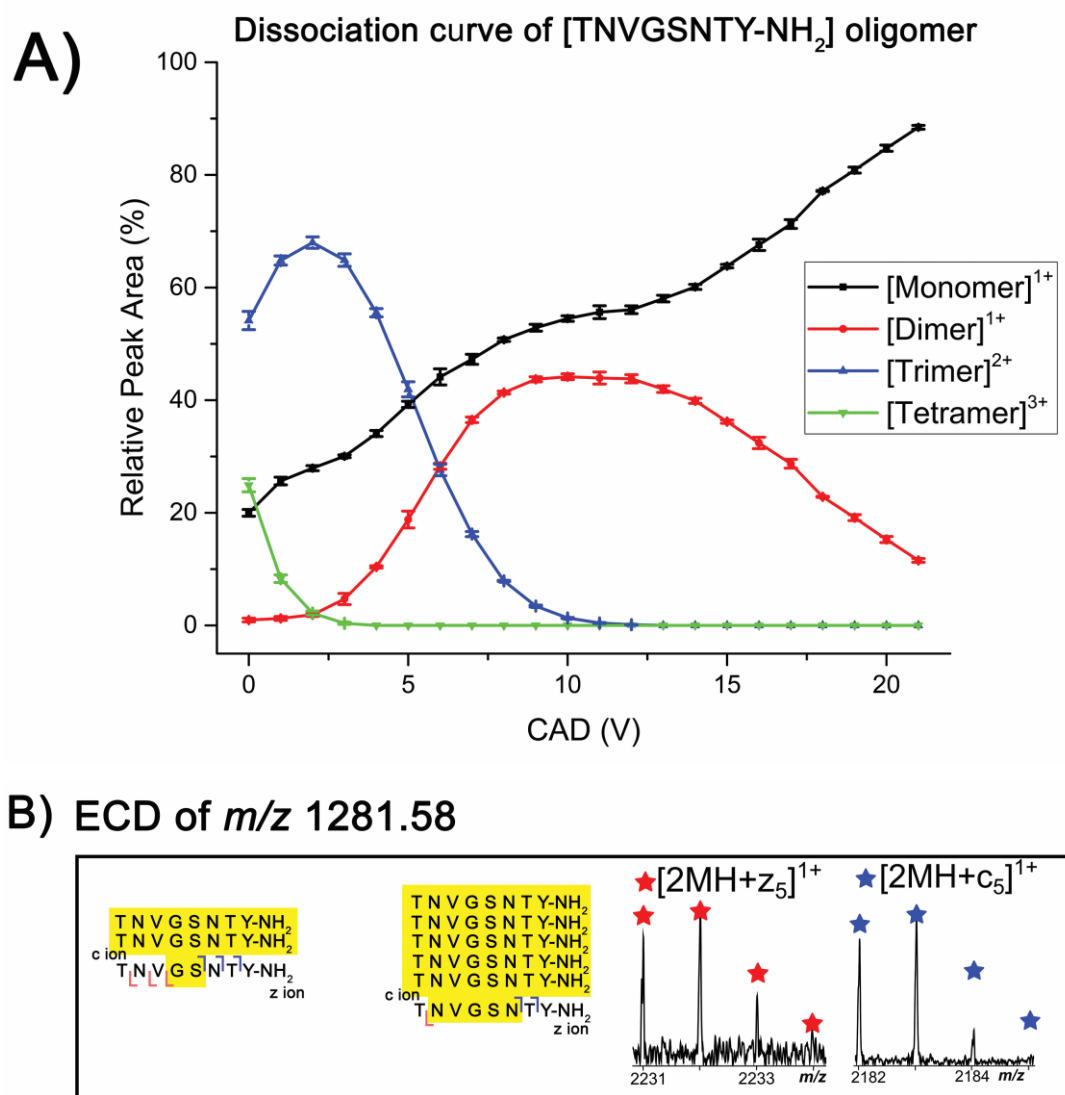
Figure S2.13. (A) The relative peak areas of monomer (black), dimer (red), trimer (blue), and tetramer (green) during the dissociation of the 3+ charge state tetramer of the eight-residue segment of hIAPP in the CAD MS/MS. The 3+ charge state tetramer is completely dissociated with CAD energy at 3 V. The relative peak area of the 2+ charge state trimer increases when CAD energy is increased from 0 V to 2 V and sharply reduces when the CAD energy exceeds 2 V. The relative peak area of the 1+ charge state dimer increases from 0 V to 10 V and begins to decrease afterwards. The relative peak area of the 1+ charge state monomer continuously increases with an increase of CAD energy. (B) Summary of the ECD MS/MS fragments of the 3+ charge state hIAPP segment in aqueous

solution. Examples of the most critical assigned trimer c- and z-ion are inserted. Highlighted sequence represents the proposed non-covalent interaction region. Coloured shapes represent theoretical calculated species overlaid onto observed pattern. The assigned fragments are listed in the Supplementary Information Table S2.9.

The table of assigned ECD MS/MS fragments of the 3+ ^{30}TNVGSNTY^{37}-NH$_2$ trimer

Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
2MH+c5	1+	2182.024503	2182.024212	-0.133
2MH+c6	1+	2297.075256	2297.074902	-0.154
2MH+c7	1+	2399.126351	2399.124459	-0.789
2MH+z5	1+	2232.011932	2232.011937	0.002
2MH+z6	1+	2330.076933	2330.074927	-0.861
2MH+z7	1+	2444.119860	2444.119499	-0.148
5MH+c6	3+	1620.088832	1620.088636	-0.121
5MH+c7	3+	1654.441566	1654.442594	0.621
5MH+z7	3+	1669.439658	1669.439893	0.141
3[MH]	2+	1281.096770	1281.096595	-0.137
3[MH]	1+	2561.185513	2561.186263	0.293
6[MH]	4+	1282.350153	1282.349911	-0.189
6[MH]	3+	1708.797730	1708.798167	0.256
MH	1+	854.400272	854.400285	0.015

			Average Error:	-0.086
			Absolute Average Error:	0.276
			Standard Deviation:	0.387

Table S2.9. List of the assigned fragments from the ECD MS/MS spectrum of the 3+ charge state trimer of the hIAPP segment. M represents one hIAPP unit.

Dissociation curve of [hIAPP+TNVGSNTY-NH$_2$] species
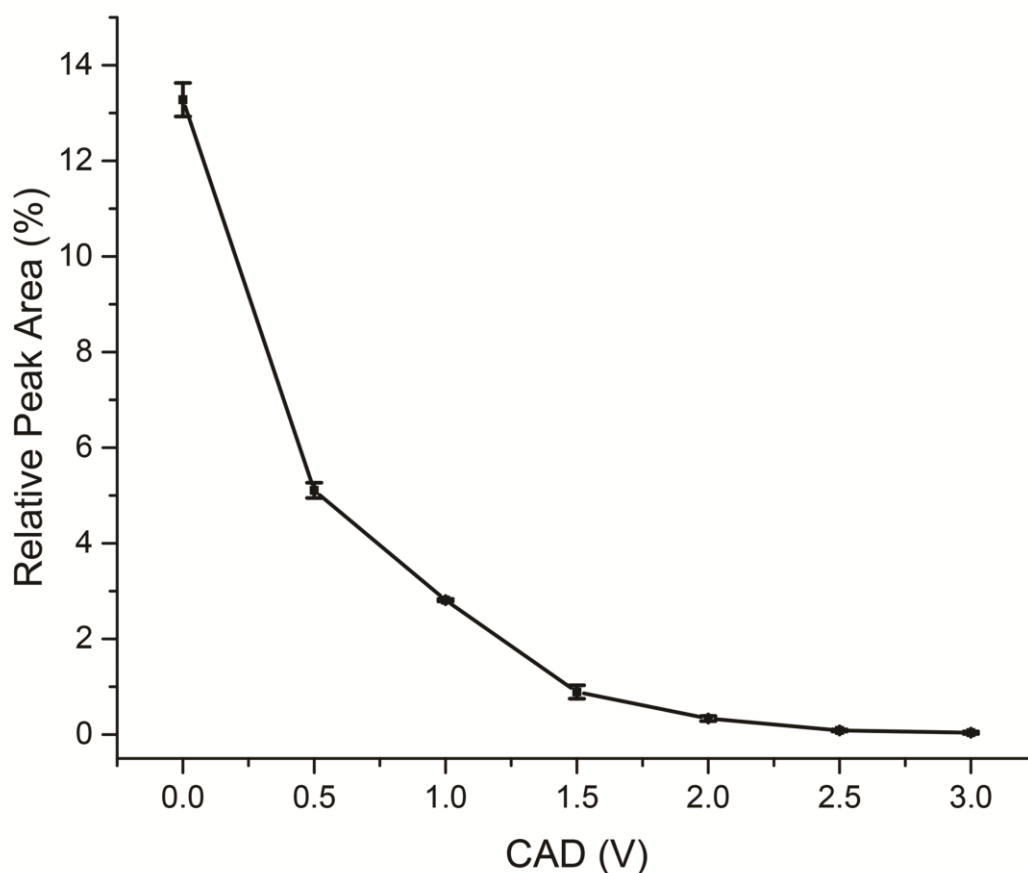
Figure S2.14. The CAD curve of the 4+ charge state hIAPP and ^{30}TNVGSNTY^{37}-NH$_2$ in 1:1 at *m/z* 1190. Only 13% of the mixture was successfully isolated and detected by the MS when no CAD energy was applied, which indicated the interaction between the biomolecules was very weak. The mixture was completely dissociated when 3 V of additional CAD energy was applied.
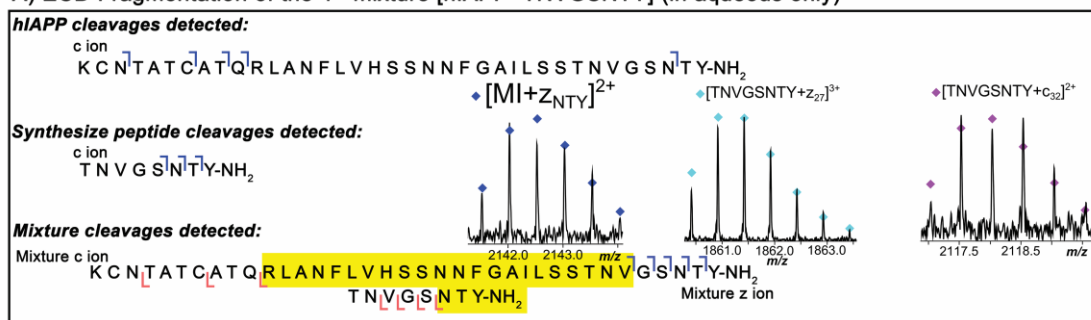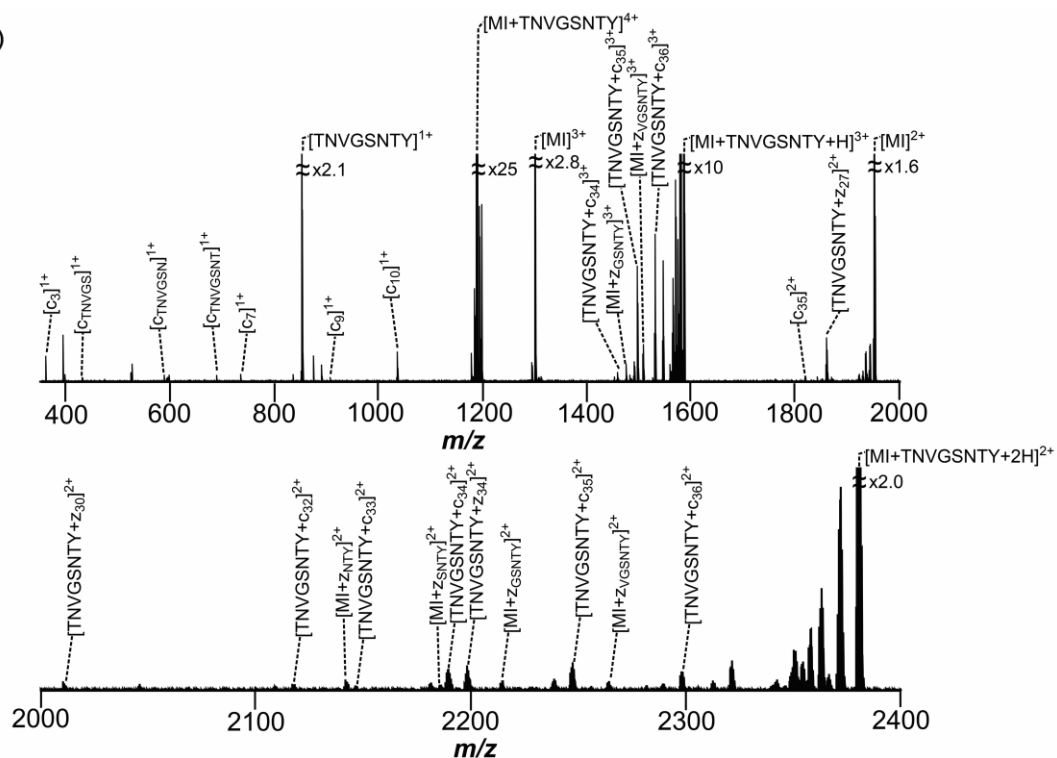
Figure S2.15. (A) Summary of c- and z-ion fragments observed in the (B) ECD MS/MS spectrum of the 4+ charge state mixture of [hIAPP + ^{30}TNVGSNTY^{37}-NH$_2$] species in aqueous solution. The most critical identified mixture fragments were hIAPP plus z ion of SNTY-NH$_2$, the mass of ^{30}TNVGSNTY^{37}-NH$_2$ plus c_{32} of hIAPP, and the mass of ^{30}TNVGSNTY^{37}-NH$_2$ plus z_{27} of hIAPP. The side chain losses are not labelled for clarity. The assigned fragments are listed in the Supplementary Information Table S2.10.

Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
[hIAPP]	3+	1301.295034	1301.295128	0.072
[hIAPP]	2+	1952.946261	1952.945652	-0.312
[hIAPP+TNVGSNTY]	4+	1189.571414	1189.571482	0.057
[hIAPP+TNVGSNTY]	3+	1586.095401	1586.095367	-0.021
[hIAPP+TNVGSNTY]	2+	2379.645083	2379.644330	-0.316
[TNVGSNTY]	1+	854.400272	854.400332	0.070
c3	1+	363.180901	363.180902	0.003
c7	1+	737.306908	737.306884	-0.033
c9	1+	909.391701	909.392057	0.391
c10	1+	1037.450279	1037.450525	0.237
c35	2+	1820.386836	1820.386597	-0.131
c ion of TNVGS	1+	476.246339	476.246325	-0.029
c ion of TNVGSN	1+	590.289267	590.289294	0.046
c ion of TNVGSNT	1+	691.336946	691.336919	-0.039
hIAPP+z ion of SNTY	2+	2185.541566	2185.540981	-0.268
hIAPP+z ion of GSNTY	3+	1476.370423	1476.369560	-0.585
hIAPP+z ion of GSNTY	2+	2214.553924	2214.552379	-0.698
hIAPP+z ion of VGSNTY	3+	1509.057288	1509.057173	-0.076
hIAPP+z ion of VGSNTY	2+	2263.586505	2263.586224	-0.124
TNVGSNTY+c32	2+	2118.536699	2118.536594	-0.050
TNVGSNTY+c33	2+	2147.049813	2147.050779	0.450
TNVGSNTY+c34	3+	1459.708148	1459.707114	-0.708
TNVGSNTY+c34	2+	2190.561113	2190.561344	0.105
TNVGSNTY+c35	3+	1497.386516	1497.386130	-0.258
TNVGSNTY+c35	2+	2246.581754	2246.581291	-0.206
TNVGSNTY+c36	3+	1531.405017	1531.405418	0.262
TNVGSNTY+c36	2+	2298.111132	2298.111262	0.057
TNVGSNTY+z27	2+	1861.421286	1861.421295	0.005
TNVGSNTY+z30	2+	2011.492975	2011.493855	0.437
TNVGSNTY+z34	2+	2198.056564	2198.055447	-0.508
			Average Error:	-0.072
			Absolute Average Error:	0.219
			Standard Deviation:	0.297

Table S2.10.   List of the assigned fragments from the ECD MS/MS spectrum of

the 4+ charge state of [hIAPP + ^{30}TNVGSNTY^{37}-NH$_2$] species.
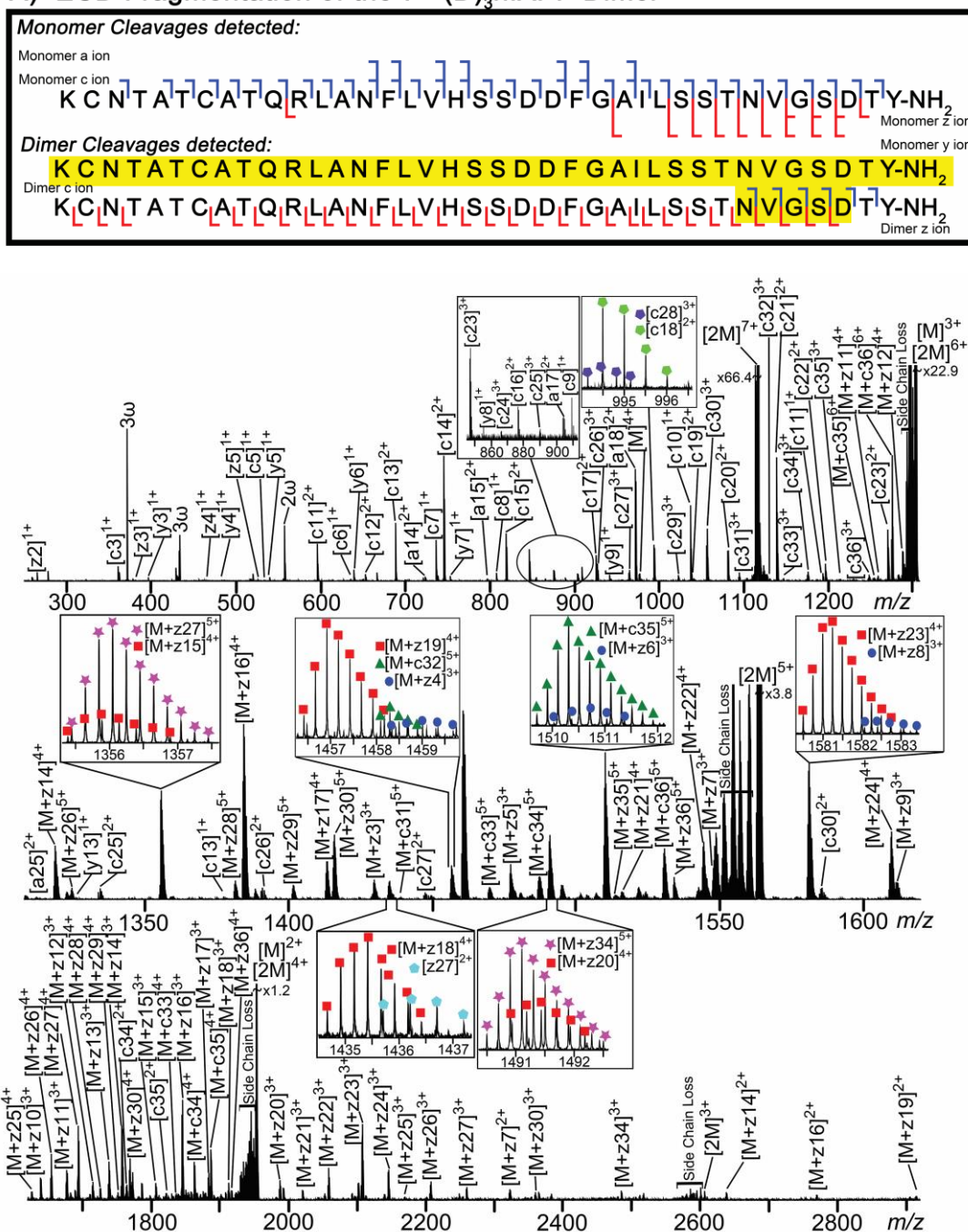
Figure S2.16. (A) Summary of a-, c-, y-, and z-ion fragments observed in the (B) ECD MS/MS spectrum of the 7+ charge state ((D)₃hIAPP) dimer (in aqueous solution with 2% residual DMSO). Highlighted sequence represents the proposed non-covalent interaction region. The side chain losses are not labelled for clarity. The assigned fragments are listed in the Supplementary Information Table S2.11.

The table of assigned ECD MS/MS fragments of the 7+ (D)$_3$hIAPP dimer (in aqueous solution with 2% DMSO)

Ion	charge	Theoretical m/z	Experimental m/z	Error (ppm)
MH	4+	977.462474	977.461786	-0.704
MH	3+	1302.279143	1302.279165	0.017
MH	2+	1952.915077	1952.915991	0.468
2[MH]	7+	1116.383162	1116.383047	-0.103
2[MH]	6+	1303.115465	1303.115560	0.073
2[MH]	5+	1563.738667	1563.738999	0.212
2[MH]	4+	1954.673471	1954.674185	0.365
2[MH]	3+	2605.897413	2605.897274	-0.053
a14	2+	724.354565	724.354485	-0.110
a15	2+	797.888772	797.888565	-0.259
a17	2+	903.965011	903.965096	0.094
a18	2+	972.494467	972.494518	0.052
a22	2+	1174.553438	1174.553758	0.272
a23	2+	1248.087645	1248.088012	0.294
a25	2+	1312.116934	1312.116940	0.005
c3	1+	362.173076	362.173008	-0.188
c5	1+	534.257869	534.257735	-0.251
c6	1+	636.313372	636.313252	-0.189
c7	1+	737.306908	737.306867	-0.056
c8	1+	808.344022	808.344002	-0.025
c9	1+	909.391701	909.391767	0.073
c10	1+	1037.450279	1037.450336	0.055
c11	2+	597.279333	597.279236	-0.162
c11	1+	1193.551390	1193.551711	0.269
c12	2+	653.821365	653.821279	-0.132
c13	2+	689.339922	689.339851	-0.103
c13	1+	1377.672568	1377.673273	0.512
c14	2+	746.361386	746.361348	-0.051
c15	2+	819.895593	819.895538	-0.067
c16	2+	876.437625	876.437608	-0.019
c17	2+	925.971832	925.971879	0.051
c18	2+	994.501288	994.501285	-0.003
c19	2+	1038.017303	1038.017350	0.045
c20	2+	1081.533317	1081.533397	0.074
c21	2+	1139.046789	1139.046866	0.068
c22	2+	1196.560261	1196.560497	0.197
c23	3+	847.065404	847.065401	-0.004
c23	2+	1270.094468	1270.094648	0.142
c24	3+	866.072559	866.072757	0.229
c25	3+	889.751597	889.751369	-0.256
c25	2+	1334.123849	1334.123808	-0.031
c26	3+	927.446285	927.446297	0.013
c26	2+	1390.665789	1390.665948	0.114
c27	3+	965.140973	965.141253	0.290
c27	2+	1447.709221	1447.709709	0.337
c28	3+	994.151649	994.151665	0.016
c29	3+	1023.162325	1023.162402	0.075
c29	2+	1535.242655	1535.243476	0.535
c30	3+	1056.844885	1056.844929	0.042
c30	2+	1584.763690	1584.764560	0.549
c31	3+	1094.859194	1094.859257	0.058
c31	2+	1642.286550	1642.287622	0.653
c32	3+	1127.881999	1127.882005	0.005
c32	2+	1692.823617	1692.824904	0.760
c33	3+	1146.889154	1146.889155	0.001
c34	3+	1175.899830	1175.899842	0.010
c34	2+	1763.346107	1763.347520	0.801
c35	3+	1214.576411	1214.576892	0.396
c35	2+	1820.859579	1820.859860	0.154
c36	3+	1247.924704	1247.924446	-0.207
y3	1+	397.171777	397.171687	-0.227
y4	1+	484.203806	484.203654	-0.314
y5	1+	541.225270	541.225155	-0.212
y6	1+	640.293684	640.293629	-0.086
y7	1+	754.336612	754.336600	-0.016
y8	1+	855.384291	855.384408	0.137
y9	1+	942.416320	942.416439	0.126
y10	1+	1029.448349	1029.448387	0.037
y13	1+	1326.653591	1326.653604	0.010
z2	1+	266.126109	266.126109	0.000
z3	1+	382.160877	382.160842	-0.092
z4	1+	467.177255	467.177105	-0.321
z5	1+	526.214369	526.214234	-0.257
z27	2+	1435.701402	1435.700996	-0.283
M+c31	5+	1438.883524	1438.884166	0.446
M+c32	5+	1458.496774	1458.496281	-0.338

(Table S2.11. Continue)

Ion	charge	Theoretical m/z	Experimental m/z	Error (ppm)
M+c32	4+	1822.618557	1822.618598	0.022
M+c33	5+	1469.700594	1469.701214	0.422
M+c33	4+	1837.376472	1837.376990	0.282
M+c34	5+	1487.107000	1487.107008	0.005
M+c34	4+	1858.883887	1858.885137	0.672
M+c35	6+	1258.761936	1258.761197	-0.587
M+c35	5+	1510.312868	1510.313214	0.229
M+c35	4+	1887.891222	1887.892061	0.444
M+c36	6+	1275.603219	1275.603240	0.016
M+c36	5+	1530.924420	1530.924741	0.210
M+z3	3+	1429.328685	1429.329245	0.392
M+z4	3+	1458.339361	1458.339476	0.079
M+z5	3+	1477.346514	1477.347034	0.352
M+z6	3+	1510.369320	1510.368816	-0.334
M+z7	3+	1548.049351	1548.049768	0.269
M+z7	2+	2323.077144	2323.078314	0.504
M+z8	3+	1582.400354	1582.401276	0.583
M+z9	3+	1611.746974	1611.746129	-0.524
M+z10	3+	1639.753258	1639.752527	-0.446
M+z11	4+	1258.839125	1258.839009	-0.092
M+z11	3+	1677.447952	1677.448418	0.278
M+z12	4+	1287.110147	1287.110616	0.364
M+z12	3+	1715.811104	1715.812415	0.764
M+z13	4+	1304.869427	1304.869520	0.071
M+z13	3+	1738.821678	1738.820724	-0.549
M+z14	4+	1318.623443	1318.623618	0.133
M+z14	3+	1757.828832	1757.829443	0.348
M+z14	2+	2637.748852	2637.749391	0.204
M+z15	4+	1355.390547	1355.390850	0.224
M+z15	3+	1807.856058	1807.856916	0.475
M+z16	4+	1384.147283	1384.147298	0.011
M+z16	3+	1846.532522	1846.533255	0.397
M+z16	2+	2769.799453	2769.799567	0.041
M+z17	4+	1412.904019	1412.904502	0.342
M+z17	3+	1884.874843	1884.874962	0.063
M+z18	4+	1434.912737	1434.913034	0.207
M+z18	3+	1913.885523	1913.885817	0.154
M+z19	4+	1456.420033	1456.420138	0.072
M+z19	3+	1942.896202	1942.895513	-0.355
M+z19	2+	2914.846456	2914.846774	0.109
M+z20	4+	1491.186124	1491.186700	0.386
M+z20	3+	1988.582507	1988.583390	0.444
M+z21	4+	1515.953231	1515.953433	0.133
M+z21	3+	2021.605320	2021.606146	0.409
M+z22	4+	1544.224252	1544.224820	0.368
M+z22	3+	2059.300020	2059.301231	0.588
M+z23	4+	1580.489984	1580.490360	0.238
M+z23	3+	2108.322843	2108.323867	0.486
M+z24	4+	1609.000716	1609.001073	0.222
M+z24	3+	2146.337153	2146.338370	0.567
M+z25	4+	1627.512008	1627.512728	0.442
M+z25	3+	2170.350335	2170.351553	0.561
M+z26	5+	1324.827880	1324.828098	0.165
M+z26	4+	1655.783031	1655.783641	0.368
M+z26	3+	2208.046833	2208.046628	-0.093
M+z27	5+	1356.048099	1356.048275	0.130
M+z27	4+	1695.060261	1695.060689	0.252
M+z27	3+	2260.080530	2260.081089	0.247
M+z28	5+	1381.659815	1381.660080	0.192
M+z28	4+	1726.824267	1726.824613	0.200
M+z29	5+	1402.069844	1402.070205	0.257
M+z29	4+	1752.086188	1752.086770	0.332
M+z30	5+	1416.076777	1416.077014	0.167
M+z30	4+	1770.096109	1770.096483	0.211
M+z30	3+	2359.792401	2359.792212	-0.080
M+z34	5+	1491.103517	1491.103695	0.119
M+z34	4+	1864.130123	1864.130663	0.290
M+z34	3+	2485.507093	2485.507949	0.344
M+z35	5+	1513.912103	1513.912030	-0.048
M+z36	5+	1534.713228	1534.713352	0.081
M+z36	4+	1918.141132	1918.141936	0.419
			Average Error:	0.124
			Absolute Average Error:	0.236
			Standard Deviation:	0.278

Table S2.11. List of the assigned fragments from the ECD MS/MS spectrum of the 7+ charge state ((D)$_3$hIAPP) dimer ion (in aqueous solution with 2% residual DMSO). M represents one ((D)$_3$hIAPP) unit.

Figure S2.17. (A) Summary of a-, c-, y-, and z-ion fragments observed in the (B) ECD MS/MS spectrum of the 7+ charge state ((isoD)$_3$hIAPP) dimer (in aqueous solution with 2% residual DMSO). Highlighted sequence represents the proposed non-covalent interaction region. The side chain losses are not labelled for clarity. The assigned fragments are listed in the Supplementary Information Table S2.12.

The table of assigned ECD MS/MS fragments of the 7+ (isoD)₃hIAPP dimer
(in aqueous solution with 2% DMSO)

Type	Ion	charge	Theoretical m/z	Experimental m/z	Error (ppm)
MH	MH	4+	977.211884	977.211677	-0.212
MH	MH	3+	1302.279143	1302.279080	-0.048
MH	MH	2+	1952.915077	1952.915655	0.296
2[MH]	2[MH]	7+	1116.383162	1116.383218	0.050
2[MH]	2[MH]	6+	1303.115465	1303.115597	0.101
2[MH]	2[MH]	5+	1563.738667	1563.738713	0.029
2[MH]	2[MH]	4+	1954.673471	1954.672757	-0.365
a	23	2+	1248.087645	1248.087793	0.119
a	24	2+	1277.099772	1277.099889	0.092
a	25	2+	1312.116934	1312.117109	0.133
c	3	1+	362.173076	362.173033	-0.119
c	6	1+	636.313372	636.313275	-0.152
c	7	1+	737.306908	737.306876	-0.043
c	8	1+	808.344022	808.343966	-0.069
c	9	1+	909.391701	909.391605	-0.106
c	10	1+	1037.450279	1037.450277	-0.002
c	11	2+	597.279333	597.279315	-0.030
c	11	1+	1193.551390	1193.551247	-0.120
c	12	2+	653.821365	653.821303	-0.095
c	13	2+	689.339922	689.339931	0.013
c	14	2+	746.361386	746.361372	-0.019
c	15	2+	819.895593	819.895597	0.005
c	16	2+	876.437625	876.437697	0.082
c	17	2+	925.971832	925.971846	0.015
c	18	2+	994.501288	994.501292	0.004
c	19	2+	1038.017303	1038.017376	0.070
c	20	2+	1081.533317	1081.533351	0.031
c	21	3+	759.700285	759.700251	-0.045
c	21	2+	1139.046789	1139.046820	0.027
c	22	3+	798.042599	798.042537	-0.078
c	22	2+	1196.560261	1196.560293	0.027
c	23	3+	847.065404	847.065386	-0.021
c	23	2+	1270.094468	1270.094418	-0.039
c	24	3+	866.072559	866.072652	0.107
c	25	3+	889.751597	889.751696	0.111
c	25	2+	1334.123849	1334.123757	-0.069
c	26	3+	927.446285	927.446286	0.001
c	26	2+	1390.665789	1390.666501	0.512
c	27	3+	965.140973	965.140963	-0.010
c	27	2+	1447.709221	1447.709778	0.385
c	29	3+	1023.162325	1023.162215	-0.108
c	29	2+	1534.239850	1534.239965	0.075
c	30	3+	1056.844885	1056.844848	-0.035
c	30	2+	1584.763690	1584.763724	0.021
c	31	3+	1094.859194	1094.859270	0.069
c	31	2+	1642.787956	1642.788594	0.388
c	32	3+	1127.881999	1127.881584	-0.368
c	32	2+	1691.319356	1691.320699	0.794
c	33	3+	1146.889154	1146.888901	-0.221
c	33	2+	1720.331490	1720.332143	0.380
c	34	3+	1175.899830	1175.899636	-0.165
c	34	2+	1763.346107	1763.346490	0.217
c	35	3+	1214.576411	1214.576568	0.129
c	35	2+	1820.859579	1820.860110	0.292
c	36	3+	1247.924704	1247.924462	-0.194
c	20i	2+	1110.032142	1110.032427	0.256
c	21i	2+	1167.545613	1167.545686	0.063
c	34i	3+	1195.233316	1195.233951	0.531
y	3	1+	397.171777	397.171827	0.126
y	4	1+	484.203806	484.203771	-0.072
y	5	1+	541.225270	541.225281	0.020
y	6	1+	640.293684	640.293709	0.039
y	7	1+	754.336612	754.336563	-0.065
y	9	1+	942.416320	942.416353	0.035
y	10	1+	1029.448349	1029.448645	0.288
z	2	1+	266.126109	266.126109	0.000
z	3	1+	382.160877	382.160841	-0.094
z	4	1+	469.192905	469.192914	0.019
z	5	1+	526.214369	526.214338	-0.059
z	6	1+	625.282783	625.282678	-0.168
z	27	2+	1435.197489	1435.196299	-0.829
M+c	31	5+	1438.482615	1438.482645	0.021
M+c	32	5+	1458.496774	1458.496042	-0.502
M+c	32	4+	1822.618557	1822.618534	-0.013
M+c	33	5+	1469.700594	1469.699673	-0.627
M+c	33	4+	1837.375068	1837.375877	0.440

(Table S2.12. Continue)

Type	Ion	charge	Theoretical *m/z*	Experimental *m/z*	Error (ppm)
M+c	34	5+	1487.105911	1487.106523	0.412
M+c	34	4+	1858.883887	1858.883845	-0.023
M+c	35	6+	1258.761936	1258.761954	0.014
M+c	35	5+	1510.312868	1510.312981	0.075
M+c	35	4+	1887.891222	1887.891506	0.150
M+c	36	6+	1275.603219	1275.603316	0.076
M+c	36	5+	1530.924420	1530.924203	-0.142
M+c	34i	5+	1498.707009	1498.707619	0.407
M+c	34i	4+	1873.131943	1873.132905	0.514
M+z	2	3+	1391.654617	1391.655808	0.856
M+z	3	3+	1429.328685	1429.329444	0.531
M+z	4	3+	1458.339361	1458.340006	0.442
M+z	5	3+	1477.680667	1477.680895	0.154
M+z	6	3+	1510.369320	1510.368775	-0.361
M+z	7	3+	1548.049351	1548.049654	0.196
M+z	7	2+	2323.578098	2323.578660	0.242
M+z	8	3+	1582.400354	1582.401094	0.468
M+z	9	3+	1611.746974	1611.746700	-0.170
M+z	10	4+	1230.568102	1230.567991	-0.090
M+z	10	3+	1639.753258	1639.753036	-0.135
M+z	11	4+	1258.839125	1258.839593	0.372
M+z	11	3+	1677.447952	1677.447828	-0.074
M+z	12	4+	1287.110147	1287.110909	0.592
M+z	12	3+	1715.811104	1715.811618	0.300
M+z	13	4+	1304.869427	1304.869611	0.141
M+z	13	3+	1738.821678	1738.821082	-0.343
M+z	14	4+	1318.623443	1318.623189	-0.193
M+z	14	3+	1757.828832	1757.829081	0.142
M+z	14	2+	2638.247419	2638.246892	-0.200
M+z	15	4+	1355.390547	1355.390795	0.183
M+z	15	3+	1807.856058	1807.856380	0.178
M+z	16	4+	1384.147283	1384.147411	0.092
M+z	16	3+	1846.532522	1846.532903	0.206
M+z	16	2+	2769.295145	2769.294672	-0.171
M+z	17	4+	1412.904019	1412.903882	-0.097
M+z	17	3+	1884.874843	1884.874708	-0.072
M+z	18	4+	1434.912737	1434.912984	0.172
M+z	18	3+	1913.885523	1913.885178	-0.180
M+z	19	4+	1456.420033	1456.420238	0.141
M+z	19	3+	1942.896202	1942.895988	-0.110
M+z	20	4+	1491.186124	1491.186435	0.209
M+z	20	3+	1988.582507	1988.582942	0.219
M+z	21	4+	1515.953231	1515.953379	0.098
M+z	21	3+	2021.605320	2021.605553	0.115
M+z	22	4+	1544.224252	1544.224504	0.163
M+z	22	3+	2059.300020	2059.300910	0.432
M+z	23	4+	1580.489984	1580.489808	-0.111
M+z	23	3+	2108.322843	2108.323448	0.287
M+z	24	4+	1609.000716	1609.000938	0.138
M+z	24	3+	2146.337153	2146.337709	0.259
M+z	25	4+	1627.512008	1627.512645	0.391
M+z	25	3+	2170.350335	2170.351009	0.311
M+z	26	5+	1324.827880	1324.827646	-0.177
M+z	26	4+	1655.783031	1655.783415	0.232
M+z	26	3+	2208.046833	2208.046104	-0.330
M+z	27	5+	1356.048099	1356.048495	0.292
M+z	27	4+	1695.060261	1695.060550	0.170
M+z	27	3+	2260.080530	2260.081289	0.336
M+z	28	5+	1381.659815	1381.660250	0.315
M+z	28	4+	1726.824267	1726.823766	-0.290
M+z	29	5+	1402.069844	1402.069738	-0.076
M+z	29	4+	1752.086188	1752.086786	0.341
M+z	30	5+	1416.076777	1416.076531	-0.174
M+z	30	4+	1769.844153	1769.844536	0.216
M+z	34	5+	1491.103517	1491.103470	-0.032
M+z	34	4+	1864.130123	1864.130353	0.123
M+z	36	5+	1534.713228	1534.713224	-0.003
M+z	36	4+	1918.141132	1918.141912	0.407
M+z	16-i	4+	1370.399228	1370.399399	0.125
M+z	16-i	3+	1827.199153	1827.199447	0.161
M+z	17-i	4+	1399.406577	1399.406737	0.114
M+z	17-i	3+	1865.875642	1865.875456	-0.100

		Average Error:	0.068
		Absolute Average Error:	0.187
		Standard Deviation:	0.245

Table S2.12.   List of the assigned fragments from the ECD MS/MS spectrum of the 7+ charge state ((isoD)$_3$hIAPP) dimer ion (in aqueous solution with 2% residual DMSO).   M represents one ((isoD)$_3$hIAPP) unit.

A) Measurement of fluorescence emission in $(D)_3$hIAPP at 490 nm

■ 5 µM $(D)_3$hIAPP
$R^2 = 0.95$
$t_{50\%} = 32$ hours
$y_{max} = 0.25$

■ 10 µM $(D)_3$hIAPP
$R^2 = 0.99$
$t_{50\%} = 33$ hours
$y_{max} = 0.96$
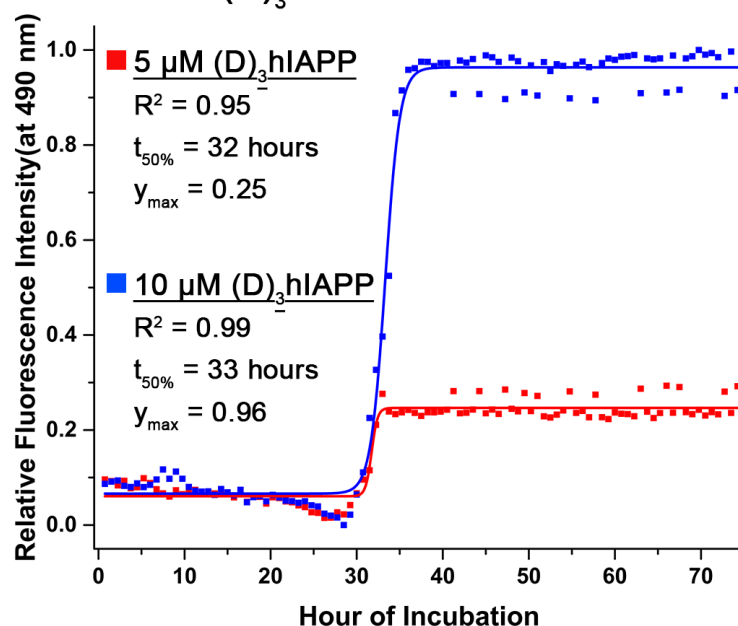
B) Measurement of fluorescence emisson in $(isoD)_3$hIAPP at 490 nm

■ 5 µM $(isoD)_3$hIAPP
$R^2 = 0.99$
$t_{50\%} = 37$ hours
$y_{max} = 0.20$

■ 10 µM $(isoD)_3$hIAPP
$R^2 = 0.99$
$t_{50\%} = 39$ hours
$y_{max} = 0.94$
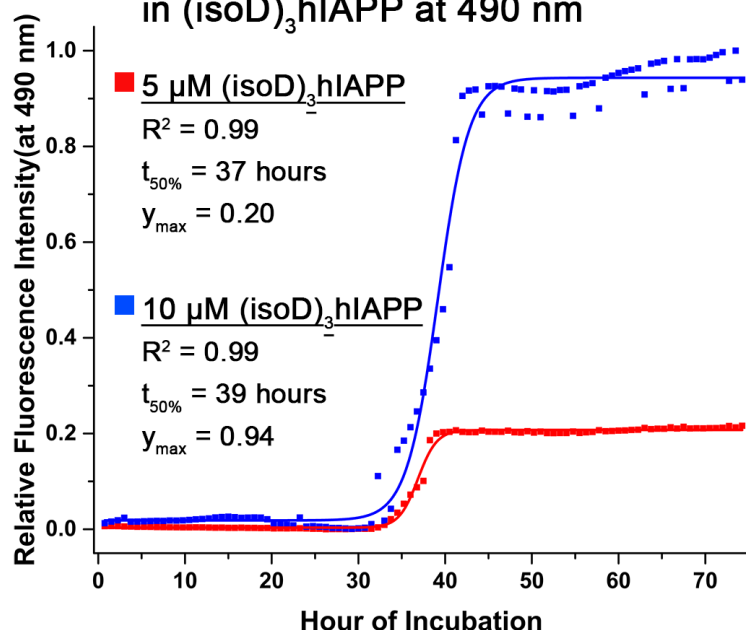
Figure S2.18. The relative fluorescence emission intensity of the µM and 10 µM mutant hIAPP solutions (A) ($(D)_3$hIAPP) and (B) ($(isoD)_3$hIAPP). The relative fluorescence activity of the solutions of ($(D)_3$hIAPP) and ($(isoD)_3$hIAPP) were measured and the lag phase of the formation of mature fibrils were 30 and 32 hours respectively. Compared to the 10 µM solutions of the

synthetic mutant hIAPPs, the 5 μM solutions of ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) had a marginally shorter half-life of mature fibrils formation, which were around 40 hours.   Even though there was no significant difference between the midpoint formations (lag phase) of mature fibrils for the 5 μM and 10 μM solutions, the relative fluorescence intensity of the 10 μM solutions of ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) were 3.8 and 4.7 times higher than their 5 μM solutions, respectively.
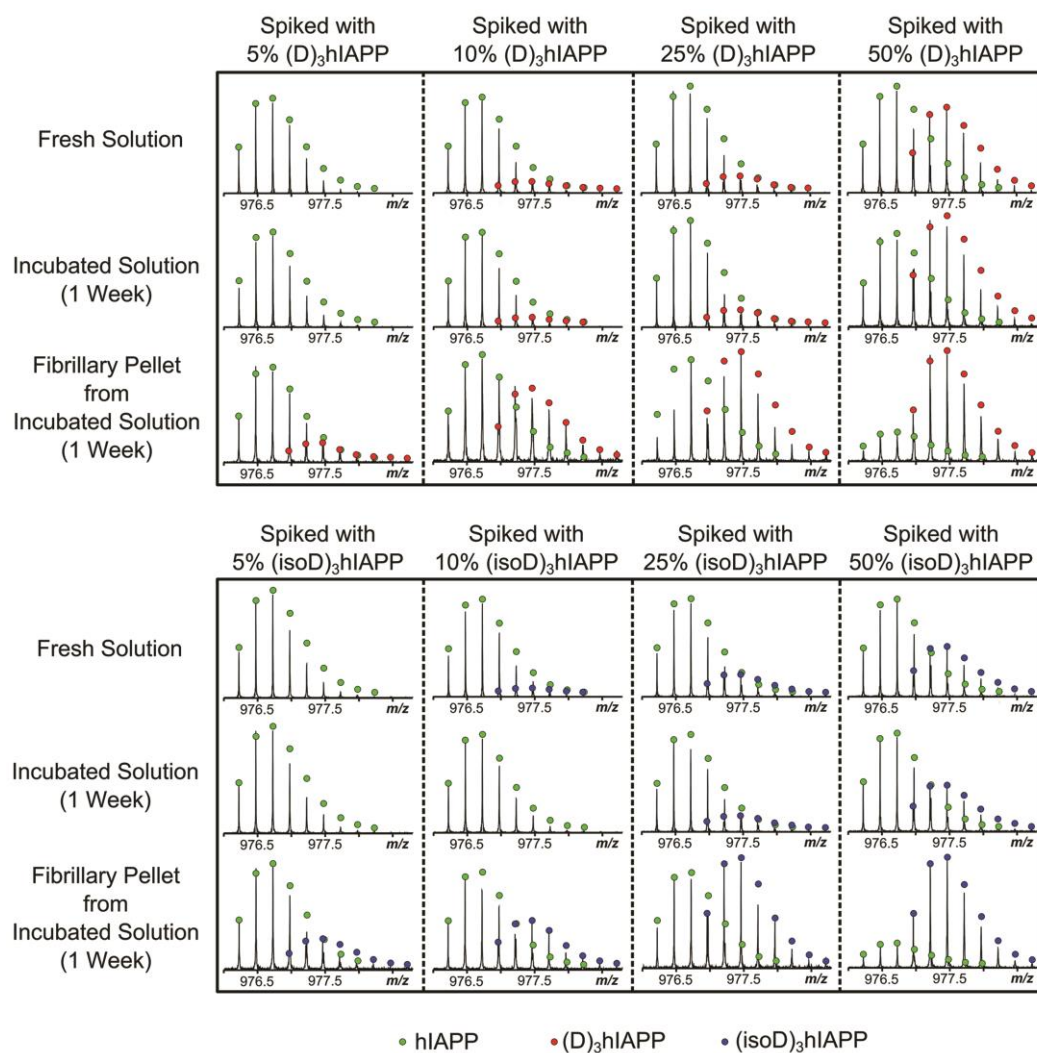
Figure. S2.19. The nESI-MS spectra showing the fresh, the incubated solutions, and the incubated fibrillary pellets of hIAPP mixed with 5%, 10%, 25%, or 50% of mutant ((D)₃hIAPP) or ((isoD)₃hIAPP).

Chapter 3 The Inhibition Pathways of Human Islet Amyloid Polypeptide

This chapter demonstrates the applications of Fourier transform ion cyclotron resonance (FTICR MS) together with fluorescence spectroscopy and transmission electron microscopy (TEM) to study the inhibition pathways of various potential drugs towards aggregation of an amyloidogenic protein − human islet amyloid polypeptides (hIAPP).

The results presented in this chapter were predominantly carried out by the thesis Author, aside from acquisition of transmission electron microscopy (TEM) data were assisted by Mr. Ian-Hands Portman from the Department of Life Science, University of Warwick.

A manuscript entitled "The Inhibition Pathways of Human Islet Amyloid Polypeptide" by Yuko P. Y. Lam, Cookson K.C. Chiu, Christopher A. Wootton, Ji-Inn Song, Meng Li, Ian Hands-Portman, Mark P. Barrow, and Peter B. O'Connor, has been prepared for submission to Nature Chemistry based on the results presented in this chapter.

Chapter 3 – The Inhibition Pathways of hIAPP

3.1 Abstract

Human islet amyloid polypeptide (hIAPP) is a highly amyloidogenic protein that aggregates rapidly in humans. Classic structure-based design for therapeutics drug development against amyloid proteins is challenging as most amyloid proteins are inherently disordered with high conformation flexibility. Many compounds are proposed as potential inhibitors of toxic oligomer formation, but target regions for inhibiting aggregation and reports on inhibitor performance are mixed and often conflict. Herein, mass spectrometry was used to identify binding regions of a range of reported hIAPP aggregation inhibitors and were correlated with fluorescence and transmission electron microscopy measurements to understand target regions for aggregation inhibition mechanisms of action for different classes of inhibitors. From our results, two mechanisms are shown to effectively inhibit the aggregation of wild-type hIAPP either by site-specific interaction with the amyloidogenic region on hIAPP (Ser-29 to Asn-35) or accelerate the formation of non-toxic amorphous aggregates. Our data further show that the aggregation of deamidated hIAPP could no longer be effectively prevented by the site-specific inhibition compounds, instead an amorphous aggregate was rapidly formed. The data herein contribute to understanding the inhibition mechanisms on wild-type and deamidated hIAPP using potential inhibitors and forms an important basis for future therapeutics development.

3.2 Introduction

Abnormal aggregation of proteins and peptides into mature amyloid fibrils is a hallmark in approximately 50 human disorders, including Parkinson's disease (PD), Alzheimer's disease (AD), Huntington's disease (HD), and type II diabetes (T2D).[1-4] Inhibiting amyloid fibrils formation is a critical goal in searching for potential therapeutic compounds. Most amyloid proteins, however, are inherently disordered with high conformation flexibility,[3, 4] making inhibition sites difficult to identify through classic therapeutic development using structure-based design.[5, 6] Thus, a better understanding of the inhibition mechanisms of current amyloid protein aggregation inhibitors would help minimise the time and cost spent in searching for therapeutics against amyloid protein aggregation, as well as accelerating the development of higher efficiency, less toxic therapeutic compounds.

Human islet amyloid polypeptide (hIAPP, also named amylin) is a 37 amino acid hormone with an intramolecular disulfide bond between Cys-2 and Cys-7.[7-11] hIAPP is stored in the pancreatic β-cell secretory granules and co-secreted with insulin to regulate blood glucose levels.[12, 13] In general, monomeric hIAPP is soluble and inherently disordered, and hIAPP amyloid fibrils are only found in T2D patients.[7, 8] The formation of hIAPP fibrils does not contribute to T2D, but the early oligomers of hIAPP do lead to the decline of pancreatic β-cell mass and the failure of islet cell transplantation for T2D treatment.[14, 15] Thus, reducing the formation of the early oligomers of hIAPP can enhance the graft survival rate.[16, 17]

The early oligomers of hIAPP are more toxic and harmful to the pancreatic β-cell than the mature hIAPP fibrils;[18] thus potential therapeutic compounds should not only inhibit mature fibril formation, but should also interact with the early oligomers of hIAPP to prevent the formation of protofibrils. (-)-

Chapter 3 – The Inhibition Pathways of hIAPP

Epigallocatechin 3-gallate (EGCG), an extract from green tea, is one of the famous inhibitors which interacts with the monomeric state of hIAPP and reduces mature amyloid fibril formation.[19-22]   The inhibition mechanism of EGCG on hIAPP aggregation is still under debate.   Hudson *et al.* showed EGCG interacted with the monomeric state of amyloid protein and maintained the protein in its pre-fibrillary state inhibiting further aggregation.[23]   However, Dagmar *et al.* showed EGCG re-directed amyloid protein into off-pathway oligomers which did not induce cell toxicity.[24]   Not only is the inhibition mechanism not clear, the interaction mode between EGCG and hIAPP is also complex because aromatic interaction and non-specific hydrophobic interactions with protein sidechains have been proposed simultaneously.[20, 25, 26]

Insulin is a hormone peptide with 51 amino acid residues over two chains (A chain contains 21 residues and B chain contains 30 residues) and linked by one intra-stand and two inter-strand disulfide bonds.[27]   The concentration of insulin is around a hundred times higher than the level of hIAPP in the β-cell secretory granules.[28]   Similar to hIAPP, insulin also aggregates and forms amyloid fibrils, but insulin mostly aggregates rapidly at low pH (~pH 2-3) while hIAPP aggregates much faster at neutral pH (~pH 7).[29-31]   Insulin has been shown to be a potential inhibitor of hIAPP fibril formation as monomeric insulin can interact with hIAPP to form a heterodimer, preventing hIAPP aggregation.[32-35]   Studies using nuclear magnetic resonance spectroscopy (NMR) and ion mobility mass spectrometry (IM-MS) with computational prediction suggested the alpha helix region of hIAPP (residue 8 – 23) and insulin B chain (residue 9 – 24) interact to form the heterodimer complex.[32, 34]

Deamidation is a non-enzymatic post-translational modification which leads to protein ageing and has been called a "molecular clock".[36, 37]   Deamidation at

asparagine (Asn) and glutamine (Gln) residues contributes to the formation of iso-aspartic acid or aspartic acid, and γ-glutamic acid or glutamic acid respectively, as well as stereoisomers.[38, 39] Presence of an extra methylene group (-CH$_2$-) in the backbone of iso-aspartic acid and γ-glutamic acid contributes to a significant alteration in protein structure, stability, folding, and function.[40, 41] Deamidation in other amyloid proteins, including amyloid-β, β$_2$-microglobulin, and αβ-crystallin, has been shown to accelerate the amyloid fibril formation.[42-47] hIAPP contains six Asn and one Gln potential deamidation sites which converted within one week of incubation at 37°C.[48] A previous study using a non-amyloidogenic variant of hIAPP segment, residues 20-29 (SNNFPAILSS) showed that the presence of less than 5% of deamidated monomer segment was required to initiate the formation of amyloid fibrils.[49] Results from Dunkelberger *et al.* further suggested the β-sheet structure of deamidated hIAPP around Leu-27 had been disrupted and accelerated the formation of hIAPP fibrils.[50] Recent studies have further narrowed the deamidation sites of hIAPP to Asn-21, Asn-22, and Asn-35 residues, and showed mutant hIAPP (with iso-aspartic acid residues replacement at the deamidation sites) aggregated into amyloid fibrils faster than the wild-type hIAPP.[48] Even though deamidated hIAPP has been reported to significantly affect the morphology of hIAPP and the aggregation rate, the interactions between the deamidated hIAPP and the inhibitors commonly used for preventing the aggregation of wild-type hIAPP have not been studied.

Fluorescence dye-binding, i.e. thioflavin T (ThT), is one of the most common techniques to monitor the kinetics of amyloid fibrils formation,[51] as well as the effectiveness of potential inhibitors against aggregation by measuring the change of onset of fibril formation,[52, 53] but the additional dyes and buffers can cause misleading results, as well as increase the aggregation rate of amyloid proteins.[54,

[55] Thus, new techniques with high resolution and sensitivity are required to monitor the interaction between the early oligomers of hIAPP and potential inhibitors.

IM-MS has recently been applied to monitor the formation of hIAPP oligomers, and the interactions between the early oligomers of hIAPP and various potential inhibitors.[21, 22, 34, 56-60] These studies demonstrated that mass spectrometry (MS) provides high sensitivity and resolution to further study the interaction between early oligomers and potential inhibitors. A high throughput screening method has also been proposed to search for the potential inhibitor against amyloid proteins aggregation using IM-MS.[22] However, understanding the underlying inhibition mechanisms between the early oligomers of amyloid proteins and the potential inhibitors used can inform further drug development directly to inhibit the formation of toxic aggregates. Electron capture dissociation tandem MS (ECD MS/MS) provides the ability to fragment a protein backbone while preserving the non-covalent interaction between proteins and protein-molecule complexes,[61-63] and thus can be used to determine the interaction region between hIAPP and the potential inhibitors.[48]

Herein, Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) is applied in monitoring the interactions between hIAPP and potential anti-aggregation drugs. Both previously reported hIAPP inhibitors (insulin and EGCG) and species able to bind to hIAPP, but believed to not inhibit amyloid fibril formation are investigated. MS quantification of soluble amyloid proteins, ThT fluorescence emission measurement of the onset of mature fibrils formation, and the observation of the overall hIAPP fibrillary structure using transmission electron microscopy (TEM) help to evaluate the short- and longer-term inhibition performance of the binding molecules. ECD MS/MS is also applied in order to

determine the interaction sites between hIAPP and the potential inhibitors, which contributes to the understanding of possible inhibition mechanisms for hIAPP. The results herein demonstrate more than one inhibition mechanism can be used to prevent the formation of hIAPP fibrils, while also showing that molecules capable of binding to hIAPP are not necessarily capable of inhibiting amyloid fibril formation. Analysis of binding locations, correlated with inhibition of fibril formation also reveal specific target locations within the protein sequence to reduce oligomerization, and provide a basis for higher throughput screening in the future.

3.3 Experimental Section

Sample preparation for MS.   Wild-type hIAPP with purity >95% (AS-60254-1; Eurogentec Ltd., England), the synthetic mutant hIAPP with purity >95% and iso-aspartic acid residue replacement at Asn-21, Asn-22, and Asn-35 ((isoD)$_3$hIAPP; Pepscan Company Ltd, The Netherlands), and the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide with purity >95% (Eurogentect Ltd., England) were dissolved in Milli-Q H$_2$O (Direct-Q$^®$ 3 UV System, Millipore Corporation, USA) with pH ~7.5 to 500 µM for storage (at -80°C) and samples were consumed within two weeks.   3-amino-1-propane sulfonic acid (3-APS), 1H-Benzimidazole-2-sulfonic acid (BISA), and (-)-epigallocatechin gallate (EGCG) were purchased from Sigma Aldrich Company Ltd, England and dissolved in Milli-Q H$_2$O to 500 µM for storage (at -80°C).   Recombinant human insulin (Sigma Aldrich Company Ltd, England) was dissolved in acidified Milli-Q H$_2$O (~pH 5.8) to 500 µM for storage (at -80°C), as performed by Susa *et al*.[34]

For the MS spectra of hIAPP with potential inhibitors, 50 µM solutions of hIAPP, (isoD)$_3$hIAPP, or ^{30}TNVGSNTY^{37}-NH$_2$ peptide were initially mixed with 50 µM solutions of 3-APS, BISA, EGCG, or insulin in Milli-Q H$_2$O separately. The mixed solutions were then vortexed for half an hour in room temperature before further dilution by 5-fold with Milli-Q H$_2$O into 10 µM solutions.   The pH values of all solutions were measured.   The values were between pH 6.8 − pH 7.5. The solutions were then centrifuged at 14,000 rpm for five minutes before MS analysis.

For MS relative quantification, 10 µM solutions of wild-type hIAPP, mutant (isoD)$_3$hIAPP, or synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide were mixed with 10 µM solutions of 3-APS, BISA, EGCG, or insulin in Milli-Q H$_2$O individually.   A 200 µL aliquot was extracted from each solution and stored in -80°C while the

remaining 300 µL solution was incubated at 37°C for seven days.   After seven days, solutions stored at -80°C (regarded as fresh solutions) and solutions incubated at 37°C were centrifuged at 14,000 rpm for one hour to separate the soluble wild-type hIAPP, mutant (isoD)$_3$hIAPP, or ^{30}TNVGSNTY^{37}-NH$_2$ peptide (supernatant) from the insoluble aggregates (precipitate) in the solutions. Supernatant solutions were further diluted 10-fold with 49.5:49.5:1 H$_2$O/acetonitrile/formic acid and centrifuged at 14,000 rpm for five minutes prior to MS analysis.

FTICR MS analysis.   MS spectra of the early oligomers of the amyloid proteins with or without potential inhibitors were acquired using nano-electrospray ionisation (nESI) FTICR MS with a capillary voltage of 0.8 – 1.3 kV and a source temperature of 80°C.   The nESI glass capillaries (World Precision Instruments, England) were pulled on a Sutter P-97 capillary puller instrument (One Digital Drive Novato, USA).   Samples were loaded onto the glass capillaries and inserted with nickel chromium wire (Jacobs, USA) to ground the electric potential for ionisation.

For the MS quantification spectra of soluble amyloid proteins, an Apollo II electrospray ionisation (ESI) source MS (Bruker Daltonics ,Bremen, Gremany) with a capillary voltage of 4 – 5 kV was applied and the source temperature was set to 200°C.   The ESI flow rate was optimised to 100 – 150 µL/h.   Five MS spectra were acquired for each solutions as technical replicates.

All mass spectra were acquired on a 12 tesla solariX FTICR MS (Bruker Daltonics, Bremen, Germany) in positive ionisation mode.   Sample ions were first accumulated from 0.1 – 1 seconds in the hexapole collision cell before being directed to an infinity ICR cell for excitation and detection.[64]   All mass spectra were acquired with four mega point (32-bit) transients and were analysed using

Bruker DataAnalysis 4.2 software MS (Bruker Daltonics, Bremen, Germany).

FTICR MS/MS analysis. For the collisionally activated dissociation tandem MS (CAD MS/MS) experiments, precursor ions were isolated using a quadrupole mass filter, collided with argon gas and accumulated for $1 - 3$ seconds in the hexapole collision cell.   The collision energies were varied from $0 - 16$ V in order to generate the breakdown curves for the precursor ions which were composed of the amyloid protein with the bound molecule of interest.   Fragments were then directed to the ICR cell for detection.   Five spectra were acquired for each collision energy level as technical replicates.

For ECD MS/MS experiments, precursor ions were first isolated with a wide mass-to-charge (*m/z*) isolation window (100 *m/z*) in the quadrupole and accumulated for 1-5 seconds in the collision cell.   Ions were further transferred to the ICR cell and isolated with MULTI-CHEF isolation with an excitation power of $50 - 55\%$.[65]   The precursor ions were then irradiated with 1.2 eV electrons from a 1.5 A heated hollow cathode and the fragments were detected by the ICR cell.

MS data analysis.   All MS data were analysed using Bruker DataAnalysis 4.2 software MS (Bruker Daltonics, Bremen, Germany).   The most intense isotopic peak from each fragment with signal-to-noise ratio (S/N) over five was manually matched with the assigned species.   Five fragments from each spectrum were chosen for the internal calibration of the spectrum and all fragments were then assigned with an uncertainty of $0.1 - 10$ part-per-million (ppm) (see Supplementary Information, Table S3.1- S3.15).

To study the binding strength between wild-type hIAPP and the potential inhibitors, the dissociation curves of each complex precursor ion, which composed of hIAPP and the binding molecule, were plotted.   The areas of all the isotopic

peaks from each charge state of hIAPP, the complexes, and the potential inhibitors (if present) were measured using the Bruker DataAnalysis 4.2 software. The relative percentage of the complex was calculated as follow:

$$\text{Complex (\%)} = \frac{Isotopic\ peak\ areas\ of\ the\ complex}{Total\ isotopic\ peak\ areas\ of\ hIAPP + complex + binding\ molecule}$$

The dissociation curves for the complexes of hIAPP with different potential inhibitors were obtained by plotting the relative percentage of each complex calculated above against the applied CAD energies. The same calculation method was used to obtain the dissociation curves of the mutant $(isoD)_3$hIAPP and the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide with various potential inhibitors.

To study the long-term inhibition performance of each binding molecule against hIAPP aggregation, the MS spectra of the incubated wild-type hIAPP solutions (with/ without inducing the potential inhibitor) were compared to their corresponding fresh MS spectra which were obtained using the solutions stored at -80°C. The absolute peak areas of the fresh and incubated solutions were first tested against a two-tailed T-test to differentiate whether the peak areas between the fresh and incubated solutions were significant. The hypothesis applied in the two-tailed T-test assumed there was no significant difference between the fresh and incubated solutions; thus a p-value that was less than 0.05 (95% confidence) meant the two solutions were significantly different and vice versa. The percentage change of soluble hIAPP was then calculated as follows after the two-tailed T-test:

$$\text{\% change} = \frac{(peak\ area\ of\ incubated\ solution - peak\ area\ of\ fresh\ solution)}{peak\ area\ of\ fresh\ solution} \, x\ 100\%$$

The same calculation method was applied to obtain the percentage change of the mutant $(isoD)_3$hIAPP and the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide experiments.

Transmission electron microscopy (TEM). TEM images of the incubated

solutions, including the 10 µM solutions of the wild-type hIAPP, the mutant (isoD)$_3$hIAPP, and the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide individually mixed with the 10 µM solutions of 3-APS, BISA, EGCG, and insulin, were acquired on a Jeol 2010F TEM which operated at 200 kV. 5 µL of the incubated solution was added onto a carbon coated grid (EM Resolutions, England) and settled for one minute, followed by using a filter paper to remove the excess aqueous solution. A 2% (w/v) uranyl acetate solution was then added onto the carbon grid and settled for one minute before removal by a filter paper. Multiple TEM images were acquired with magnification from x10,000 to x50,000. The contrast of the TEM images were auto-corrected by IrfanView 64 software (Irfan, Austria).

Thioflavin T (ThT) fluorescence reactivity. Fluorescence emission intensities of the 50 µM solutions of the wild-type hIAPP, the mutant (isoD)$_3$hIAPP, the mixtures of wild-type hIAPP with various potential inhibitors were measured using a GloMax$^®$-Multi Detection System (Promega, USA) at room temperature. All samples were placed in a black 96 well-plate and mixed with a 100 µM ThT solution. Fluorescence emission for the samples in the 96 well-plate was measured every 45 minutes for 73 hours (99 sampled points) with excitation at 405 nm and emission measurement at 490 nm.[66, 67] The fluorescence emission intensities were normalised to one using the maximum and the minimum readings. The lag phase of each sample was then calculated according to the previous studies.[68, 69]

The corrected ThT fluorescence emission intensities of the wild-type hIAPP and the mutant (isoD)$_3$hIAPP solutions were obtained by using the absolute fluorescence emission intensity obtained from the amyloid protein solutions minus the absolute fluorescence emission intensity of the 100 µM ThT solution. The corrected ThT fluorescence emission intensities of the mixtures of wild-type

Chapter 3 – The Inhibition Pathways of hIAPP

hIAPP/ mutant $(isoD)_3$hIAPP with various potential inhibitors were calculated using the absolute fluorescence emission intensities obtained from the mixtures minus the absolute fluorescence emission intensities of the corresponding potential inhibitor.

3.4 Results

The interactions between early oligomers of the wild-type hIAPP and potential aggregation inhibitors. A wide range of molecules has been shown to bind to wild-type hIAPP; however, only a small portion of these molecules are useful in inhibiting hIAPP aggregation.[22, 34] When 10 µM wild-type hIAPP was diluted in an aqueous solution, it aggregated rapidly to form early oligomers within minutes; as a result, the early oligomers of hIAPP, from monomer up to pentamer, were observed in the MS spectrum (see Figure 3.1A). When hIAPP was mixed with an equimolar insulin solution, hIAPP and insulin rapidly formed a heterodimer complex (one hIAPP plus one insulin protein) in the solution, together with the homodimers of hIAPP and insulin were observed in the MS spectrum (see Figure 3.1B). The hIAPP-insulin heterodimer was the only reaction product observed in the MS spectrum. The homodimer-to-monomer ratio of hIAPP was significantly reduced with addition of insulin.

Sequence of human insulin

G I V E Q C C T S I C S L Y Q L E N Y C N

F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T

Sequence of hIAPP

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y - NH$_2$
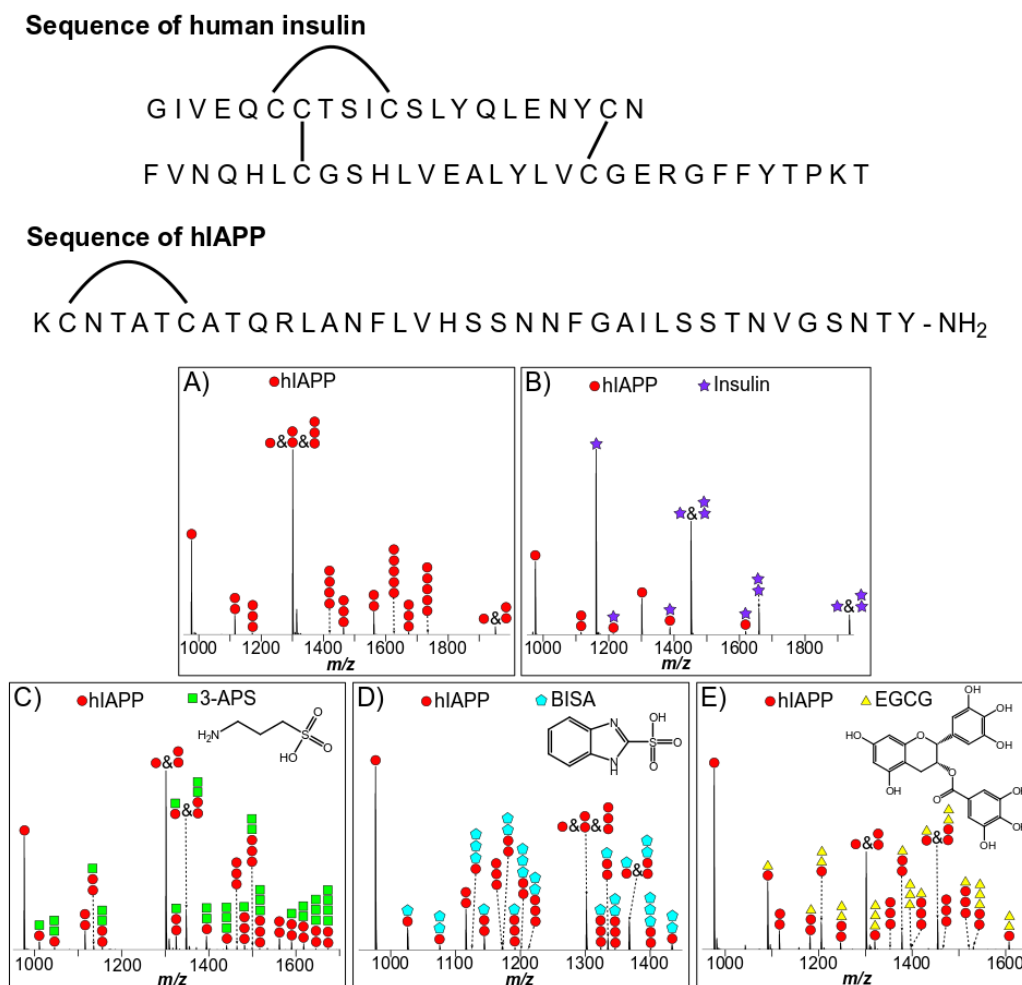
Figure 3.1. The sequences of human insulin and hIAPP. The MS spectra of the 10 µM solution of the wild-type hIAPP (A; red circle) only, the 10 µM aqueous solutions of the wild-type hIAPP mixed with the 10 µM insulin (B; purple star), 10 µM 3-APS (C; green square), 10 µM BISA (D; cyan pentagon), and 10 µM EGCG (E; yellow triangle).

The MS spectra of equimolar mixture of the wild-type hIAPP and 3-APS/ BISA/ EGCG were also acquired individually (see Figure 3.1C – 3.1E) and the results obtained from the spectra showed a similar trend. The early oligomers of hIAPP, from monomer to trimer, together with the complexes of the early oligomers of hIAPP with 3-APS/ BISA/ EGCG molecule(s) were simultaneously observed in the spectrum. The molecular binding ratios between hIAPP and 3-

Chapter 3 – The Inhibition Pathways of hIAPP

APS/ BISA/ EGCG ranged from 1:1 to 1:3 respectively.

The performance of potential inhibitors against hIAPP aggregation.
Fluorescence emission intensity is commonly used to measure the change in lag phase and the relative amount of amyloid fibril formation which reflects the inhibition effects of the compounds.[51, 68, 69]   The calculated lag phase of the 50 µM solution of wild-type hIAPP was around 18 hours (see Figure 3.2A).   The lag phase of the 50 µM solution of hIAPP mixed with equimolar of 3-APS was slightly reduced to 15 hours, while the lag phase of the 50 µM solution of hIAPP mixed with equimolar of BISA was slightly prolonged to 20 hours.   The lag phases were significantly extended to 24 and 29 hours respectively, when equimolar insulin and EGCG were induced into the hIAPP solutions individually, showing promising inhibition of fibril formation.   The amount of mature fibrils formed in the mixture solutions of hIAPP and various compounds compared to pure hIAPP were as follows, insulin 55%, 3-APS 95%, BISA 114%, and EGCG 33%.   A corrected ThT fluorescence emission intensity was also calculated by deducting the emission obtained from the potential inhibitor in order to minimise the emission influence by the added molecule (see Supplementary Information, Figure S3.1C).   After the correction, the amount of hIAPP fibrils that could be detected in the mixture solution of hIAPP and insulin was essentially zero, showing potential inhibition of hIAPP aggregation.

A)

Model	Boltzmann Fitting				
Sample	hIAPP only	hIAPP+ Insulin	hIAPP + 3-APS	hIAPP + BISA	hIAPP + EGCG
Calculated lag phase	17.9	24.5	11.0	18.5	28.1
Δ% of fibrils detected	0	-45.6	-4.8	+13.7	-63.6

B)

	hIAPP only	hIAPP + Insulin	hIAPP + 3-APS	hIAPP + BISA	hIAPP + EGCG
% Change	-49.71	-0.06	-20.63	-33.00	-55.35
p value	5.16E-08	0.92	2.94E-05	3.27E-04	9.40E-09

C) hIAPP only

D) hIAPP + Insulin

E) hIAPP + 3-APS

F) hIAPP + BISA

G) hIAPP + EGCG
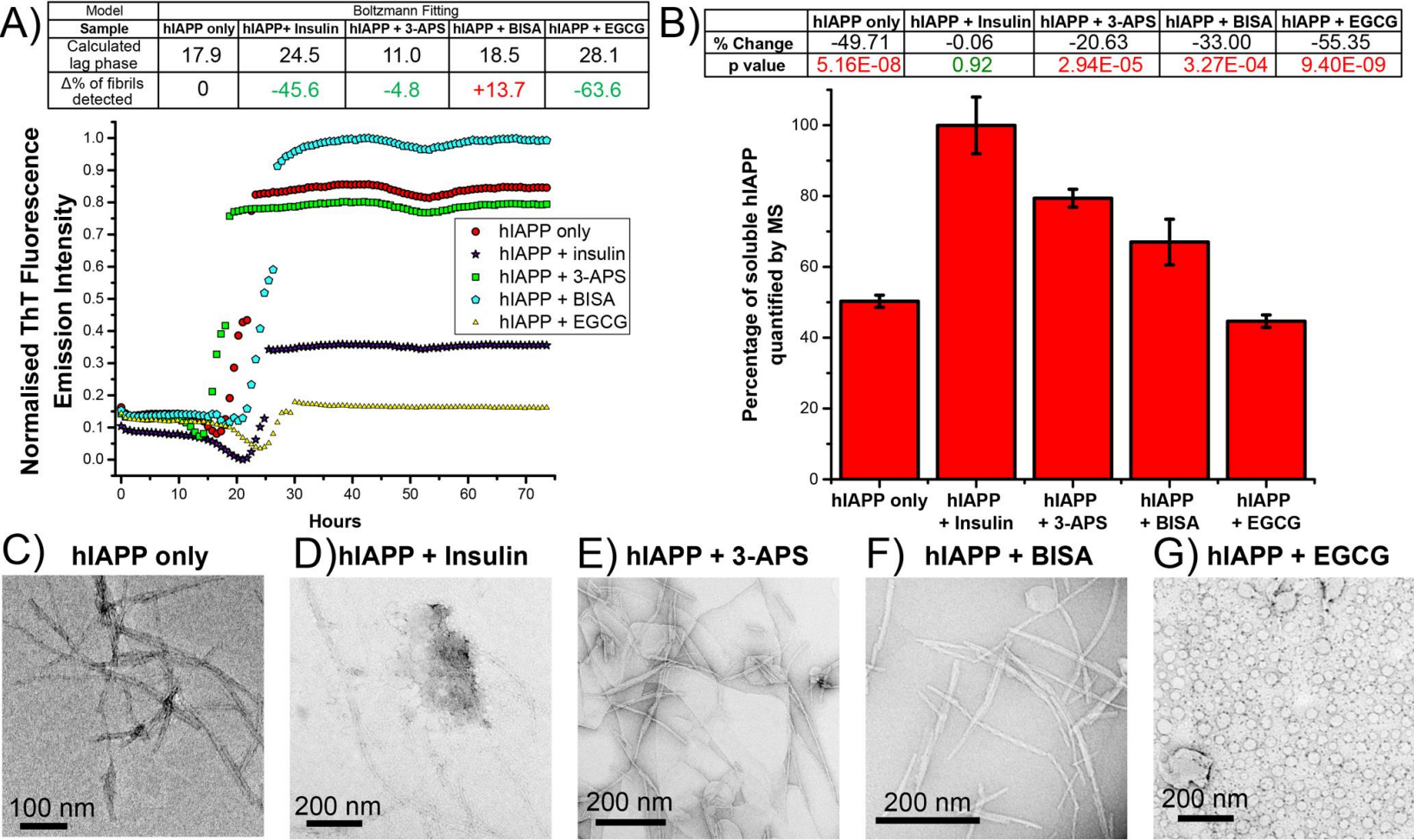
Figure 3.2.   (A) The normalised ThT fluorescence emission intensities of the 50 µM wild-type hIAPP only (red circle), the solution mixtures of the 50 µM wild-type hIAPP induced with the 50 µM solutions of insulin (purple star), 3-APS (green rectangle), BISA (cyan pentagon), and EGCG (yellow triangle).   (B) The percentage of the remaining soluble wild-type hIAPP in the 10 µM solutions, incubated for one week, followed by the results of the two-tailed T-test.   TEM images of the 10 µM, one week incubated solutions of (C) the wild-type hIAPP only, the mixtures of the wild-type hIAPP and equal molar of (D) insulin, (E) 3-APS, (F) BISA, and (G) EGCG molecules.   Scale bar for each image shown were inset.

The performances of the potential inhibitors were also measured by quantifying the amount of remaining soluble hIAPP between the fresh and one week incubated samples (see Figure 3.2B). Two-tailed t-tests were performed to determine whether the differences measured between the fresh and incubated solutions were significant. Pure hIAPP aggregated significantly over 1-week incubation period, resulting in 50% remaining soluble hIAPP quantifiable by MS, and the t-test p value was less than 0.05. The remaining soluble hIAPP in the hIAPP and insulin solution after one week was ~100% and the t-test p value of the fresh and incubated solutions was also greater than 0.05. The remaining soluble insulin quantified in the solution was 88% and the p value was less than 0.05 (see Supplementary Information, Figure S3.2). For other compounds, the quantity of the soluble hIAPP was significantly reduced by 21% (3-APS), 33% (BISA) and 55% (EGCG) in the incubated solutions and the t-test p values were all less than 0.05.

TEM images do not inherently provide quantitation information, but allow large-scale fibril structure to be observed in the amyloid aggregates. Pure hIAPP incubated sample produced elongated, crosslinked fibrils observed in the TEM images (see Figure 3.2C). Similar results were observed in the mixtures of hIAPP with equimolar solutions of 3-APS or BISA (see Figure 3.2E & 3.2F). Fewer fibrils were observed in the TEM image of the incubated solution of hIAPP with insulin (see Figure 3.2D). Numerous circular vesicle-like structures with an estimated average diameter of 48 nm were observed in the TEM images of hIAPP with EGCG (see Figure 3.2G). Each circular vesicle seemed to contain several dense spots which had an estimated average diameter of 4.2 nm.

The interaction sites between hIAPP and the potential inhibitors. CAD

MS/MS was applied to determine the relative binding energy between the monomer unit of hIAPP and the potential inhibitors (see Supplementary Information, Figure S3.3), the dissociation curves were then obtained (see Supplementary Information, Figure S3.4). The reaction product ions were dissociated completely to the hIAPP monomer ion using low collision energy thus strength of interaction can be indirectly compared, but no interaction site information was available using CAD MS/MS.

ECD MS/MS was then applied to the same range of reaction product ions. The 7+ molecular ion, which composed of a hIAPP monomer unit with an insulin molecule, was fragmented by the electrons in the FTICR cell, 136 fragments were identified from the ECD MS/MS spectrum (see Figure 3.3A; Supplementary Information, Figure S3.5 and Table S3.1.). The most critical ECD fragments were the z_3 and c_{30} ions (highlighted in Figure 3.3A and 3.3B) indicating the intact insulin molecule binds to hIAPP between Ser-29 and Asn-35 residues (see Figure 3.3C).

A) Insulin fragment ▢ hIAPP fragment ● hIAPP fragment with Insulin ★ Insulin fragment with hIAPP ▲ hIAPP plus Insulin ★●
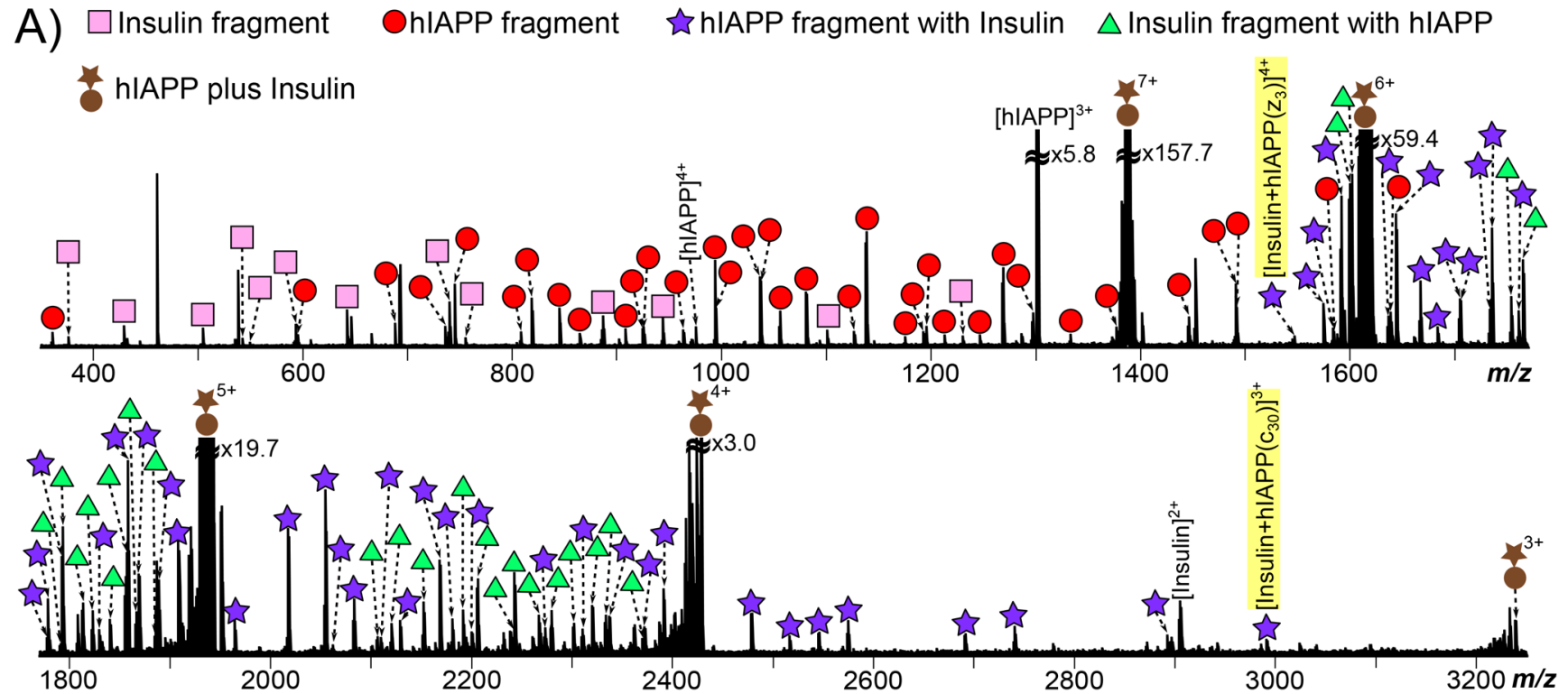
B) **ECD fragments of hIAPP with intact Insulin**

c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y - NH₂

z ion

(Figure 3.3. Continue)
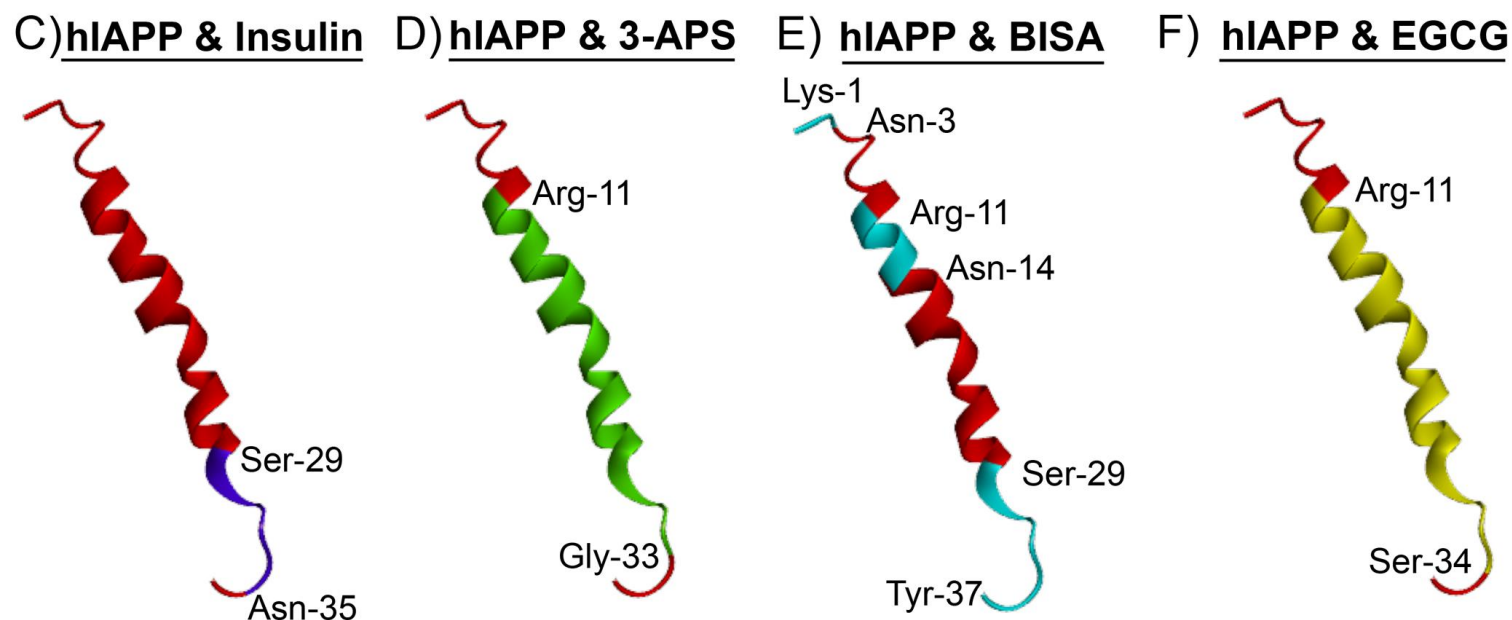
Figure 3.3. (A) The ECD spectrum of the 7+ molecular ion of hIAPP plus insulin. The highlighted fragments indicate the interaction region of hIAPP with insulin. (B) Summarised ECD MS/MS fragments of the 7+ molecular ion from (A). The proposed interaction region of the monomer unit of hIAPP with an attachment of (C) insulin (purple region), (D) 3-APS (green region), (E) BISA (cyan region), and (F) EGCG (yellow region). Non-interacting regions are shown in red.

ECD MS/MS was then applied to determine the interaction region between hIAPP monomer unit and the other potential inhibitors, i.e. 3-APS, BISA, and EGCG. ECD spectra of each complex were generated (see Supplementary Information, Figure S3.6 – S3.8 and Table S3.2 – S3.4) and the detected fragments from each spectrum were used to identify the interaction region between hIAPP and the potential inhibitors. For 3-APS and EGCG, the interaction regions were similar and located between Arg-11 to Gly-33 (3-APS) and Arg-11 to Ser-34 (EGCG) (see Figure 3.3D and 3.3F).

BISA was shown to interact in multiple regions, Lys-1 to Asn-3, Arg-11 to Asn-14, and Ser-29 to Tyr-37. The ECD MS/MS result also indicated the sulfonic acid group of BISA bound to the monomer unit of hIAPP between Arg-11 and Asn-14 residues, while the benzimidazole group of BISA was attached to the residues of hIAPP at Lys-1 to Asn-3 and Ser-29 to Try-37 (see Supplementary Information, Figure S3.7 and Table S3.3).

ECD MS/MS of a hIAPP dimer bound to one 3-APS or one BISA molecule showed that both 3-APS and BISA effectively bind between Lys-1 and Asn-14 (see Figure 3.4B and 3.4C, respectively) which indicates a non-disruptive binding site of potential inhibitors.
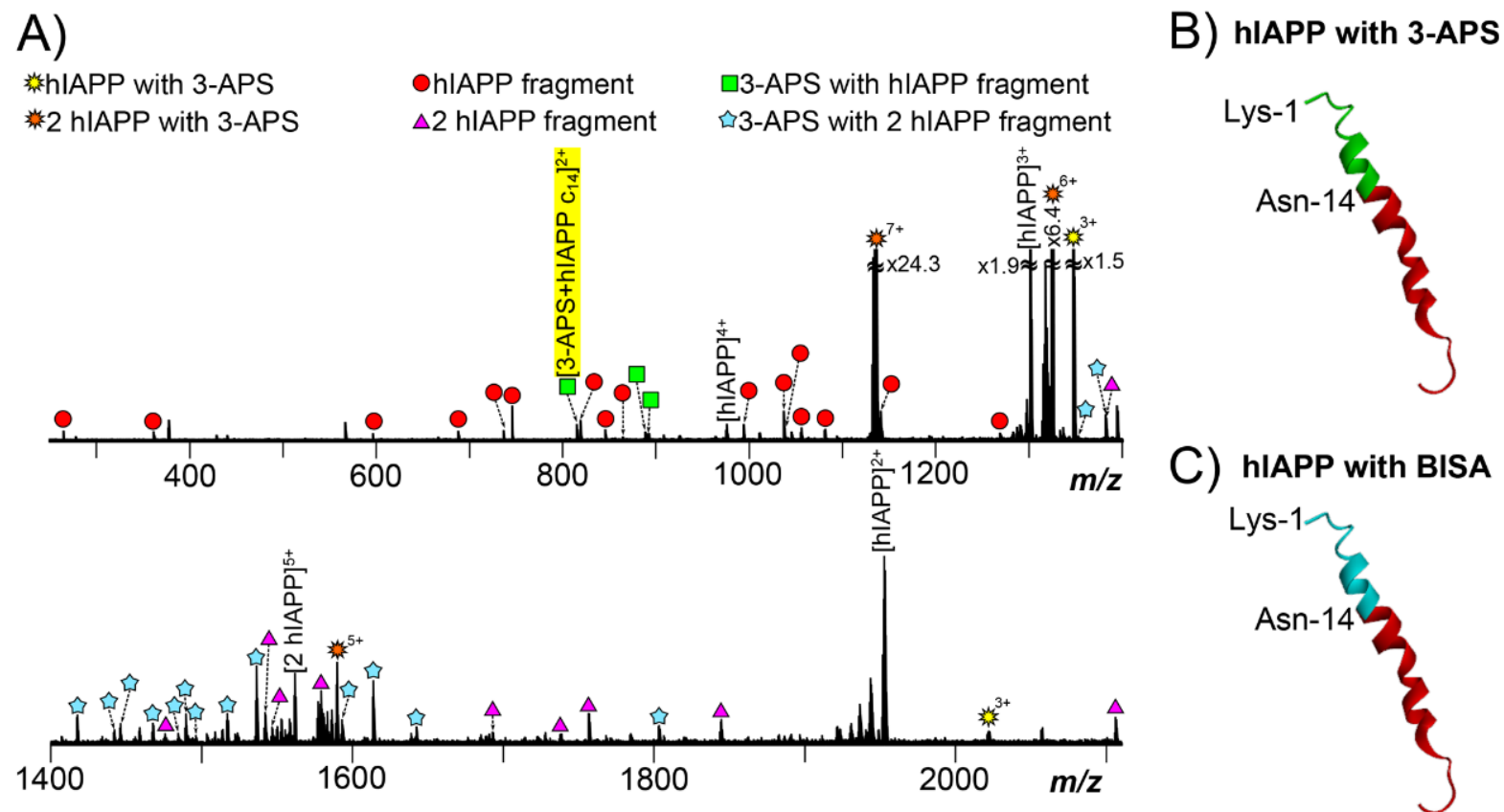
Figure 3.4. (A) ECD MS/MS spectrum of the 7+ molecular ion of hIAPP dimer bound with one 3-APS molecule. Critical fragments are highlighted. Binding regions are summarised on the protein structure for (B) 3-APS and (C) BISA.

The interactions between the target ^{30}TNVGSNTY^{37}-NH$_2$ peptide and the potential inhibitors. The aggregation region of hIAPP fibrils has previously been identified to be between Gly-33 and Ser-34, and the synthetic peptide ^{30}TNVGSNTY^{37}-NH$_2$ containing this region was shown to be amyloidogenic in previous studies.[48] Thus, the potential inhibitors were mixed with the target peptide in equimolar amount to investigate their inhibition effects against the amyloidogenic ^{30}TNVGSNTY^{37}-NH$_2$ peptide.

The early oligomers, from monomer to hexamer, of ^{30}TNVGSNTY^{37}-NH$_2$ peptide were observed in the nESI MS spectrum (see Figure 3.5A). When the peptide was incubated in the presence of insulin, only the low order oligomers, monomer and dimer, of the synthetic peptide were observed in the MS spectrum (see Figure 3.5B). The complexes, which composed of ^{30}TNVGSNTY^{37}-NH$_2$ and insulin, were also observed in the MS spectrum. The molecular binding ratios between ^{30}TNVGSNTY^{37}-NH$_2$ peptide and insulin ranged from 1:1 to 2:1.
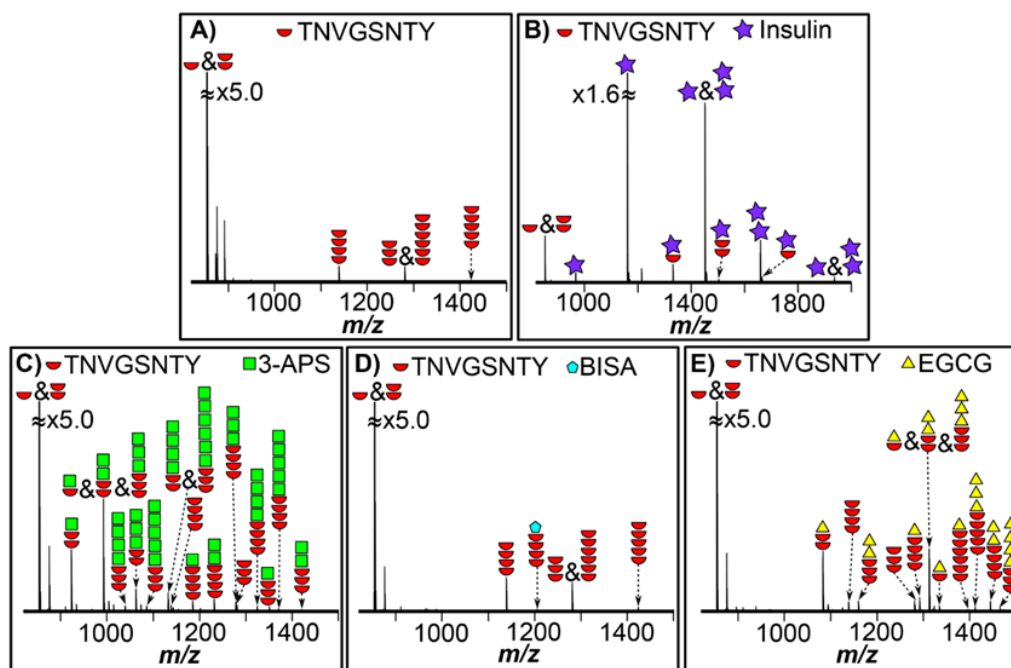
Figure 3.5. The MS spectra of 10 µM solution of ^{30}TNVGSNTY^{37}-NH$_2$ (A; red semi-circle) peptide only, mixed with equimolar of insulin (B; purple star), 3-APS (C; green square), BISA (D; cyan pentagon), and EGCG (E; yellow triangle).

Similar spectra were observed between 3-APS and EGCG which were mixed equimolar with ^{30}TNVGSNTY^{37}-NH$_2$ peptide as the early oligomers, from monomer to trimer, of ^{30}TNVGSNTY^{37}-NH$_2$ peptide were observed, along with those oligomers bound to 3-APS or EGCG (see Figure 3.5B & 3.5D). The molecular binding ratio between ^{30}TNVGSNTY^{37}-NH$_2$ peptide and 3-APS or EGCG ranged from 1:0.25 to 1:2 (for 3-APS) 1:0.2 to 1:1.3 (for EGCG). Less interaction was observed between ^{30}TNVGSNTY^{37}-NH$_2$ peptide and BISA molecule as only one reaction product was observed, the tetramer of ^{30}TNVGSNTY^{37}-NH$_2$ peptide and a BISA molecule (see Figure 3.5C), otherwise the MS spectrum was very similar to that of pure peptide (see Figure 3.5A).

The inhibition effects of the binding molecules on the mutant (isoD)₃hIAPP. hIAPP has been previously shown to be deamidated at Asn-21, Asn-22, and Asn-35.[48] Furthermore, it was also shown that deamidation mimics of hIAPP (mutant (isoD)₃hIAPP) could accelerate amyloid fibril formation, thus it is of particular interest to investigate whether potential inhibitors are effective not just in preventing amyloid protein aggregation, but also deamidated amyloid aggregation. The MS spectra of the mutant (isoD)₃hIAPP mixed with equimolar amounts of potential inhibitors, i.e. insulin, 3-APS, BISA, and EGCG, were acquired (see Supplementary Information, Figure S3.17). The MS spectrum of (isoD)₃hIAPP mixed with insulin showed abundant heterodimer peaks consisting of (isoD)₃hIAPP and insulin in 1:1 ratio. Other reaction mixtures with 3-APS, BISA, or EGCG showed a range of hIAPP monomers and oligomers bound to potential inhibitors in ratio of 1:0.5 to 1:5 (hIAPP to potential inhibitor).

The dissociation curves of (isoD)₃hIAPP bound to potential inhibitors are shown in Figure 3.6A – 3.6D. The small molecule species (3-APS, BISA, and EGCG) all showed a stronger binding towards the (isoD)₃hIAPP than the wild type hIAPP. Whereas the insulin species was dissociated more easily and so was bound more weakly to the (isoD)₃hIAPP compared to the wildtype. ECD MS/MS of the mutant (isoD)₃hIAPP with potential inhibitors showed the interaction regions were very similar to those regions observed above for the wild-type hIAPP (see Supplementary Information, Figure S3.18 – S3.22 and Table S3.11 – S3.15).
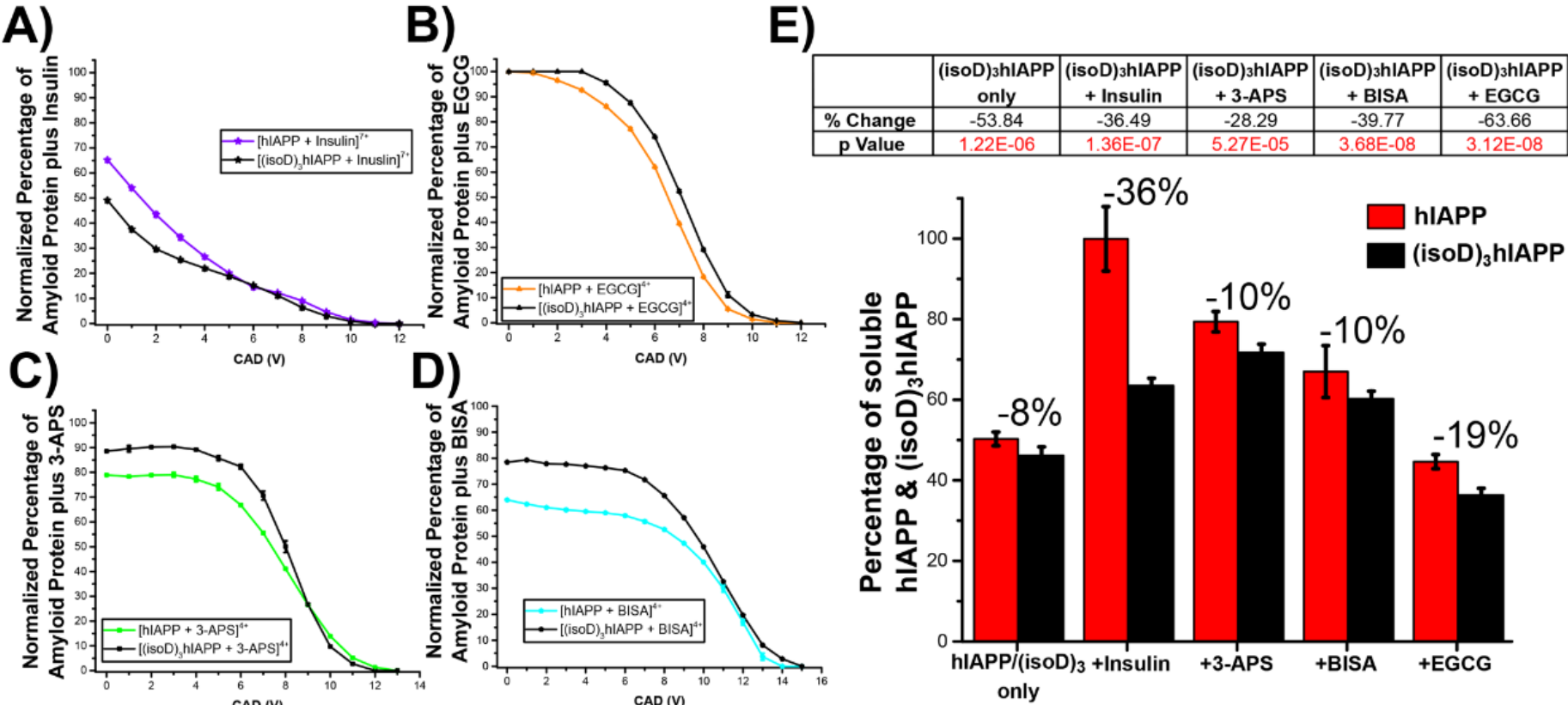
(Figure 3.6. Continue)
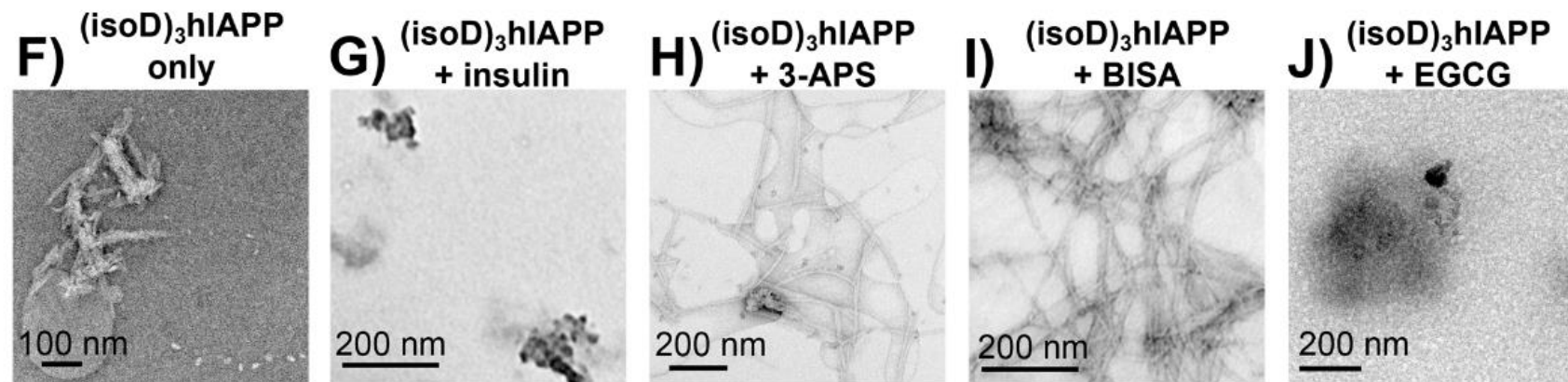
Figure 3.6. CAD MS/MS dissociation curves for wild-type hIAPP with potential inhibitors; insulin (A; purple), EGCG (B; orange), 3-APS (C; green), and BISA (D; cyan) against the mutant (isoD)₃hIAPP (coloured in black) mixed with the same molecules. (E) The percentage of the soluble hIAPP (red bar) and (isoD)₃hIAPP (black bar) present in the 10 µM incubated solutions incubated with potential inhibitors for one week, including t-test results for changes. TEM images of the one week incubated 10 µM solutions of (F) (isoD)₃hIAPP only, and the mixtures of (isoD)₃hIAPP with (G) insulin, (H) 3-APS, (I) BISA, and (J) EGCG molecules

The ThT fluorescence emission of the mutant (isoD)$_3$hIAPP with potential inhibitors showed the lag phases of aggregation also to be very similar to those observed for the wild-type hIAPP with the EGCG species showing the lowest amount of fibril formation (see Supplementary Information, Figure S3.23).

MS quantification of the (isoD)$_3$hIAPP with potential inhibitors incubated for one week were comparable to the trends observed for wild-type hIAPP with potential inhibitors above (see Figure 3.6E).   The amount of remaining soluble (isoD)$_3$hIAPP only solution was 46% after one week incubation, which was 8% less than the wild-type hIAPP.   Similar percentage differences were observed for the reaction mixtures containing 3-APS or BISA between wild-type hIAPP and (isoD)$_3$hIAPP.   However, EGCG showed a 19% change between hIAPP and (isoD)$_3$hIAPP solutions.   In the wild-type hIAPP solution incubated with insulin, the recovery of amyloid protein was 99.94% and the two-tailed t-test also showed that there was no significant difference between the fresh and incubated solutions, indicating essentially no aggregation of wild-type hIAPP occurring.   However, for (isoD)$_3$hIAPP incubated with insulin, only 64% of soluble amyloid protein was detected after one week incubation and the two-tailed t-test indicated there was a significant difference between the fresh and incubated solutions, indicating aggregation of (isoD)$_3$hIAPP.

TEM images of the incubated solutions of the (isoD)$_3$hIAPP with potential inhibitors showed the structure of the aggregates.   Short, dense fibrils were observed for the (isoD)$_3$hIAPP only (see Figure 3.6F); whereas elongated fibrils were observed in the TEM images of (isoD)$_3$hIAPP mixed with 3-APS or BISA (see Figure 3.6H and 3.6I).   No significant fibrils were observed in the TEM image of the incubated solution of (isoD)$_3$hIAPP mixed with insulin, however, several dense black blots were observed (see Figure 3.6G).   A TEM image of the

mutant (isoD)$_3$hIAPP mixed with EGCG showed large globular structures with no

distinct fibrillary structure (see Figure 3.6J).

3.5 Discussion

Reducing the formation of the early oligomers of hIAPP, as well as inhibiting the amyloid fibril formation, helps to enhance the success rate in islet cell transplantation for T2D patients. [14, 15]   Many studies focus on developing new compounds to inhibit the aggregation of hIAPP,[3, 10, 11] and recently a high throughput screening method was developed to quickly search for potential targets out of a hundred of compounds using IM-MS experiments.[22]   Understanding the mechanisms of potential inhibitors not only reduces the trial-and-error searching for potential compounds but, more importantly, can also provide guidance to create more effective therapeutics with fewer undesirable side reactions to work against aggregation-based conditions.   FTICR MS, together with fluorescence spectrometry and TEM, have been applied to explore the different inhibition mechanisms adopted by potential inhibitors of interest, i.e. insulin, 3-APS, BISA, and EGCG, to inhibit the amyloid fibril formation.

Specific binding to the aggregation region of the wild-type hIAPP and inhibition of fibril formation.   Incubated solutions of wild-type hIAPP aggregate rapidly to form early oligomers (see Figure 3.1A), followed by mature fibrils (see Figure 3.22C), which is consistent with the previous studies.[21, 48] However, when co-incubated with insulin, only monomer and dimer of hIAPP were detected, indicating inhibition of normal aggregation behaviour (see Figure 3.1B).   In addition, a heterodimer was detected indicating clear interactions between the two proteins, preventing usual aggregation.

The lag phase of ThT fluorescence emission and the MS quantification of the incubated solution represent the short- and long-term inhibition performance of insulin respectively.   The lag phase of the 50 µM solutions of the mixture of

hIAPP and insulin was significantly prolonged compared to the 50 µM solution of hIAPP alone, which shows insulin can extend the time frame of mature amyloid fibril formation (see Figure 3.2A). Significant amounts of mature fibrils were detected by fluorescence experiments, however the bulk emission was also, in part, due to insulin aggregation (see Supplementary Information, Figure S3.1). The MS quantification results of the one week hIAPP with insulin incubation solution showed a significant decrease in the percentage of soluble insulin (see Supplementary Information, Figure S3.2) and no significant change in the percentage of soluble hIAPP (see Figure 3.2B). Less dense fibrils (most likely insulin) and small dense regions (presumably hIAPP and insulin heterodimer) were simultaneously observed in the TEM image (see Figure 3.2D). The data acquired agree insulin is a promising short- and long-term potential inhibitor for wild-type hIAPP aggregation, and agrees with the previous studies using insulin to inhibit hIAPP aggregation.[32, 34]

Since insulin was shown to be an effective inhibitor in preventing the aggregation of wild-type hIAPP, it is important to explore the effective interaction sites between hIAPP and insulin molecules which may help for understanding and development of new therapeutics. The interaction between hIAPP and insulin is weak and non-covalently bound, which was demonstrated by the CAD MS/MS (see Supplementary Information, Figure S3.4). Using ECD MS/MS, the non-covalent interaction between hIAPP and insulin was shown to be between residues Ser-29 and Asn-35, located within the β-strand region of hIAPP (see Figure 3.3A – 3.3C). Residues of Asn-31 to Ser-34 on hIAPP molecule have previously been proposed as the most critical aggregation site for the formation of hIAPP early oligomers.[48] The result shown herein suggests insulin attaches to the residues of Ser-29 to Asn-35, which is the same aggregation region from which hIAPP forms

early oligomers, thus the hIAPP involved in the heterodimer complex cannot further aggregate to form higher order oligomers, as well as, mature fibrils.

The hypothesis that insulin is binding to the critical aggregation region of hIAPP and results in decelerating, and even inhibiting the formation of mature fibrils is also supported by the experiment that mixed the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide, which is a segment of proposed aggregation region of hIAPP with insulin. Early oligomers of ^{30}TNVGSNTY^{37}-NH$_2$ peptide were significantly reduced when insulin was incubated in the solution (see Figure 3.5B), indicating insulin can bind to the critical aggregation region of hIAPP and prevent the formation of early oligomers in aqueous solution.

Insulin has successfully demonstrated one of the inhibition mechanisms for hIAPP aggregation that is non-covalently and specifically bound to the critical aggregation region of hIAPP (between Ser-29 and Asn-35), which results in preventing the formation of the early oligomers of hIAPP, as well as, mature hIAPP fibrils (see Figure 3.7A).
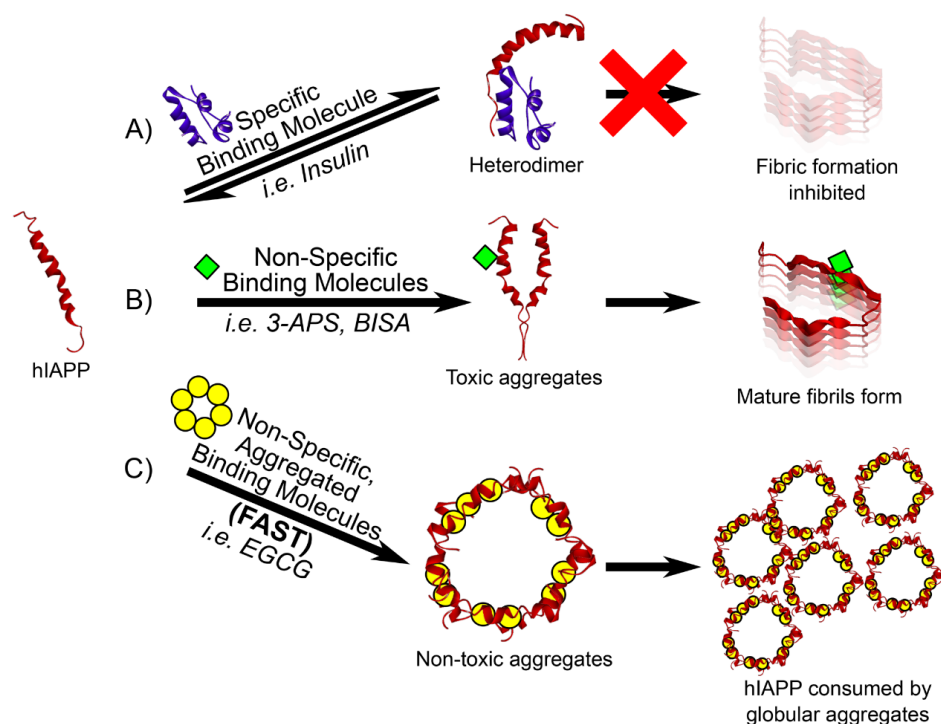
Figure 3.7. Summarised inhibition mechanisms of various compounds on wild-type hIAPP aggregation.

Ineffective, non-specific binding of potential inhibitors towards hIAPP. Despite effectively binding to hIAPP, 3-APS and BISA molecules bind in a non-specific way at multiple points across hIAPP (see Figure 3.2C and 3.2D). Higher order oligomers, i.e. trimers, etc., of hIAPP with an attachment of 3-APS/ BISA molecules were formed readily demonstrating 3-APS and BISA molecules are not effective in inhibiting the formation of hIAPP early oligomers.

The lag phases shown by the mixtures of hIAPP and 3-APS/ BISA were similar to the lag phase of hIAPP only solution, indicating these molecules do not prevent the initial formation of amyloid fibrils either (see Figure 3.2A). The apparent different of the percentage of soluble hIAPP present in between the fresh and incubated solution with addition of 3-APS/ BISA molecules also agrees both molecules are not effective inhibitors for hIAPP aggregation (see Figure 3.2B).

However, the amounts of soluble hIAPP in the incubated solutions with an addition of equal molar of 3-ASP/ BISA molecules were 29% and 17% higher than the incubated solution of hIAPP only, suggesting 3-APS and BISA molecules do have a slight long-term effect in preventing some formation of mature hIAPP fibrils, and 3-APS has a higher inhibition effect than BISA.   Normal, elongated fibrils were observed in the correlating TEM images which also shows 3-APS and BISA do not effectively inhibit the formation of mature hIAPP fibrils (see Figure 3.2E and 3.2F), agreeing with the results shown in the previous studies.[22]

CAD MS/MS results clearly demonstrate the interactions between hIAPP and 3-APS/ BISA are fragile and non-covalent (see Supplementary Information, Figure S3.4).   The interaction region between a monomer unit of hIAPP and a 3-APS molecule is between Arg-11 and Gly-33 (see Figure 3.3D); while the interaction region between a monomer unit of hIAPP and BISA molecule is between Lys-1 to Asn-3, Arg-11 to Asn-14, and Ser-29 to Tyr-37 residues (see Figure 3.3E).   ECD MS/MS results indicate the sulfonic group of the BISA molecule is interacting between Arg-11 and Asn-14, while the benzimidazole group of the BISA molecule interacts with Lys-1 to Ans-3 and Ser-29 to tyr-37 regions, demonstrating one BISA molecule interacts with two different regions of hIAPP simultaneously. These results also show both 3-APS and BISA can interact with the critical aggregation region of hIAPP (Asn-31 to ser-34 residues), but do not significantly reduce the aggregation rate of wild-type hIAPP.

Non-specific binding was observed in the incubated samples of ^{30}TNVGSNTY^{37}-NH$_2$ peptide and 3-APS (see Figure 3.5C), which shows the capability of 3-APS interacted with the critical aggregation of hIAPP.   However, BISA showed limited binding to the peptide target (see Figure 3.5D), indicating the interaction between BISA and the critical aggregation of hIAPP is weak and

possibly easily reversible.   The lower extent of reaction between BISA and ^{30}TNVGSNTY^{37}-NH$_2$ peptide suggests BISA can interact in a non-specific fashion towards hIAPP, but mainly outside of the critical aggregation region (Thr-30 and Tyr-37). However 3-APS showed extensive binding to both the critical aggregation region peptide and the full hIAPP, indicating a non-specific binding behaviour directly including the aggregation region (Thr-30 and Tyr-37), yet still with limited inhibition of aggregation.

Even though the early oligomers of hIAPP were still formed despite binding to 3-APS/ BISA molecules, it is also important to explore the binding regions on hIAPP which do not interrupt the formation of hIAPP early oligomers, so called non-target regions for compounds.   The ECD MS/MS spectra of the mixtures of the hIAPP dimer bound to a 3-APS/ BISA molecule shows the uninterrupted binding region is between Lys-1 and Asn-14 on hIAPP molecule (see Figure 3.4), which suggests there is no inhibition effect on hIAPP oligomers formation when the molecules bind to this region.

3-APS and BISA molecules demonstrate non-specific binding mechanisms to various sites/regions of hIAPP, which may or may not include the binding towards the critical aggregation site of hIAPP.   However, the non-specific binding does not seem to decelerate/inhibit the formation of mature hIAPP fibrils effectively in short-term and only slightly decrease amyloid fibril formation in long-term (see Figure 3.7B).

Non-specific binding to the wild-type hIAPP to accelerate formation of non-toxic aggregates.   Numerous studies have demonstrated that EGCG not only effectively inhibits the aggregation of hIAPP, but is also useful in preventing other amyloid proteins aggregation.[21-24, 26, 70]   EGCG was shown to bind non-

specifically to the wild-type hIAPP molecule with several EGCG molecules interacting with even single hIAPP molecules. Early oligomers of hIAPP bound to many EGCG molecules were also observed in the MS spectrum (see Figure 3.1E), which indicates EGCG is very similar to 3-APS and BISA that they can bind to hIAPP in a non-specific way.

EGCG showed effective short-term inhibition effects against hIAPP aggregation as the lag phase of the mature fibril formation measured by fluorescence experiments was significantly prolonged and the amount of fibrils formed was the lowest amongst all compounds studied (see Figure 3.2A), which indicates EGCG is a potential compound for preventing the formation of mature amyloid fibrils. However, the long-term inhibition effect of EGCG on hIAPP aggregation measured by MS quantification shows the opposite conclusion that the amount of soluble hIAPP remaining in solution was the least among all the incubated solutions, including the incubated solution of purely hIAPP (see Figure 3.2B). These data indicate that EGCG may not actually purely inhibit mature fibril formation, but instead accelerate formation of globular aggregates which do not further aggregate to detectable levels for ThT fluorescence. TEM images of the incubated solution of hIAPP with EGCG helps to explain the results obtained from the fluorescence spectrometry and MS quantification results as no mature fibril were observed in the TEM images of hIAPP with EGCG, however many small circular aggregates were (see Figure 3.2G). TEM image help to explain the minor ThT fluorescence emission in the solution of hIAPP with EGCG as ThT only binds to the mature β-sheets of amyloid fibrils,[51, 71] when there is no extended β-sheet formation, i.e. the small, circular aggregates, ThT molecules cannot attach to those aggregates effectively and there is no significant fluorescence emission detected. However, the TEM image clearly showed hIAPP aggregated into

amorphous structures and this structure is dense enough to be easily separated by centrifugation, thus the MS quantification of the incubated solution of EGCG and hIAPP showed a large reduction in available, soluble hIAPP. These findings are very similar to the previous studies stating EGCG could redirect amyloid protein aggregation into a non-toxic pathway,[24] results herein further show that when EGCG is co-incubated with hIAPP, no detectable amyloid fibrils are formed, but many amorphous aggregates are observed.

Since the EGCG has a very different mechanism in inhibiting the amyloid fibril formation, it is interesting to explore the interaction between hIAPP and EGCG molecule. Similar to the other small molecule species, the interaction between EGCG and hIAPP is fragile and non-covalent in nature, as reflected by CAD MS/MS results (see Supplementary Information, Figure S3.4). The binding region of EGCG to the monomer unit of hIAPP molecule was found to be between Arg-11 and Ser-34 (see Figure 3.3F), very similar to the proposed interaction region of 3-APS and hIAPP (Arg-11 to Gly-33). The MS spectrum of the target peptide ^{30}TNVGSNTY^{37}-NH$_2$ and EGCG also suggests a non-specific interaction, and higher ordered oligomers with the attachment of EGCG were observed (see Figure 3.5E). Data from hIAPP with EGCG samples suggest similar interactions to those from hIAPP with 3-APS, but there is drastically different aggregation behaviour, possible due to EGCG's own aggregation behaviour.

TEM images of purely EGCG solutions showed huge numbers of globular aggregates (see Supplementary Information, Figure S3.25A). TEM image of the ^{30}TNVGSNTY^{37}-NH$_2$ peptide with EGCG shows the circular aggregates increased in size compared to those in the EGCG solution alone (see Supplementary Information, Figure S3.25D) and the aggregates further increased in size when

hIAPP was incorporated (see Figure 3.2G). These data suggested the circular aggregates observed in the TEM images are composed of the EGCG and available protein/ peptide species as the size of the vesicle changed when the length of the protein/ peptide varies.

The formation of the amorphous aggregates and the inclusion of amyloid biomolecules support the MS quantitation and fluorescence data above that hIAPP is being consumed into off-fibril aggregation pathways and the rate of EGCG aggregation and inclusion of biomolecules seems to be much faster than that of amyloid fibril formation as the TEM images are dominated by the amorphous aggregates, which do not readily solubilise back into the aqueous solution (see Figure 3.7D). The non-specific binding of EGCG and the creation of globular aggregates could also help explain similar effects on results for other amyloid proteins observed previously.[21-24, 26, 70, 72]

The inhibition performance of the binding molecules towards the mutant (isoD)$_3$hIAPP. Mutant (isoD)$_3$hIAPP is a synthetic deamidation mimic of hIAPP, with a replacement of iso-aspartic acid residue at Asn-21, Asn-22, and Asn-35, and has been reported as an acceleration factor for hIAPP aggregation and it is important to explore the interaction between the (isoD)$_3$hIAPP and the binding molecules.[48] MS spectra (see Supplementary Information, Figure S3.17) and ECD MS/MS data (see Supplementary Information, Figure S3.18 – 3.22) showed there was no significant difference between the interaction behaviour (specific/ non-specific) or the interaction sites between wild-type hIAPP and the mutant (isoD)$_3$hIAPP. However, there was a significant difference of the interaction strength between the molecules and the amyloid proteins (see Figure 3.6A – 3.6D). Insulin showed a weaker interaction to the mutant (isoD)$_3$hIAPP, while the

interactions between $(isoD)_3hIAPP$ and 3-APS/ BISA/ EGCG were stronger than normal hIAPP. The results indicate a change in the structure of $(isoD)_3hIAPP$ can affect the strength of interaction with large, specific binding insulin molecules distinctly, the same behavior did not transfer to small, non-specific binding molecules, i.e. 3-APS, BISA, and EGCG.

The amount of soluble $(isoD)_3hIAPP$ remaining when incubated with insulin for one week was significantly less compared to the wild-type hIAPP with insulin, indicating that the insulin molecule is much less effective at inhibiting the aggregation of the mutant $(isoD)_3hIAPP$ compared to the wild-type (see Figure 3.6B). The amount of the soluble $(isoD)_3hIAPP$ incubated with the small molecules decreased in similar percentage when compared with the amount of soluble hIAPP in the incubated solutions, agreeing with the previous studies showing that the mutant $(isoD)_3hIAPP$ can aggregate faster than the wild-type.[48] However, ThT fluorescence measurement showed the lag phase and the amount of mature fibrils formed did not dramatically change compared to the wild-type hIAPP with insulin incubation (see Supplementary Information, Figure S3.21). TEM image of the mutant $(isoD)_3hIAPP$ with insulin incubated solution showed amorphous aggregates were formed, and as a result it did not produce a detectable ThT emission signal, in a similar way to wild-type hIAPP with EGCG above.

Mature fibrils were still observed in purely $(isoD)_3hIAPP$ solution, as well as when incubated with 3-APS/ BISA, indicating 3-APS/ BISA does not have any significant effect on $(isoD)_3hIAPP$ aggregation compared to wild-type hIAPP. TEM image of EGCG with $(isoD)_3hIAPP$ showed larger, denser black blot formation compared to the wild-type, suggesting a more extensive inclusion of protein into the circular aggregates. These were reflected in the MS quantitation results, showing a larger decrease in soluble $(isoD)_3hIAPP$ compared to the wild-

type.

The results here show that 3-APS and BISA molecules do not have significant effect on inhibiting the mutant $(isoD)_3$hIAPP. Previously, insulin was good at inhibiting the wild-type hIAPP aggregation and preserving the soluble hIAPP monomer in the solution. The mutant $(isoD)_3$hIAPP incubated with insulin forms amorphous structures which are different from the expected amyloid fibrils, and this complex is no longer reversibly soluble into aqueous solution. The EGCG molecule still accelerates the aggregation of $(isoD)_3$hIAPP into amorphous structures and the rate of this aggregation is also faster than the $(isoD)_3$hIAPP solution aggregation.

3.6 Conclusion

The experimental results here demonstrated two potential inhibition pathways are occurring to prevent the formation of toxic hIAPP aggregates.   Inhibitors like insulin adopt a direct pathway that interact with hIAPP between Ser-29 and Asn-35 residues which were proposed to be the most critical aggregation region for the formation of hIAPP early oligomers.   These inhibitors are not only effective in preventing the formation of toxic hIAPP aggregates, but also able to dissociate the hetero-complexes back into the monomer units of hIAPP and inhibitor, which is a possible key feature for potential drug development.   These inhibitors, however, are only effective towards a limited range of amyloid protein sequences and structures, as a result, they are not viable to prevent the aggregation of mutant hIAPPs or other amyloid proteins.   On the other hand, other inhibitors, such as EGCG, can self-aggregate and trigger the formation of non-toxic, amorphous hIAPP aggregates faster than forming the toxic hIAPP aggregates.   These inhibitors interact non-specifically to the amyloid proteins, thus they are effective to prevent the formation of toxic aggregates regardless of the sequences and structures of amyloid proteins.   Despite the inhibition effect, the formation of amorphous aggregates seems to indicate the hIAPP units are continuously consumed once mixed with these inhibitors, in a long term, very low or even no free hIAPP unit is available for the normal physiological function inside the human body.

3.7 References

1. Soto, C., Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat. Rev. Neurosci.* **2003,** *4* (1), 49-60.

2. Ross, C. A.; Poirier, M. A., Protein aggregation and neurodegenerative disease. *Nat. Med.* **2004,** *10*, S10-S17.

3. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human disease. In *Annu. Rev. Biochem*, Annual Reviews: Palo Alto, 2006; Vol. 75, pp 333-366.

4. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M., The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** *15* (6), 384-396.

5. Noble, M. E.; Endicott, J. A.; Johnson, L. N., Protein kinase inhibitors: insights into drug design from structure. *Science* **2004,** *303* (5665), 1800-1805.

6. Hajduk, P. J.; Greer, J., A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery* **2007,** *6* (3), 211-219.

7. Westermark, P.; Wernstedt, C.; Wilander, E.; Hayden, D. W.; O'Brien, T. D.; Johnson, K. H., Amyloid fibrils in human insulinoma and islets of Langerhans of the diabetic cat are derived from a neuropeptide-like protein also present in normal islet cells. *Proc. Natl. Acad. Sci. USA* **1987,** *84* (11), 3881-3885.

8. Cooper, G.; Willis, A.; Clark, A.; Turner, R.; Sim, R.; Reid, K., Purification and characterization of a peptide from amyloid-rich pancreases of type 2 diabetic patients. *Proc. Natl. Acad. Sci. USA* **1987,** *84* (23), 8628-8632.

9. Luca, S.; Yau, W. M.; Leapman, R.; Tycko, R., Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *J. Biochem.* **2007,** *46* (47), 13505-13522.

10. Cao, P.; Marek, P.; Noor, H.; Patsalo, V.; Tu, L.-H.; Wang, H.; Abedini, A.; Raleigh, D. P., Islet amyloid: from fundamental biophysics to mechanisms of cytotoxicity. *FEBS Lett.* **2013,** *587* (8), 1106-1118.

11. Akter, R.; Cao, P.; Noor, H.; Ridgway, Z.; Tu, L. H.; Wang, H.; Wong, A. G.; Zhang, X. X.; Abedini, A.; Schmidt, A. M.; Raleigh, D. P., Islet Amyloid Polypeptide: Structure, Function, and Pathophysiology. *J. Diabetes Res.* **2016**.

12. Lukinius, A.; Wilander, E.; Westermark, G.; Engström, U.; Westermark, P., Co-localization of islet amyloid polypeptide and insulin in the B cell secretory granules of the human pancreatic islets. *Diabetologia* **1989,** *32* (4), 240-244.

13. Cooper, G. J. S.; Day, A. J.; Willis, A. C.; Roberts, A. N.; Reid, K. B. M.; Leighton, B., Amylin and the amylin gene: structure, function and relationship to islet amyloid and to diabetes mellitus. *Biochim. Biophys. Acta* **1989,** *1014* (3), 247-258.

14. Clark, A.; Wells, C.; Buley, I.; Cruickshank, J.; Vanhegan, R.; Matthews, D.; Cooper, G.; Holman, R.; Turner, R., Islet amyloid, increased A-cells, reduced B-cells and exocrine fibrosis: quantitative changes in the pancreas in type 2 diabetes. *Diabetes research (Edinburgh, Scotland)* **1988,** *9* (4), 151-159.

15. Lorenzo, A.; Razzaboni, B.; Weir, G. C.; Yankner, B. A., Pancreatic islet cell toxicity of amylin associated with type-2 diabetes mellitus. *Nature* **1994,** *368* (6473), 756-760.

16. Westermark, G. T.; Westermark, P.; Berne, C.; Korsgren, O.; Nordic Network Clin Islet, T., Widespread amyloid deposition in transplanted human pancreatic islets. *New Engl. J. Med.* **2008,** *359* (9), 977-979.

17. Potter, K.; Abedini, A.; Marek, P.; Klimek, A.; Butterworth, S.; Driscoll, M.; Baker, R.; Nilsson, M.; Warnock, G.; Oberholzer, J., Islet amyloid deposition limits the viability of human islet grafts but not porcine islet grafts. *Proc. Natl. Acad. Sci. USA* **2010,** *107* (9), 4305-4310.

18. Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C., Islet amyloid in type 2 diabetes, and the toxic oligomer hypothesis. *Endocr. Rev.* **2008,** *29* (3), 303-316.

19. Meng, F.; Abedini, A.; Plesner, A.; Verchere, C. B.; Raleigh, D. P., The flavanol (−)-epigallocatechin 3-gallate inhibits amyloid formation by islet amyloid polypeptide, disaggregates amyloid fibrils, and protects cultured cells against IAPP-induced toxicity. *J. Biochem.* **2010,** *49* (37), 8127-8133.

20. Cao, P.; Raleigh, D. P., Analysis of the inhibition and remodeling of islet amyloid polypeptide amyloid fibers by flavanols. *J. Biochem.* **2012,** *51* (13), 2670-2683.

21. Young, L. M.; Cao, P.; Raleigh, D. P.; Ashcroft, A. E.; Radford, S. E., Ion Mobility Spectrometry-Mass Spectrometry Defines the Oligomeric Intermediates in Amylin Amyloid Formation and the Mode of Action of Inhibitors. *J. Am. Chem. Soc.* **2014,** *136* (2), 660-670.

22. Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E., Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nat. Chem.* **2015,** *7* (1), 73-81.

23. Hudson, S. A.; Ecroyd, H.; Dehle, F. C.; Musgrave, I. F.; Carver, J. A., (−)-Epigallocatechin-3-gallate (EGCG) maintains κ-casein in its pre-fibrillar state without redirecting its aggregation pathway. *J. Mol. Biol.* **2009,** *392* (3), 689-700.

24. Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E., EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.* **2008,** *15* (6), 558-566.

25.   Porat, Y.; Abramowitz, A.; Gazit, E., Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism. *Chem. Biol. Drug Des.* **2006,** *67* (1), 27-37.

26.   Huang, R.; Vivekanandan, S.; Brender, J. R.; Abe, Y.; Naito, A.; Ramamoorthy, A., NMR characterization of monomeric and oligomeric conformations of human calcitonin and its interaction with EGCG. *J. Mol. Biol.* **2012,** *416* (1), 108-120.

27.   Brange, J.; Langkjœr, L., Insulin structure and stability. In *Stability and Characterization of Protein and Peptide Drugs*, Springer: 1993; pp 315-350.

28.   Jaikaran, E.; Nilsson, M. R.; Clark, A., Pancreatic beta-cell granule peptides form heteromolecular complexes which inhibit islet amyloid polypeptide fibril formation. *Biochem. J* **2004,** *377*, 709-716.

29.   Brange, J.; Andersen, L.; Laursen, E. D.; Meyn, G.; Rasmussen, E., Toward understanding insulin fibrillation. *J. Pharm. Sci.* **1997,** *86* (5), 517-525.

30.   Jimenez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R., The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci. USA* **2002,** *99* (14), 9196-9201.

31.   Ahmad, A.; Uversky, V. N.; Hong, D.; Fink, A. L., Early events in the fibrillation of monomeric insulin. *J. Biol. Chem.* **2005,** *280* (52), 42669-42675.

32.   Wiltzius, J. J. W.; Sievers, S. A.; Sawaya, M. R.; Eisenberg, D., Atomic structures of IAPP (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process. *Protein Sci.* **2009,** *18* (7), 1521-1530.

33.   Landreh, M.; Alvelius, G.; Johansson, J.; Jornvall, H., Insulin, islet amyloid polypeptide and C-peptide interactions evaluated by mass spectrometric analysis. *Rapid Commun. Mass Spectrom.* **2014,** *28* (2), 178-184.

34.   Susa, A. C.; Wu, C.; Bernstein, S. L.; Dupuis, N. F.; Wang, H.; Raleigh, D. P.; Shea, J. E.; Bowers, M. T., Defining the Molecular Basis of Amyloid Inhibitors: Human Islet Amyloid Polypeptide-Insulin Interactions. *J. Am. Chem. Soc.* **2014,** *136* (37), 12912-12919.

35.   Nedumpully-Govindan, P.; Ding, F., Inhibition of IAPP aggregation by insulin depends on the insulin oligomeric state regulated by zinc ion concentration. *Sci. Rep.* **2015,** *5*, 7.

36.   Robinson, N. E.; Robinson, A. B., Molecular clocks. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (3), 944-949.

37.   Robinson, N. E.; Robinson, A. B., Deamidation of human proteins. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (22), 12409-12413.

38.   Robinson, A. B.; Rudd, C. J., Deamidation of glutaminyl and asparaginyl residues in peptides and proteins. *Curr. Top. Cell. Regul.* **1974,** *8* (0), 247-95.

39. Robinson, N. E.; Robinson, A., *Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins*. Althouse press: Oregon, 2004.

40. Mamula, M. J.; Gee, R. J.; Elliott, J. I.; Sette, A.; Southwood, S.; Jones, P. J.; Blier, P. R., Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J. Biol. Chem.* **1999,** *274* (32), 22321-22327.

41. Doyle, H. A.; Mamula, M. J., Post-translational protein modifications in antigen recognition and autoimmunity. *Trends Immunol.* **2001,** *22* (8), 443-449.

42. Shimizu, T.; Watanabe, A.; Ogawara, M.; Mori, H.; Shirasawa, T., Isoaspartate formation and neurodegeneration in Alzheimer's disease. *Arch. Biochem. Biophys.* **2000,** *381* (2), 225-234.

43. Kad, N. M.; Thomson, N. H.; Smith, D. P.; Smith, D. A.; Radford, S. E., beta(2)-microglobulin and its deamidated variant, N17D form amyloid fibrils with a range of morphologies in vitro. *J. Mol. Biol.* **2001,** *313* (3), 559-571.

44. Shimizu, T.; Fukuda, H.; Murayama, S.; Izumiyama, N.; Shirasawa, T., Isoaspartate formation at position 23 of amyloid beta peptide enhanced fibril formation and deposited onto senile plaques and vascular amyloids in Alzheimer's disease. *J. Neurosci. Res.* **2002,** *70* (3), 451-461.

45. Gupta, R.; Srivastava, O. P., Effect of deamidation of asparagine 146 on functional and structural properties of human lens alpha B-crystallin. *Invest. Ophthalmol. Vis. Sci.* **2004,** *45* (1), 206-214.

46. Tomidokoro, Y.; Rostagno, A.; Neubert, T. A.; Lu, Y.; Rebeck, G. W.; Frangione, B.; Greenberg, S. M.; Ghiso, J., Iowa Variant of Familial Alzheimer's Disease Accumulation of Posttranslationally Modified A beta D23N in Parenchymal and Cerebrovascular Amyloid Deposits. *Am. J. Pathol.* **2010,** *176* (4), 1841-1854.

47. Soulby, A. J.; Heal, J. W.; Barrow, M. P.; Roemer, R. A.; O'Connor, P. B., Does deamidation cause protein unfolding? A top-down tandem mass spectrometry study. *Protein Sci.* **2015,** *24* (5), 850-860.

48. Lam, Y. P. Y.; Wootton, C. A.; Hands-Portman, I.; Wei, J.; Chiu, C. K. C.; I. Romero-Canelon; Lermyte, F.; Barrow, M. P.; O'Connor, P. B., Exploring the Aggregation Mechanisms of Amyloid Protofibrils Using Electron Capture Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2018,** *Submitted*.

49. Nilsson, M. R.; Driscoll, M.; Raleigh, D. P., Low levels of asparagine deamidation can have a dramatic effect on aggregation of amyloidogenic peptides: Implications for the study of amyloid formation. *Protein Sci.* **2002,** *11* (2), 342-349.

50. Dunkelberger, E. B.; Buchanan, L. E.; Marek, P.; Cao, P.; Raleigh, D. P.;

Zanni, M. T., Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. *J. Am. Chem. Soc.* **2012,** *134* (30), 12658-12667.

51.   Biancalana, M.; Koide, S., Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta, Proteins Proteomics* **2010,** *1804* (7), 1405-1412.

52.   Bartolini, M.; Bertucci, C.; Cavrini, V.; Andrisano, V., β-Amyloid aggregation induced by human acetylcholinesterase: inhibition studies. *Biochem. Pharmacol.* **2003,** *65* (3), 407-416.

53.   Meier, J. J.; Kayed, R.; Lin, C.-Y.; Gurlo, T.; Haataja, L.; Jayasinghe, S.; Langen, R.; Glabe, C. G.; Butler, P. C., Inhibition of human IAPP fibril formation does not prevent β-cell death: evidence for distinct actions of oligomers and fibrils of human IAPP. *Am. J. Physiol. Endocrinol. Metab.* **2006,** *291* (6), E1317-E1324.

54.   Hudson, S. A.; Ecroyd, H.; Kee, T. W.; Carver, J. A., The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds. *The FEBS journal* **2009,** *276* (20), 5960-5972.

55.   Coelho-Cerqueira, E.; Pinheiro, A. S.; Follmer, C., Pitfalls associated with the use of Thioflavin-T to monitor anti-fibrillogenic activity. *Bioorg. Med. Chem. Lett.* **2014,** *24* (14), 3194-3198.

56.   Dupuis, N. F.; Wu, C.; Shea, J. E.; Bowers, M. T., Human Islet Amyloid Polypeptide Monomers Form Ordered beta-hairpins: A Possible Direct Amyloidogenic Precursor. *J. Am. Chem. Soc.* **2009,** *131* (51), 18283-18292.

57.   Dupuis, N. F.; Wu, C.; Shea, J. E.; Bowers, M. T., The Amyloid Formation Mechanism in Human IAPP: Dimers Have beta-Strand Monomer-Monomer Interfaces. *J. Am. Chem. Soc.* **2011,** *133* (19), 7240-7243.

58.   Young, L. M.; Mahood, R. A.; Saunders, J. C.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E., Insights into the consequences of co-polymerisation in the early stages of IAPP and Abeta peptide assembly from mass spectrometry. *The Analyst* **2015,** *140* (20), 6990-9.

59.   Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Ashcroft, A. E.; Radford, S. E., ESI-IMS-MS: A method for rapid analysis of protein aggregation and its inhibition by small molecules. *Methods* **2016,** *95*, 62-69.

60.   Young, L. M.; Ashcroft, A. E.; Radford, S. E., Small molecule probes of protein aggregation. *Curr. Opin. Chem. Biol.* **2017,** *39*, 90-99.

61.   Breuker, K.; McLafferty, F. W., Native electron capture dissociation for the structural characterization of noncovalent interactions in native cytochrome c. *Angew. Chem. Int. Ed.* **2003,** *42* (40), 4900-4904.

62.   O'Connor, P. B., Uses of electron capture dissociation in biomolecules. *Abstr.*

Pap. Am. Chem. Soc. **2010,** *240*.

63.   Qi, Y.; Liu, Z.; Li, H.; Sadler, P. J.; O'Connor, P. B., Mapping the protein-binding sites for novel iridium(III) anticancer complexes using electron capture dissociation. *Rapid Commun. Mass Spectrom.* **2013,** *27* (17), 2028-2032.

64.   Caravatti, P.; Allemann, M., The 'Infinity Cell': a New Trapped-ion Cell With Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991,** *26* (5), 514-518.

65.   De Koning, L.; Nibbering, N.; Van Orden, S.; Laukien, F., Mass selection of ions in a Fourier transform ion cyclotron resonance trap using correlated harmonic excitation fields (CHEF). *Int. J. Mass Spectrom. Ion Processes* **1997,** *165*, 209-219.

66.   Chan, F. T. S.; Schierle, G. S. K.; Kumita, J. R.; Bertoncini, C. W.; Dobson, C. M.; Kaminski, C. F., Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst* **2013,** *138* (7), 2156-2162.

67.   Yang, T.; Wang, X. H.; Zhang, C. L.; Ma, X.; Wang, K.; Wang, Y. Q.; Luo, J.; Yang, L.; Yao, C.; Wang, X. Y., Specific self-monitoring of metal-associated amyloid-beta peptide disaggregation by a fluorescent chelator. *Chem. Commun.* **2016,** *52* (11), 2245-2248.

68.   Arosio, P.; Knowles, T. P.; Linse, S., On the lag phase in amyloid fibril formation. *PCCP* **2015,** *17* (12), 7606-7618.

69.   Shoffner, S. K.; Schnell, S., Estimation of the lag time in a subsequent monomer addition model for fibril elongation. *PCCP* **2016,** *18* (31), 21259-21268.

70.   Bieschke, J.; Russ, J.; Friedrich, R. P.; Ehrnhoefer, D. E.; Wobst, H.; Neugebauer, K.; Wanker, E. E., EGCG remodels mature α-synuclein and amyloid-β fibrils and reduces cellular toxicity. *Proc. Natl. Acad. Sci. USA* **2010,** *107* (17), 7710-7715.

71.   Groenning, M., Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils—current status. *J. Chem. Biol.* **2010,** *3* (1), 1-18.

72.   Chandrashekaran, I. R.; Adda, C. G.; MacRaild, C. A.; Anders, R. F.; Norton, R. S., EGCG disaggregates amyloid-like fibrils formed by Plasmodium falciparum merozoite surface protein 2. *Arch. Biochem. Biophys.* **2011,** *513* (2), 153-157.

73.   Huang, Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H., Statistical characterization of the charge state and residue dependence of low-energy CID peptide dissociation patterns. *Anal. Chem.* **2005,** *77* (18), 5800-5813.

3.8 Supplementary Information
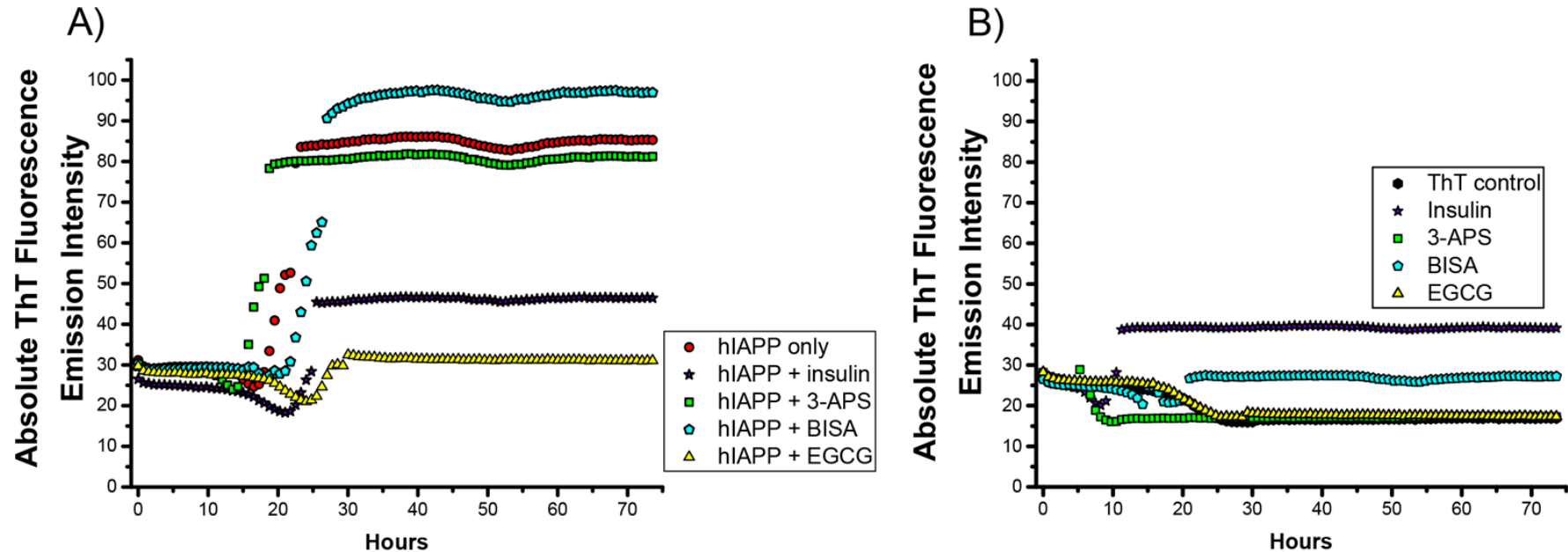
(Figure S3.1. Continue)
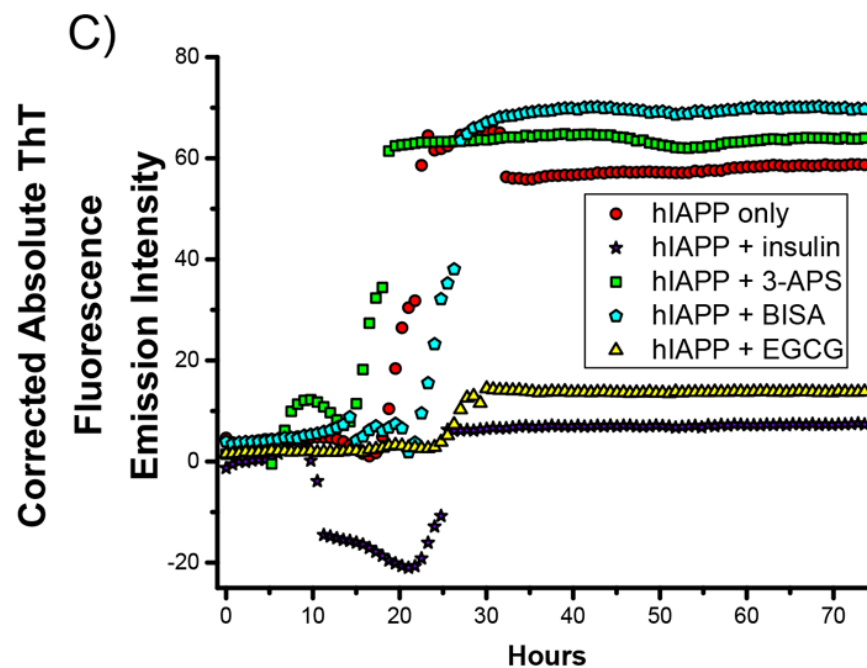
Figure S3.1. (A) The absolute ThT fluorescence emission of a 50 µM solution of the wild-type hIAPP (red circles), and the 50 µM solutions of

the wild-type hIAPP mixed with equal molar of insulin (purple star), 3-APS (green rectangle), BISA (cyan pentagon), and EGCG (yellow triangle).

(B) The absolute ThT fluorescence emission of the 100 µM ThT solution (black hexagon), and the 50 µM solutions of insulin (purple star), 3-APS

(green rectangle), BISA (cyan pentagon), and EGCG (yellow triangle) alone.   (C) The corrected absolute ThT fluorescence emission of the 50

µM solution of the wild-type hIAPP (red circles), and the 50 µM solutions of the wild-type hIAPP mixed with equal molar of insulin (purple star),

3-APS (green rectangle), BISA (cyan pentagon), and EGCG (yellow triangle).   The corrected absolute ThT fluorescence emission was calculated

by using the absolute ThT fluorescence emission of hIAPP mixed with various binding molecules (see Figure S3.1A) minus the absolute ThT

fluorescence emission of the binding molecules alone (see Figure S3.1B).

	hIAPP only	hIAPP + Insulin	hIAPP + 3-APS	hIAPP + BISA	hIAPP + EGCG	Insulin Only
% Change	-49.71	-0.06	-20.63	-33.00	-55.35	-9.46
t-Test Value	5.16E-08	0.92	2.94E-05	3.27E-04	9.40E-09	0.01
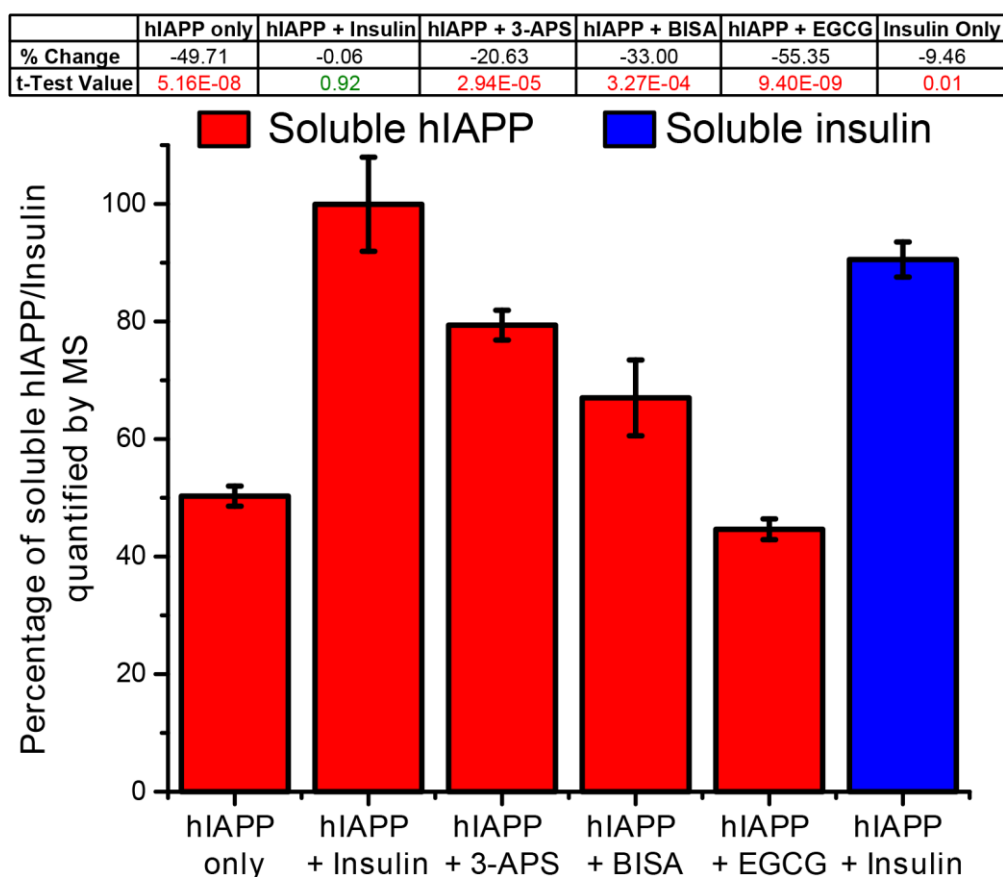
Figure S3.2. The percentage of the soluble wild-type hIAPP (red bar)/ insulin (blue bar) which remained in the 10 µM solutions incubated at 37°C for one week. The percentage of soluble hIAPP in the fresh solutions and the incubated solutions were test against a two-tailed T-test. The p value that is <0.05 (red font) indicates there is a significant different of the soluble hIAPP/ insulin percentage between the fresh and incubated solutions, while the p value >0.05 (green font) indicates there is no significant difference of the soluble hIAPP/ insulin percentage between the fresh and incubated solutions.
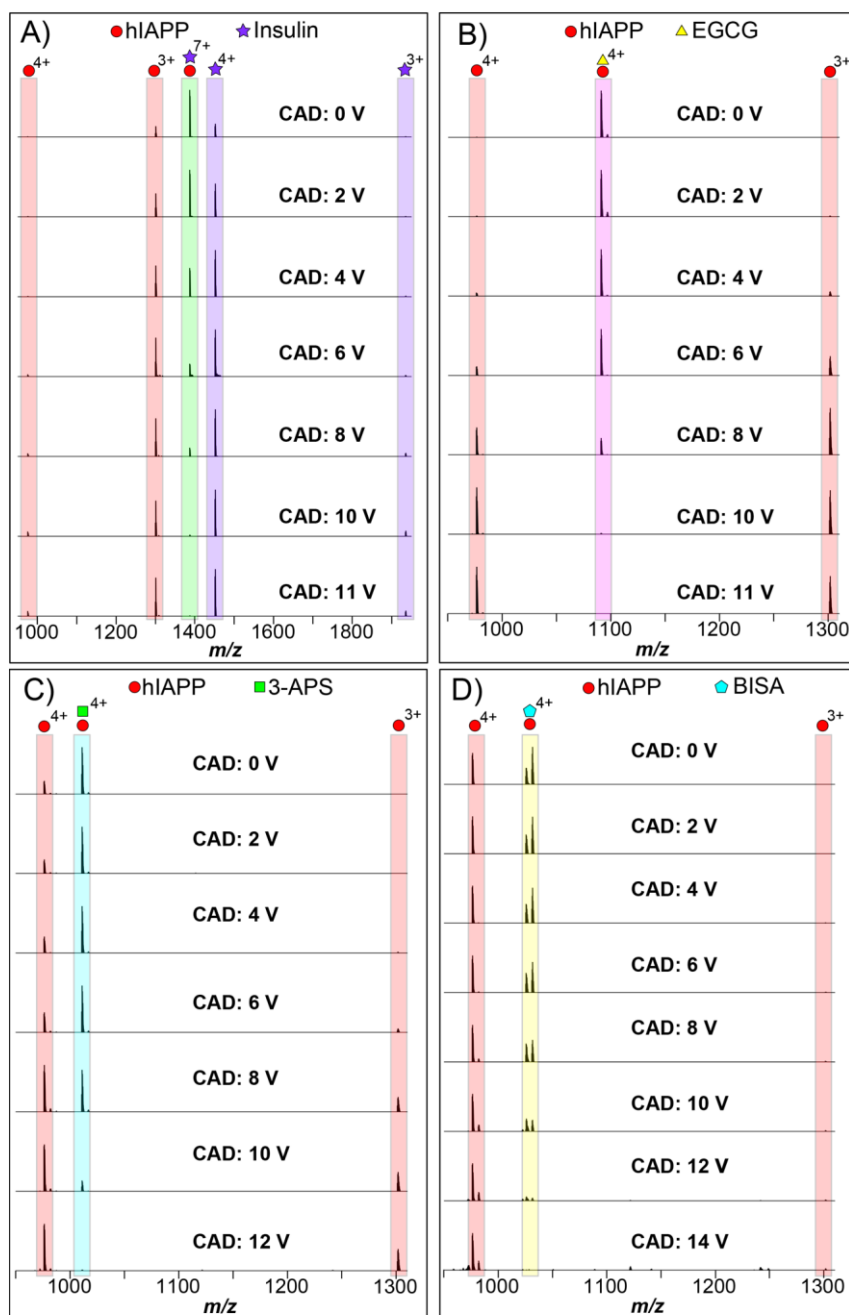
Figure S3.3. The representative CAD spectra of (A) the 7+ charge state molecular ion of the wild-type hIAPP monomer unit with an intact insulin molecule and the 4+ charge state molecular ion of the wild-type hIAPP monomer unit with an attachment of (B) an EGCG, (C) a 3-APS, and (D) a BISA molecules.
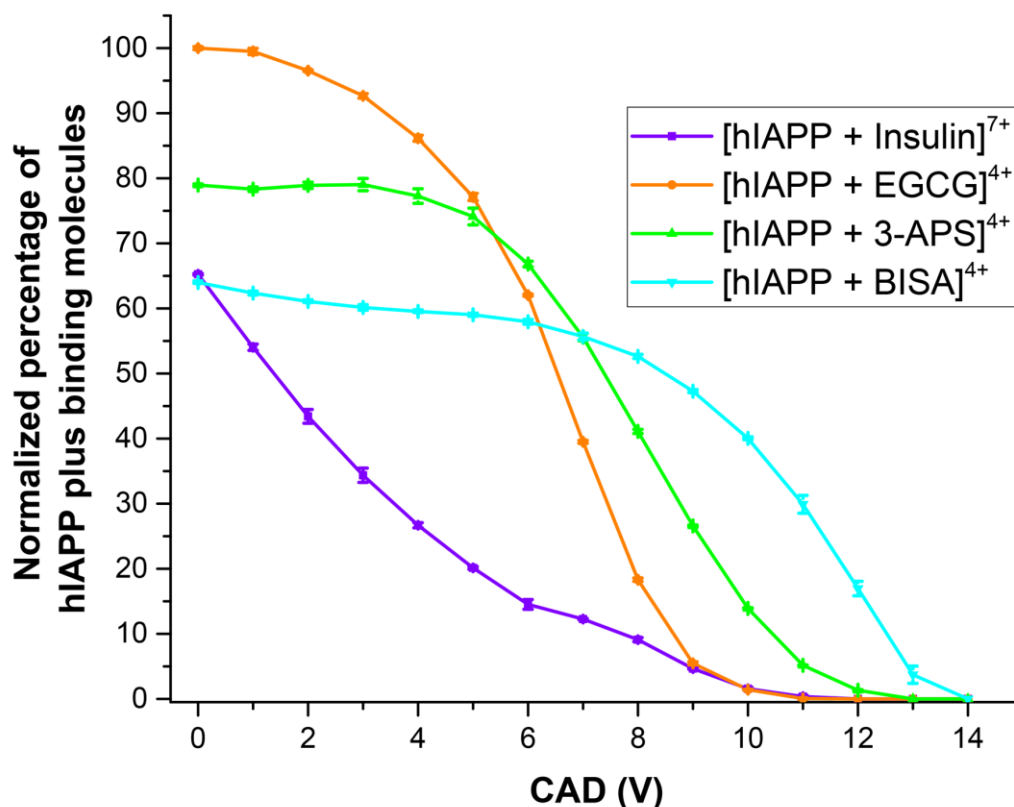
Figure S3.4. The dissociation curves of the 7+ hIAPP monomer unit with an insulin molecule (purple), and the 4+ hIAPP monomer unit with an EGCG (orange), a 3-APS (green), and a BISA (cyan) molecules which were generated from the Supplementary Information Figure S3.3. The dissociation curve of the 7+ charge state of hIAPP with insulin is the lowest among all the other dissociation curves, however, this is a bias comparison as the charge state of the molecular ion of hIAPP with insulin (7+) is higher than the others molecular ions (4+), and previous study has shown variation in the charge state of the peptide leads to different fragmentation pattern.[73] The dissociation curves of hIAPP monomer unit with the binding molecules clearly indicate that the interaction between hIAPP and the binding molecules are fragile and non-covalent bind as the dissociations resulted in obtaining the hIAPP monomer unit directly and no CAD MS/MS fragments were obtained.

A)

ECD fragments of Insulin Chain A

c ion

G I V E Q C C T S I C S L Y Q L E N Y C N

ECD fragments of Insulin Chain B

c ion

F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T

z ion

ECD fragments of hIAPP

c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y - NH$_2$

ECD fragments of hIAPP plus insulin
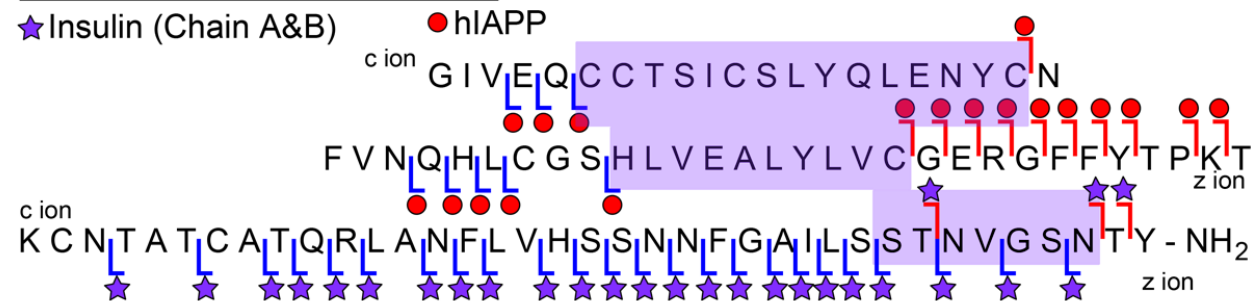
★ Insulin (Chain A&B)   ● hIAPP

c ion   G I V E Q C C T S I C S L Y Q L E N Y C N

F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T

z ion

c ion

K C N T A T C A T Q R L A N F L V H S S N N F G A I L S S T N V G S N T Y - NH$_2$

z ion

226

(Figure S3.5. Continue)
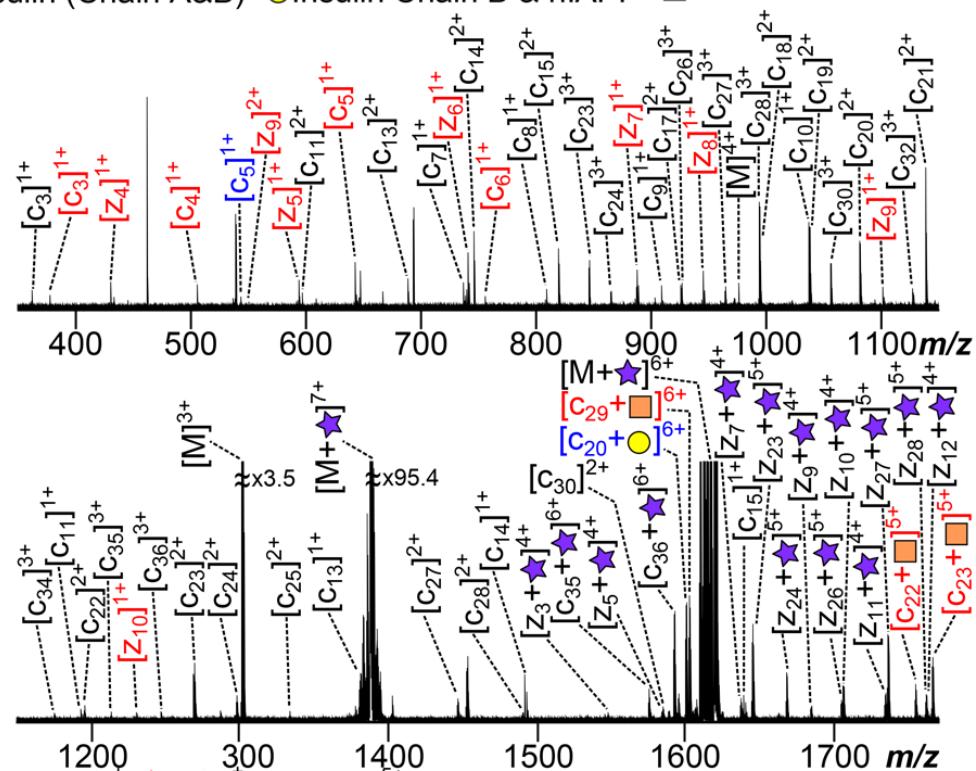
(Figure S3.5. Continue)

Figure S3.5. (A) The summarised ECD fragments of the 7+ charge state molecular ion of hIAPP monomer unit and an intact insulin molecule. The highlighted area (coloured in purple) indicates the proposed interaction region between the monomer unit of hIAPP and an insulin molecule. (B) The ECD MS/MS spectrum of the 7+ charge state molecular ion of hIAPP monomer unit and an intact insulin molecule. The assigned fragments are listed in the Supplementary Information Table S3.1.

Assigned ECD fragments of the 7+ molecular ion of hIAPP with insulin

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	2+	1952.444365	1952.445111	0.38
hIAPP only	hIAPP	M	3+	1301.295128	1301.295201	0.06
hIAPP only	hIAPP	M	4+	976.473865	976.473886	0.02
hIAPP only	hIAPP	c3	1+	363.180901	363.180902	0.00
hIAPP only	hIAPP	c7	1+	737.306908	737.306648	-0.35
hIAPP only	hIAPP	c8	1+	808.344022	808.343748	-0.34
hIAPP only	hIAPP	c9	1+	909.391701	909.391436	-0.29
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450161	-0.11
hIAPP only	hIAPP	c11	1+	1193.551390	1193.551587	0.17
hIAPP only	hIAPP	c11	2+	597.279333	597.279053	-0.47
hIAPP only	hIAPP	c13	1+	1377.672568	1377.673361	0.58
hIAPP only	hIAPP	c13	2+	689.339922	689.339635	-0.42
hIAPP only	hIAPP	c14	1+	1491.715496	1491.716126	0.42
hIAPP only	hIAPP	c14	2+	746.361386	746.361106	-0.38
hIAPP only	hIAPP	c15	1+	1638.783910	1638.784576	0.41
hIAPP only	hIAPP	c15	2+	819.895593	819.895342	-0.31
hIAPP only	hIAPP	c17	2+	925.971832	925.971641	-0.21
hIAPP only	hIAPP	c18	2+	994.501288	994.501061	-0.23
hIAPP only	hIAPP	c19	2+	1038.017303	1038.017126	-0.17
hIAPP only	hIAPP	c20	2+	1081.533315	1081.532980	-0.31
hIAPP only	hIAPP	c21	2+	1138.554781	1138.554657	-0.11
hIAPP only	hIAPP	c22	2+	1195.576245	1195.576315	0.06
hIAPP only	hIAPP	c23	2+	1269.110452	1269.110634	0.14
hIAPP only	hIAPP	c23	3+	846.743643	846.743554	-0.11
hIAPP only	hIAPP	c24	2+	1297.621184	1297.620936	-0.19
hIAPP only	hIAPP	c24	3+	865.750796	865.750609	-0.22
hIAPP only	hIAPP	c25	2+	1333.641112	1333.641354	0.18
hIAPP only	hIAPP	c26	3+	926.790274	926.789975	-0.32
hIAPP only	hIAPP	c27	2+	1446.223805	1446.223638	-0.12
hIAPP only	hIAPP	c27	3+	964.484962	964.484964	0.00
hIAPP only	hIAPP	c28	2+	1490.241206	1490.241217	0.01
hIAPP only	hIAPP	c28	3+	993.829896	993.829891	-0.01
hIAPP only	hIAPP	c30	2+	1583.779674	1583.779878	0.13
hIAPP only	hIAPP	c30	3+	1056.523134	1056.523072	-0.06
hIAPP only	hIAPP	c32	3+	1127.560247	1127.560364	0.10
hIAPP only	hIAPP	c34	3+	1175.912156	1175.912589	0.37
hIAPP only	hIAPP	c35	3+	1213.258129	1213.258544	0.34
hIAPP only	hIAPP	c36	3+	1247.274945	1247.274420	-0.42
Insulin A Chain only	Insulin A Chain	c5	1+	544.308939	544.308688	-0.46
Insulin B Chain only	Insulin B Chain	c3	1+	378.213581	378.213589	0.02
Insulin B Chain only	Insulin B Chain	c4	1+	506.272159	506.271966	-0.38
Insulin B Chain only	Insulin B Chain	c5	1+	643.331071	643.330778	-0.46
Insulin B Chain only	Insulin B Chain	c6	1+	756.415135	756.414724	-0.54
Insulin B Chain only	Insulin B Chain	z4	1+	431.250026	431.249939	-0.20
Insulin B Chain only	Insulin B Chain	z5	1+	594.313354	594.313110	-0.41
Insulin B Chain only	Insulin B Chain	z6	1+	741.381768	741.381527	-0.33
Insulin B Chain only	Insulin B Chain	z7	1+	888.450182	888.450083	-0.11
Insulin B Chain only	Insulin B Chain	z8	1+	945.471646	945.471448	-0.21
Insulin B Chain only	Insulin B Chain	z9	1+	1101.572757	1101.573033	0.25
Insulin B Chain only	Insulin B Chain	z9	2+	550.786105	550.785891	-0.39
Insulin B Chain only	Insulin B Chain	z10	1+	1230.615350	1230.615841	0.40
hIAPP plus Insulin	hIAPP plus Insulin	M	3+	3238.855716	3238.854187	-0.47
hIAPP plus Insulin	hIAPP plus Insulin	M	4+	2429.141494	2429.142002	0.21
hIAPP plus Insulin	hIAPP plus Insulin	M	5+	1943.313046	1943.313443	0.20
hIAPP plus Insulin	hIAPP plus Insulin	M	6+	1619.427447	1619.427392	-0.03
hIAPP plus Insulin	hIAPP plus Insulin	M	7+	1388.223731	1388.222704	-0.74
hIAPP plus Insulin	Insulin A Chain	c20	6+	1599.754276	1599.754214	-0.04
hIAPP plus Insulin	Insulin A Chain	z16	4+	2293.064081	2293.064589	0.22
hIAPP plus Insulin	Insulin A Chain	z17	4+	2325.078733	2325.078653	-0.03
hIAPP plus Insulin	Insulin A Chain	z18	5+	1885.871364	1885.871611	0.13
hIAPP plus Insulin	Insulin B Chain	c19	4+	2107.482415	2107.483135	0.34
hIAPP plus Insulin	Insulin B Chain	c20	4+	2121.487385	2121.487270	-0.05
hIAPP plus Insulin	Insulin B Chain	c21	4+	2153.497521	2153.497499	-0.01
hIAPP plus Insulin	Insulin B Chain	c22	4+	2192.773330	2192.773427	0.04
hIAPP plus Insulin	Insulin B Chain	c22	5+	1754.219685	1754.219472	-0.12
hIAPP plus Insulin	Insulin B Chain	c23	4+	2207.781551	2207.781806	0.12
hIAPP plus Insulin	Insulin B Chain	c23	5+	1765.623980	1765.623979	0.00
hIAPP plus Insulin	Insulin B Chain	c24	4+	2244.548693	2244.547962	-0.33
hIAPP plus Insulin	Insulin B Chain	c24	5+	1795.236892	1795.236856	-0.02
hIAPP plus Insulin	Insulin B Chain	c25	4+	2280.814911	2280.815231	0.14
hIAPP plus Insulin	Insulin B Chain	c25	5+	1824.450207	1824.450622	0.23

(Table S3.1. Continue)

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP plus Insulin	Insulin B Chain	c26	4+	2321.580766	2321.581731	0.42
hIAPP plus Insulin	Insulin B Chain	c26	5+	1856.862493	1856.862979	0.26
hIAPP plus Insulin	Insulin B Chain	c28	4+	2371.356458	2371.355481	-0.41
hIAPP plus Insulin	Insulin B Chain	c28	5+	1896.883404	1896.883827	0.22
hIAPP plus Insulin	Insulin B Chain	c29	6+	1602.087178	1602.087134	-0.03
hIAPP plus Insulin	Insulin B Chain	z21	4+	2178.522447	2178.523282	0.38
hIAPP plus Insulin	Insulin B Chain	z24	4+	2239.285863	2239.286889	0.46
hIAPP plus Insulin	Insulin B Chain	z24	5+	1792.029854	1792.029364	-0.27
hIAPP plus Insulin	Insulin B Chain	z25	4+	2268.308542	2268.308948	0.18
hIAPP plus Insulin	Insulin B Chain	z25	5+	1814.245858	1814.245772	-0.05
hIAPP plus Insulin	Insulin B Chain	z26	4+	2302.573277	2302.573352	0.03
hIAPP plus Insulin	Insulin B Chain	z26	5+	1842.258864	1842.258245	-0.34
hIAPP plus Insulin	Insulin B Chain	z27	4+	2335.088870	2335.090524	0.71
hIAPP plus Insulin	Insulin B Chain	z27	5+	1867.871757	1867.871917	0.09
hIAPP plus Insulin	hIAPP	c30	3+	2992.071476	2992.071146	-0.11
hIAPP plus Insulin	hIAPP	c35	5+	1889.887284	1889.887505	0.12
hIAPP plus Insulin	hIAPP	c35	6+	1575.240953	1575.240464	-0.31
hIAPP plus Insulin	hIAPP	c36	5+	1910.497623	1910.497296	-0.17
hIAPP plus Insulin	hIAPP	c36	6+	1591.581151	1591.581085	-0.04
hIAPP plus Insulin	hIAPP	z3	3+	2062.944772	2062.944905	0.06
hIAPP plus Insulin	hIAPP	z3	4+	1547.458964	1547.458220	-0.48
hIAPP plus Insulin	hIAPP	z5	3+	2110.962606	2110.963800	0.57
hIAPP plus Insulin	hIAPP	z5	4+	1583.722822	1583.722980	0.10
hIAPP plus Insulin	hIAPP	z7	3+	2182.667922	2182.668342	0.19
hIAPP plus Insulin	hIAPP	z7	4+	1637.251212	1637.252150	0.57
hIAPP plus Insulin	hIAPP	z9	4+	1683.520936	1683.521294	0.21
hIAPP plus Insulin	hIAPP	z10	3+	2274.037750	2274.037266	-0.21
hIAPP plus Insulin	hIAPP	z10	4+	1705.778658	1705.779637	0.57
hIAPP plus Insulin	hIAPP	z11	3+	2312.064759	2312.065994	0.53
hIAPP plus Insulin	hIAPP	z11	4+	1733.799170	1733.799257	0.05
hIAPP plus Insulin	hIAPP	z12	4+	1762.320723	1762.321406	0.39
hIAPP plus Insulin	hIAPP	z13	3+	2374.108297	2374.108691	0.17
hIAPP plus Insulin	hIAPP	z13	4+	1779.578921	1779.578564	-0.20
hIAPP plus Insulin	hIAPP	z14	3+	2392.781614	2392.782476	0.36
hIAPP plus Insulin	hIAPP	z14	4+	1793.834286	1793.834601	0.18
hIAPP plus Insulin	hIAPP	z15	4+	1831.102502	1831.102190	-0.17
hIAPP plus Insulin	hIAPP	z16	3+	2480.152623	2480.153239	0.25
hIAPP plus Insulin	hIAPP	z16	4+	1859.613225	1859.613843	0.33
hIAPP plus Insulin	hIAPP	z17	3+	2517.833081	2517.833039	-0.02
hIAPP plus Insulin	hIAPP	z17	4+	1888.624934	1888.625324	0.21
hIAPP plus Insulin	hIAPP	z18	3+	2547.177629	2547.178108	0.19
hIAPP plus Insulin	hIAPP	z18	4+	1909.881974	1909.882584	0.32
hIAPP plus Insulin	hIAPP	z19	3+	2575.520295	2575.521536	0.48
hIAPP plus Insulin	hIAPP	z19	4+	1932.391433	1932.391192	-0.12
hIAPP plus Insulin	hIAPP	z20	4+	1966.155222	1966.155322	0.05
hIAPP plus Insulin	hIAPP	z22	3+	2692.258282	2692.259678	0.52
hIAPP plus Insulin	hIAPP	z22	4+	2018.944275	2018.943853	-0.21
hIAPP plus Insulin	hIAPP	z23	3+	2741.615295	2741.615532	0.09
hIAPP plus Insulin	hIAPP	z23	4+	2055.460798	2055.461004	0.10
hIAPP plus Insulin	hIAPP	z23	5+	1644.769445	1644.769952	0.31
hIAPP plus Insulin	hIAPP	z24	4+	2084.723243	2084.723765	0.25
hIAPP plus Insulin	hIAPP	z24	5+	1667.578032	1667.578644	0.37
hIAPP plus Insulin	hIAPP	z26	4+	2130.752175	2130.752715	0.25
hIAPP plus Insulin	hIAPP	z26	5+	1704.402280	1704.402822	0.32
hIAPP plus Insulin	hIAPP	z27	3+	2893.705069	2893.704358	-0.25
hIAPP plus Insulin	hIAPP	z27	4+	2169.277708	2169.277893	0.09
hIAPP plus Insulin	hIAPP	z27	5+	1735.422057	1735.422001	-0.03
hIAPP plus Insulin	hIAPP	z28	4+	2201.542923	2201.543457	0.24
hIAPP plus Insulin	hIAPP	z28	5+	1761.033774	1761.033616	-0.09
hIAPP plus Insulin	hIAPP	z29	5+	1781.243313	1781.243788	0.27
hIAPP plus Insulin	hIAPP	z31	4+	2270.063030	2270.063299	0.12
hIAPP plus Insulin	hIAPP	z34	4+	2338.598121	2338.598584	0.20
hIAPP plus Insulin	hIAPP	z34	5+	1871.078772	1871.078728	-0.02

Average Error (ppm):		0.03
Average Absolute Error (ppm):		0.23
Standard Deviation:		0.29

Table S3.1. List of the assigned ECD fragments of the 7+ molecular ion of hIAPP monomer unit with an intact insulin molecule.

231

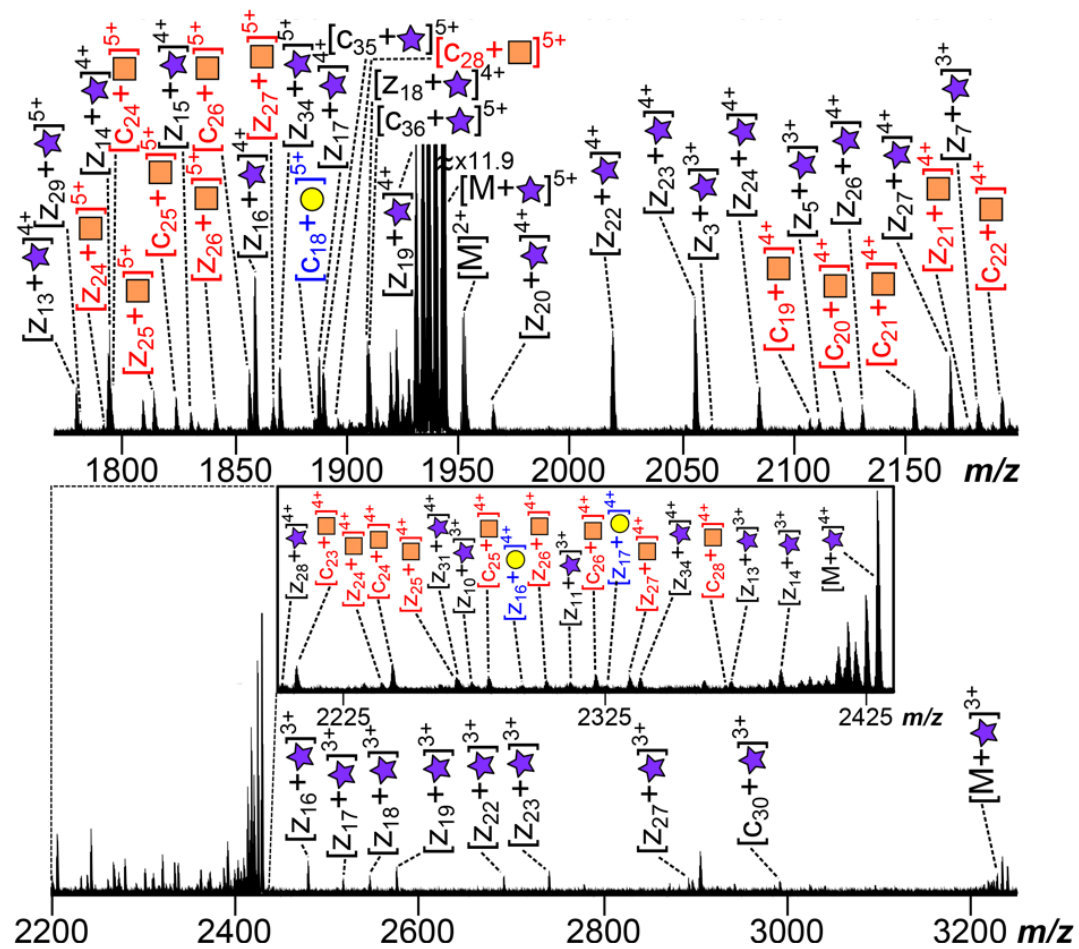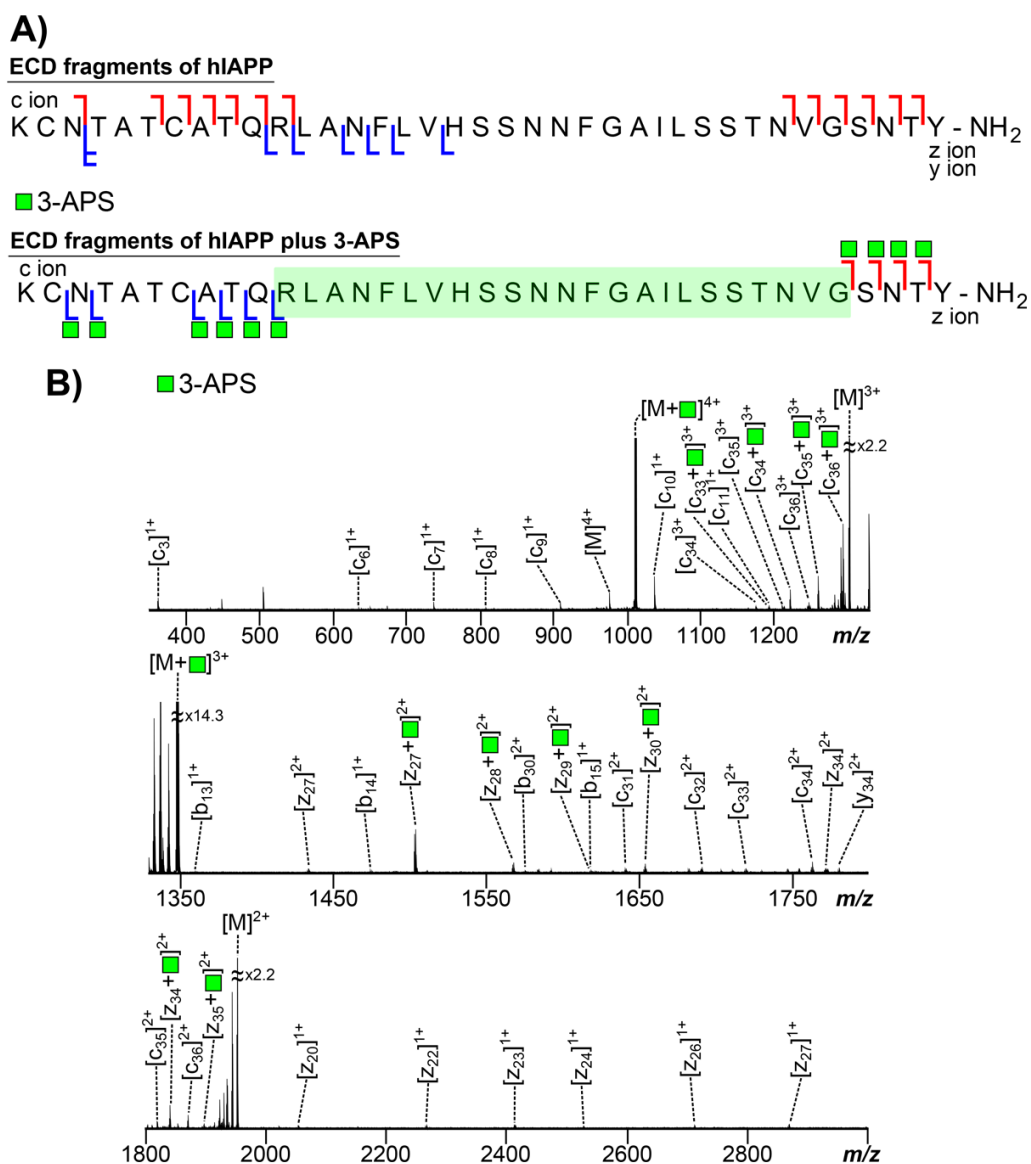**Figure S3.6.** (A) The summarised ECD fragments of the 4+ charge state molecular ion of hIAPP monomer unit and a 3-APS molecule. The highlighted area (coloured in green) indicates the proposed interaction region between the monomer unit of hIAPP and a 3-APS molecule. (B) The ECD MS/MS spectrum of the 4+ charge state molecular ion of hIAPP monomer unit and a 3-APS molecule. The assigned fragments are listed in the Supplementary Information Table S3.2.

Assigned ECD fragments of the 4+ molecular ion of hIAPP with 3-APS

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	2+	1952.948278	1952.948254	-0.01
hIAPP only	hIAPP	M	3+	1301.631069	1301.630261	-0.62
hIAPP only	hIAPP	M	4+	976.223165	976.222690	-0.49
hIAPP only	hIAPP	c3	1+	362.173076	362.173076	0.00
hIAPP only	hIAPP	c6	1+	635.305548	635.305566	0.03
hIAPP only	hIAPP	c7	1+	737.306908	737.306847	-0.08
hIAPP only	hIAPP	c8	1+	808.344022	808.344047	0.03
hIAPP only	hIAPP	c9	1+	909.391701	909.391718	0.02
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450437	0.15
hIAPP only	hIAPP	c11	1+	1193.551390	1193.552000	0.51
hIAPP only	hIAPP	c31	2+	1640.801138	1640.800801	-0.21
hIAPP only	hIAPP	c32	2+	1691.338138	1691.338261	0.07
hIAPP only	hIAPP	c33	2+	1719.349684	1719.349984	0.17
hIAPP only	hIAPP	c34	2+	1762.362091	1762.361757	-0.19
hIAPP only	hIAPP	c34	3+	1175.576214	1175.576225	0.01
hIAPP only	hIAPP	c35	2+	1819.383550	1819.382525	-0.56
hIAPP only	hIAPP	c35	3+	1213.258129	1213.258327	0.16
hIAPP only	hIAPP	c36	2+	1870.912691	1870.912382	-0.17
hIAPP only	hIAPP	c36	3+	1246.940688	1246.940086	-0.48
hIAPP only	hIAPP	y34	2+	1780.374752	1780.374624	-0.07
hIAPP only	hIAPP	z20	1+	2053.957373	2053.955469	-0.93
hIAPP only	hIAPP	z22	1+	2268.116400	2268.117725	0.58
hIAPP only	hIAPP	z23	1+	2414.181675	2414.183249	0.65
hIAPP only	hIAPP	z24	1+	2528.224602	2528.225646	0.41
hIAPP only	hIAPP	z26	1+	2713.351107	2713.348975	-0.79
hIAPP only	hIAPP	z27	1+	2867.443481	2867.444368	0.31
hIAPP only	hIAPP	z27	2+	1433.721470	1433.721878	0.28
hIAPP only	hIAPP	z34	2+	1771.360072	1771.358333	-0.98
hIAPP plus 3-APS	hIAPP plus 3-APS	M	3+	1347.974507	1347.973686	-0.61
hIAPP plus 3-APS	hIAPP plus 3-APS	M	4+	1010.980743	1010.979900	-0.83
hIAPP plus 3-APS	hIAPP	c33	3+	1192.910834	1192.910304	-0.44
hIAPP plus 3-APS	hIAPP	c34	3+	1221.585569	1221.585336	-0.19
hIAPP plus 3-APS	hIAPP	c35	3+	1259.265622	1259.265032	-0.47
hIAPP plus 3-APS	hIAPP	c36	3+	1293.618379	1293.618217	-0.13
hIAPP plus 3-APS	hIAPP	z27	2+	1503.236628	1503.236613	-0.01
hIAPP plus 3-APS	hIAPP	z28	2+	1567.265917	1567.266163	0.16
hIAPP plus 3-APS	hIAPP	z29	2+	1618.293664	1618.294162	0.31
hIAPP plus 3-APS	hIAPP	z30	2+	1653.308313	1653.307960	-0.21
hIAPP plus 3-APS	hIAPP	z34	2+	1840.371311	1840.370997	-0.17
hIAPP plus 3-APS	hIAPP	z35	2+	1898.398092	1898.396810	-0.68

Average Error (ppm):	-0.14
Average Absolute Error (ppm):	0.33
Standard Deviation:	0.41

Table S3.2.   List of the assigned ECD fragments of the 4+ molecular ion of hIAPP monomer unit with a 3-APS molecule.

Figure S3.7. (A) The summarised ECD fragments of the 4+ charge state molecular ion of hIAPP monomer unit and a BISA molecule. The highlighted area (coloured in pink) indicates the proposed interaction region between the monomer unit of hIAPP and the benzimidazole group from BISA molecule, the highlighted area (coloured in yellow) indicates the proposed interaction region between the monomer unit of hIAPP and sulfonic acid group from BISA molecule, and the highlighted area (coloured in cyan) indicates the proposed interaction region between the monomer unit of hIAPP and a BISA molecule. (B) The ECD MS/MS spectrum of the 4+ charge state molecular ion of the monomer unit of hIAPP and a BISA molecule. The assigned fragments are listed in the Supplementary Information Table S3.3.

Assigned ECD fragments of the 4+ molecular ion of hIAPP with BISA

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	1+	3904.889893	3904.887418	-0.63
hIAPP only	hIAPP	M	2+	1951.439053	1951.440443	0.71
hIAPP only	hIAPP	M	3+	1301.295128	1301.294226	-0.69
hIAPP only	hIAPP	c2	1+	249.137973	249.137977	0.02
hIAPP only	hIAPP	c3	1+	362.173076	362.173078	0.01
hIAPP only	hIAPP	c4	1+	464.228579	464.228503	-0.16
hIAPP only	hIAPP	c6	1+	635.305548	635.305584	0.06
hIAPP only	hIAPP	c7	1+	737.306908	737.306980	0.10
hIAPP only	hIAPP	c8	1+	808.344022	808.344148	0.16
hIAPP only	hIAPP	c9	1+	909.391701	909.391767	0.07
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450521	0.23
hIAPP only	hIAPP	c11	1+	1193.551390	1193.551031	-0.30
hIAPP only	hIAPP	c14	1+	1491.715496	1491.714904	-0.40
hIAPP only	hIAPP	c15	1+	1639.787321	1639.788269	0.58
hIAPP only	hIAPP	c15	2+	819.895593	819.895480	-0.14
hIAPP only	hIAPP	c21	2+	1138.554781	1138.554629	-0.13
hIAPP only	hIAPP	c24	2+	1297.117268	1297.117010	-0.20
hIAPP only	hIAPP	c28	2+	1490.742244	1490.742770	0.35
hIAPP only	hIAPP	c30	2+	1583.275757	1583.276918	0.73
hIAPP only	hIAPP	c31	2+	1640.801138	1640.801100	-0.02
hIAPP only	hIAPP	c32	2+	1690.335340	1690.334057	-0.76
hIAPP only	hIAPP	c33	2+	1719.349984	1719.349194	-0.46
hIAPP only	hIAPP	c34	2+	1754.350208	1754.350050	-0.09
hIAPP only	hIAPP	c34	2+	1762.863477	1762.863011	-0.26
hIAPP only	hIAPP	c34	3+	1175.241931	1175.241881	-0.04
hIAPP only	hIAPP	c35	2+	1820.887786	1820.888080	0.16
hIAPP only	hIAPP	c35	3+	1213.258129	1213.257571	-0.46
hIAPP only	hIAPP	c36	2+	1870.408779	1870.409875	0.59
hIAPP only	hIAPP	c36	3+	1246.940688	1246.941084	0.32
hIAPP only	hIAPP	y3	1+	396.187761	396.187747	-0.04
hIAPP only	hIAPP	y4	1+	483.219790	483.219762	-0.06
hIAPP only	hIAPP	y5	1+	540.241254	540.241213	-0.08
hIAPP only	hIAPP	y6	1+	639.309668	639.309730	0.10
hIAPP only	hIAPP	y7	1+	753.352596	753.352383	-0.28
hIAPP only	hIAPP	y8	1+	854.400275	854.400858	0.68
hIAPP only	hIAPP	y9	1+	941.432304	941.432826	0.55
hIAPP only	hIAPP	y20	1+	2068.968279	2068.968296	0.01
hIAPP only	hIAPP	y30	2+	1592.303929	1592.303293	-0.40
hIAPP only	hIAPP	y31	2+	1643.805834	1643.806789	0.58
hIAPP only	hIAPP	z3	1+	381.176861	381.176848	-0.03
hIAPP only	hIAPP	z4	1+	468.208889	468.208859	-0.06
hIAPP only	hIAPP	z7	1+	738.341694	738.341792	0.13
hIAPP only	hIAPP	z21	1+	2154.029199	2154.030537	0.62
hIAPP only	hIAPP	z22	1+	2266.109851	2266.110527	0.30
hIAPP only	hIAPP	z23	1+	2413.178813	2413.178808	0.00
hIAPP only	hIAPP	z24	1+	2528.224602	2528.224676	0.03
hIAPP only	hIAPP	z26	1+	2711.342370	2711.344627	0.83
hIAPP only	hIAPP	z27	1+	2868.446892	2868.447024	0.05
hIAPP only	hIAPP	z27	2+	1433.721470	1433.720079	-0.97
hIAPP only	hIAPP	z28	1+	2995.497644	2995.498257	0.20
hIAPP only	hIAPP	z29	1+	3097.553148	3097.555456	0.75
hIAPP only	hIAPP	z34	1+	3542.720681	3542.721346	0.19
hIAPP only	hIAPP	z34	2+	1771.360072	1771.361137	0.60
hIAPP only	hIAPP	z35	1+	3657.771433	3657.771686	0.07
hIAPP only	hIAPP	z35	2+	1828.381530	1828.381755	0.12
hIAPP only	hIAPP	z36	2+	1880.385028	1880.384550	-0.25
hIAPP plus benzi- group	hIAPP	M	3+	1340.310219	1340.310236	0.01
hIAPP plus benzi- group	hIAPP	c2	1+	365.175421	365.175412	-0.02
hIAPP plus benzi- group	hIAPP	c3	1+	479.218349	479.218359	0.02
hIAPP plus benzi- group	hIAPP	c31	2+	1699.825159	1699.824288	-0.51
hIAPP plus benzi- group	hIAPP	c32	2+	1749.359370	1749.359057	-0.18

(Table S3.3. Continue)

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP plus benzi- group	hIAPP	c33	2+	1777.368713	1777.368084	-0.35
hIAPP plus benzi- group	hIAPP	c35	2+	1877.906192	1877.906867	0.36
hIAPP plus benzi- group	hIAPP	c35	3+	1251.937275	1251.937532	0.21
hIAPP plus benzi- group	hIAPP	z3	1+	497.214310	497.214304	-0.01
hIAPP plus benzi- group	hIAPP	z4	1+	584.246339	584.246451	0.19
hIAPP plus benzi- group	hIAPP	z5	1+	641.267803	641.267988	0.29
hIAPP plus benzi- group	hIAPP	z6	1+	740.336217	740.336096	-0.16
hIAPP plus benzi- group	hIAPP	z7	1+	854.379145	854.379188	0.05
hIAPP plus benzi- group	hIAPP	z8	1+	955.426824	955.426851	0.03
hIAPP plus benzi- group	hIAPP	z13	1+	1426.696124	1426.696751	0.44
hIAPP plus benzi- group	hIAPP	z14	1+	1483.718588	1483.718036	-0.37
hIAPP plus benzi- group	hIAPP	z22	2+	1191.577291	1191.577204	-0.07
hIAPP plus benzi- group	hIAPP	z23	2+	1265.111498	1265.111194	-0.24
hIAPP plus benzi- group	hIAPP	z27	2+	1492.244107	1492.244991	0.59
hIAPP plus benzi- group	hIAPP	z28	2+	1556.273396	1556.272245	-0.74
hIAPP plus Sulf- group	hIAPP	M	3+	1328.283341	1328.283239	-0.08
hIAPP plus Sulf- group	hIAPP	c14	1+	1572.680137	1572.680764	0.40
hIAPP plus Sulf- group	hIAPP	c15	1+	1719.748551	1719.749599	0.61
hIAPP plus Sulf- group	hIAPP	c21	2+	1179.037098	1179.037729	0.54
hIAPP plus Sulf- group	hIAPP	c23	2+	1309.592773	1309.591806	-0.74
hIAPP plus Sulf- group	hIAPP	c24	2+	1338.103505	1338.104196	0.52
hIAPP plus Sulf- group	hIAPP	c25	2+	1373.622062	1373.621721	-0.25
hIAPP plus Sulf- group	hIAPP	c26	2+	1430.665464	1430.666689	0.86
hIAPP plus Sulf- group	hIAPP	c27	2+	1486.706126	1486.705662	-0.31
hIAPP plus Sulf- group	hIAPP	c28	2+	1530.222140	1530.223458	0.86
hIAPP plus Sulf- group	hIAPP	c29	2+	1573.738150	1573.736931	-0.77
hIAPP plus Sulf- group	hIAPP	c30	2+	1624.261994	1624.262247	0.16
hIAPP plus Sulf- group	hIAPP	c31	2+	1681.283453	1681.282477	-0.58
hIAPP plus Sulf- group	hIAPP	c32	2+	1759.829776	1759.829492	-0.16
hIAPP plus Sulf- group	hIAPP	c34	2+	1802.844406	1802.843646	-0.42
hIAPP plus Sulf- group	hIAPP	z27	1+	2949.406903	2949.406429	-0.16
hIAPP plus BISA	hIAPP	M	2+	2051.953229	2051.954220	0.48
hIAPP plus BISA	hIAPP	M	3+	1367.634374	1367.633946	-0.31
hIAPP plus BISA	hIAPP	M	4+	1025.725643	1025.725077	-0.55
hIAPP plus BISA	hIAPP	c2	1+	446.140062	446.140048	-0.03
hIAPP plus BISA	hIAPP	c3	1+	560.182990	560.183063	0.13
hIAPP plus BISA	hIAPP	c14	1+	1689.725410	1689.726446	0.61
hIAPP plus BISA	hIAPP	c24	2+	1396.626141	1396.625443	-0.50
hIAPP plus BISA	hIAPP	c28	2+	1588.744777	1588.744252	-0.33
hIAPP plus BISA	hIAPP	c29	2+	1632.260791	1632.259716	-0.66
hIAPP plus BISA	hIAPP	c30	2+	1682.280713	1682.281258	0.32
hIAPP plus BISA	hIAPP	c31	2+	1739.806090	1739.805997	-0.05
hIAPP plus BISA	hIAPP	c32	2+	1789.340302	1789.339691	-0.34
hIAPP plus BISA	hIAPP	c33	2+	1818.352414	1818.352893	0.26
hIAPP plus BISA	hIAPP	c34	2+	1860.863130	1860.863421	0.16
hIAPP plus BISA	hIAPP	c34	3+	1240.911179	1240.911175	0.00
hIAPP plus BISA	hIAPP	c35	2+	1268.936489	1268.936668	0.14
hIAPP plus BISA	hIAPP	c35	3+	1278.925488	1278.925303	-0.14
hIAPP plus BISA	hIAPP	c36	3+	1312.943993	1312.943927	-0.05
hIAPP plus BISA	hIAPP	z9	1+	1042.458853	1042.459888	0.99
hIAPP plus BISA	hIAPP	z27	1+	3066.456804	3066.457839	0.34
hIAPP plus BISA	hIAPP	z27	2+	1532.726427	1532.726848	0.27
hIAPP plus BISA	hIAPP	z28	2+	1596.755716	1596.756728	0.63
hIAPP plus BISA	hIAPP	z29	2+	1647.780966	1647.779427	-0.93
hIAPP plus BISA	hIAPP	z34	2+	1870.365029	1870.365191	0.09

Average Error (ppm):	0.03
Average Absolute Error (ppm):	0.32
Standard Deviation:	0.42

Table S3.3. List of the assigned ECD fragments of the 4+ molecular ion of hIAPP monomer unit with a BISA molecule.
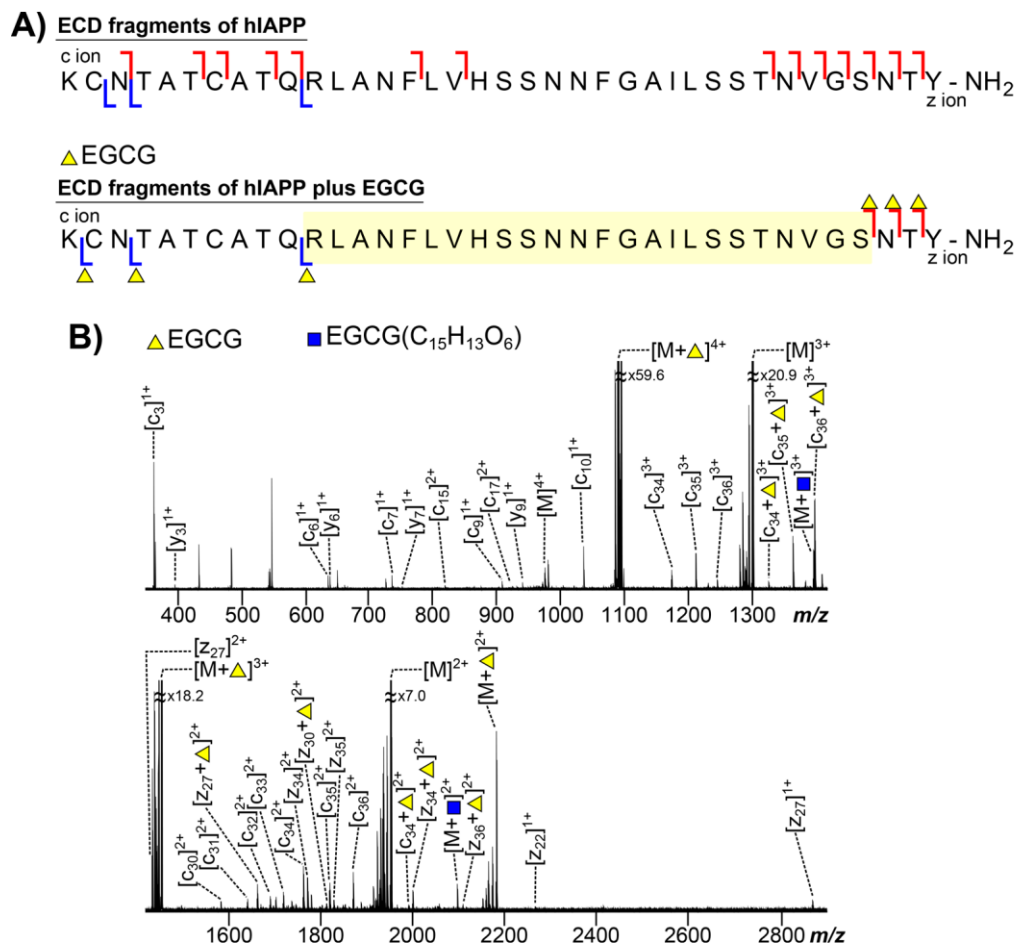
Figure S3.8. (A) The summarised ECD fragments of the 4+ charge state molecular ion of hIAPP monomer unit and an EGCG molecule. The highlighted area (coloured in yellow) indicates the proposed interaction region between the monomer unit of hIAPP and an EGCG molecule. (B) The ECD MS/MS spectrum of the 4+ charge state molecular ion of hIAPP monomer unit and an EGCG molecule. The assigned fragments are listed in the Supplementary Information Table S3.4.
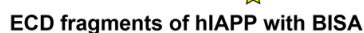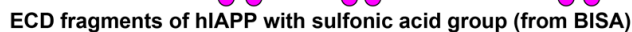
Assigned ECD fragments of the 4+ molecular ion of hIAPP with EGCG

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	2+	1952.446878	1952.445693	0.61
hIAPP only	hIAPP	M	3+	1301.965336	1301.965433	-0.07
hIAPP only	hIAPP	M	4+	976.223165	976.223115	0.05
hIAPP only	hIAPP	c3	1+	363.180901	363.180903	-0.01
hIAPP only	hIAPP	c6	1+	636.313372	636.313307	0.10
hIAPP only	hIAPP	c7	1+	737.306908	737.306793	0.16
hIAPP only	hIAPP	c9	1+	909.391701	909.392330	-0.69
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450491	-0.20
hIAPP only	hIAPP	c15	2+	819.895593	819.895681	-0.11
hIAPP only	hIAPP	c17	2+	925.971832	925.972160	-0.35
hIAPP only	hIAPP	c30	2+	1584.281062	1584.280627	0.27
hIAPP only	hIAPP	c31	2+	1641.302522	1641.302312	0.13
hIAPP only	hIAPP	c32	2+	1690.335345	1690.334878	0.28
hIAPP only	hIAPP	c33	2+	1719.347463	1719.347399	0.04
hIAPP only	hIAPP	c34	2+	1762.362086	1762.362107	-0.01
hIAPP only	hIAPP	c34	3+	1175.912156	1175.912079	0.07
hIAPP only	hIAPP	c35	2+	1820.388715	1820.388850	-0.07
hIAPP only	hIAPP	c35	3+	1213.592384	1213.592237	0.12
hIAPP only	hIAPP	c36	2+	1870.411302	1870.411105	0.11
hIAPP only	hIAPP	c36	3+	1246.940685	1246.940038	0.52
hIAPP only	hIAPP	y3	1+	396.187761	396.187732	0.07
hIAPP only	hIAPP	y6	1+	639.309668	639.309634	0.05
hIAPP only	hIAPP	y7	1+	753.352596	753.353056	-0.61
hIAPP only	hIAPP	y9	1+	941.432304	941.432732	-0.45
hIAPP only	hIAPP	z22	1+	2266.104890	2266.104193	0.31
hIAPP only	hIAPP	z27	1+	2868.446331	2868.445815	0.18
hIAPP only	hIAPP	z27	2+	1434.726804	1434.726964	-0.11
hIAPP only	hIAPP	z34	2+	1771.360072	1771.360101	-0.02
hIAPP only	hIAPP	z35	2+	1828.381536	1828.381635	-0.05
hIAPP plus EGCG ($C_{15}H_{13}O_6$)	hIAPP	M	2+	2096.979994	2096.979343	0.31
hIAPP plus EGCG ($C_{15}H_{13}O_6$)	hIAPP	M	3+	1397.986480	1397.985924	0.40
hIAPP plus EGCG	hIAPP	M	2+	2181.990765	2181.990604	0.07
hIAPP plus EGCG	hIAPP	M	3+	1454.994485	1454.994299	0.13
hIAPP plus EGCG	hIAPP	M	4+	1090.995108	1090.994632	0.44
hIAPP plus EGCG	hIAPP	c34	2+	1991.404542	1991.405413	-0.44
hIAPP plus EGCG	hIAPP	c34	3+	1327.937130	1327.936824	0.23
hIAPP plus EGCG	hIAPP	c35	3+	1366.285572	1366.285710	-0.10
hIAPP plus EGCG	hIAPP	c36	3+	1399.969938	1399.969107	0.59
hIAPP plus EGCG	hIAPP	z27	2+	1663.265384	1663.264553	0.50
hIAPP plus EGCG	hIAPP	z30	2+	1813.337064	1813.337861	-0.44
hIAPP plus EGCG	hIAPP	z34	2+	2001.405284	2001.404286	0.50
hIAPP plus EGCG	hIAPP	z36	2+	2110.432358	2110.432229	0.06

Average Error (ppm): 0.06
Average Absolute Error (ppm): 0.24
Standard Deviation: 0.31

Table S3.4. List of the assigned ECD fragments of the 4+ molecular ion of hIAPP monomer unit with an EGCG molecule.
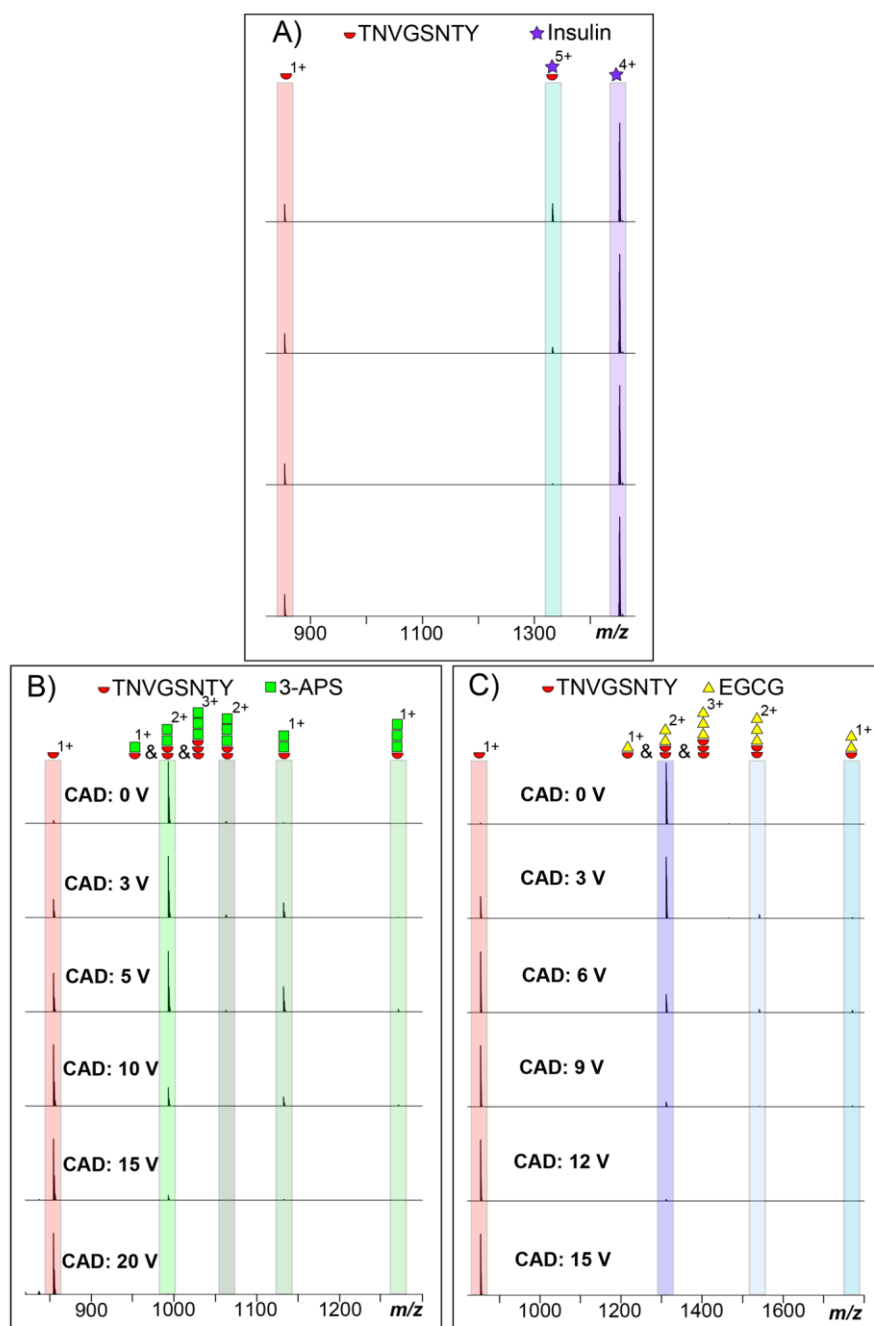
Figure S3.9. The representative CAD spectra of (A) the 5+ charge state molecular ion of ^{30}TNVGSNTY^{37}-NH$_2$ peptide mixed with an intact insulin molecule in equal molar, and the molecular ion of ^{30}TNVGSNTY^{37}-NH$_2$ peptide mixed with equal molar of (B) 3-APS and (C) EGCG molecules.
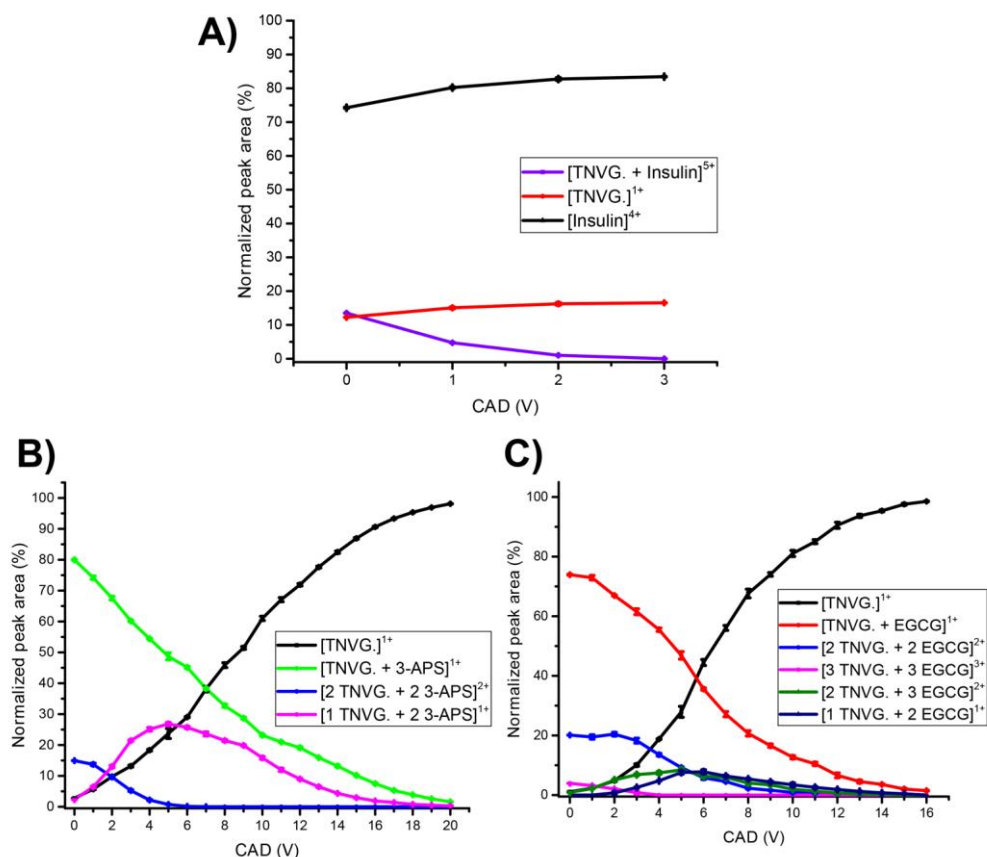
Figure S3.10. The dissociation curves of (A) the 5+ ^{30}TNVGSNTY^{37}-NH$_2$ peptide mixed with equal molar of insulin molecule (purple), and the ^{30}TNVGSNTY^{37}-NH$_2$ peptide mixed with equal molar of (B) 3-APS (green) and (C) EGCG (red) molecules which were generated from the Supplementary Information Figure S3.9. The dissociation curve of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with insulin is lower than the dissociation curves of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with 3-APS and EGCG as different charge states of molecular ions were isolated and fragmented.[73] Figure S3.10A clearly shows that the complex of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with insulin simply dissociates into the monomer unit of ^{30}TNVGSNTY^{37}-NH$_2$ peptide and insulin molecule. In Figure S3.10B, 1:1 and 2:2 complexes of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with 3-APS molecule(s) were co-isolated (different charge state but same *m/z* ratio). Dissociation of the 2:2 complex of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with 3-APS results in an increase of

1:2 complex ion of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with 3-APS which indicates the binding between the 2:2 complex is either two 3-APS molecules are attaching to one ^{30}TNVGSNTY^{37}-NH$_2$ peptide or the ^{30}TNVGSNTY^{37}-NH$_2$ peptide and 3-APS molecule are in sandwich pattern that a 3-APS molecule is placed in the middle of two ^{30}TNVGSNTY^{37}-NH$_2$ peptides.   The relative percentage of the 1:1 complex of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with 3-APS decreases while the percentage of the ^{30}TNVGSNTY^{37}-NH$_2$ peptide increases proportionally to the CAD energy applied. In Figure S3.10C, the 1:1, 2:2, and 3:3 complexes of ^{30}TNVGSNTY^{37}-NH$_2$ peptide with EGCG molecule(s) were co-isolated; and similar dissociation patterns which is described above are also observed in the CAD curve of EGCG mixed with ^{30}TNVGSNTY^{37}-NH$_2$ peptide.

A)

ECD fragments of Insulin Chain B

c ion

F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T

z ion

y ion

ECD fragments of TNVGSNTY-NH$_2$

c ion

T N V G S N T Y - NH$_2$

ECD fragments of TNVGSNTY-NH$_2$ plus insulin

⭐ Insulin (Chain A&B)

🟩 Insulin Chain A & TNVGSNTY-NH$_2$ 🟣 Insulin Chain B & TNVGSNTY-NH$_2$

c ion

G I V E Q C C T S I C S L Y Q L E N Y C N

z ion

c ion

F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T

z ion

c ion

T N V G S N T Y - NH$_2$

z ion

(Figure S3.11. Continue)

Figure S3.11. (A) The summarised ECD fragments of the 5+ charge state molecular ion of the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide with an intact insulin molecule. The fragments here show there is an interaction between the synthetic peptide and the insulin molecule; however, the interaction site between ^{30}TNVGSNTY^{37}-NH$_2$ peptide and insulin cannot be further narrowed down using ECD MS/MS fragmentation. (B) The ECD MS/MS spectrum of the 5+ charge state molecular ion of the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide with an intact insulin molecule. The assigned fragments are listed in the Supplementary Information Table S3.5.

Assigned ECD fragments of the 5+ molecular ion of TNVG. with insulin

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
TNVG. only	TNVG.	M	1+	854.400272	854.400154	-0.14
TNVG. only	TNVG.	c5	1+	476.246339	476.246346	0.01
TNVG. only	TNVG.	c6	1+	590.289267	590.289171	-0.16
TNVG. only	TNVG.	c7	1+	691.336946	691.336894	-0.08
Insulin only	Insulin	M	2+	2906.339245	2906.338207	-0.36
Insulin only	Insulin	M	3+	1938.560905	1938.560423	-0.25
Insulin only	Insulin	M	4+	1452.417985	1452.417681	-0.21
Insulin B only	Insulin B Chain	c4	1+	506.272159	506.272165	0.01
Insulin B only	Insulin B Chain	c5	1+	643.331071	643.331019	-0.08
Insulin B only	Insulin B Chain	c6	1+	756.415135	756.415114	-0.03
Insulin B only	Insulin B Chain	y5	1+	609.324255	609.324235	-0.03
Insulin B only	Insulin B Chain	y7	1+	903.461083	903.461131	0.05
Insulin B only	Insulin B Chain	y8	1+	960.482547	960.482192	-0.37
Insulin B only	Insulin B Chain	z4	1+	430.242202	430.242208	0.01
Insulin B only	Insulin B Chain	z5	1+	593.305531	593.305463	-0.11
Insulin B only	Insulin B Chain	z6	1+	740.373945	740.373861	-0.11
Insulin B only	Insulin B Chain	z7	1+	887.442359	887.442415	0.06
Insulin B only	Insulin B Chain	z8	1+	944.463823	944.463997	0.18
Insulin A & B Chain	Insulin B Chain	c29	3+	1902.210980	1902.210122	-0.45
TNVG. Plus Insulin	TNVG. Plus Insulin	M	2+	3333.039039	3333.038156	-0.26
TNVG. Plus Insulin	TNVG. Plus Insulin	M	3+	2221.357803	2221.357271	-0.24
TNVG. Plus Insulin	TNVG. Plus Insulin	M	4+	1665.515653	1665.515932	0.17
TNVG. Plus Insulin	TNVG. Plus Insulin	M	5+	1333.215286	1333.214946	-0.26
TNVG. Plus Insulin	Insulin A Chain	c20	4+	1638.512256	1638.512536	0.17
TNVG. Plus Insulin	Insulin A Chain	z20	4+	1647.255608	1647.255558	-0.03
TNVG. Plus Insulin	Insulin B Chain	c22	2+	2859.300768	2859.300628	-0.05
TNVG. Plus Insulin	Insulin B Chain	c22	3+	1908.205678	1908.205747	0.04
TNVG. Plus Insulin	Insulin B Chain	c24	2+	2962.848431	2962.849484	0.36
TNVG. Plus Insulin	Insulin B Chain	c24	3+	1976.569637	1976.569694	0.03
TNVG. Plus Insulin	Insulin B Chain	c25	2+	3036.883545	3036.882738	-0.27
TNVG. Plus Insulin	Insulin B Chain	c25	3+	2024.924789	2024.924427	-0.18
TNVG. Plus Insulin	Insulin B Chain	c26	2+	3117.413630	3117.413523	-0.03
TNVG. Plus Insulin	Insulin B Chain	c26	3+	2077.609408	2077.608310	-0.53
TNVG. Plus Insulin	Insulin B Chain	c29	4+	1640.507393	1640.507780	0.24
TNVG. Plus Insulin	Insulin B Chain	z24	2+	2954.831001	2954.831406	0.14
TNVG. Plus Insulin	Insulin B Chain	z24	3+	1969.889155	1969.889898	0.38
TNVG. Plus Insulin	Insulin B Chain	z25	2+	3012.875532	3012.875432	-0.03
TNVG. Plus Insulin	Insulin B Chain	z25	3+	2006.579793	2006.579614	-0.09
TNVG. Plus Insulin	Insulin B Chain	z26	3+	2053.602094	2053.602808	0.35
TNVG. Plus Insulin	Insulin B Chain	z27	3+	2096.288310	2096.287101	-0.58
TNVG. Plus Insulin	Insulin B Chain	z28	4+	1601.227667	1601.227547	-0.07
TNVG. Plus Insulin	TNVG.	c5	3+	2094.968618	2094.968272	-0.17
TNVG. Plus Insulin	TNVG.	c6	3+	2133.316900	2133.316551	-0.16
TNVG. Plus Insulin	TNVG.	c6	4+	1599.486538	1599.486589	0.03
TNVG. Plus Insulin	TNVG.	c7	3+	2166.999476	2166.999006	-0.22
TNVG. Plus Insulin	TNVG.	c7	4+	1625.249470	1625.250468	0.61
TNVG. Plus Insulin	TNVG.	z3	3+	2062.942860	2062.941486	-0.67
TNVG. Plus Insulin	TNVG.	z4	3+	2092.957368	2092.956062	-0.62
TNVG. Plus Insulin	TNVG.	z5	3+	2112.298431	2112.298119	-0.15
TNVG. Plus Insulin	TNVG.	z6	3+	2143.649478	2143.648099	-0.64
TNVG. Plus Insulin	TNVG.	z6	4+	1608.741898	1608.741346	-0.34
TNVG. Plus Insulin	TNVG.	z7	3+	2182.333775	2182.333009	-0.35
TNVG. Plus Insulin	TNVG.	z7	4+	1636.249053	1636.248028	-0.63

Average Error (ppm):	-0.12
Average Absolute Error (ppm):	0.22
Standard Deviation:	0.27

Table S3.5. List of the assigned ECD fragments of the 5+ molecular ion of the synthetic peptide ^{30}TNVGSNTY^{37}-NH$_2$ with an insulin molecule.
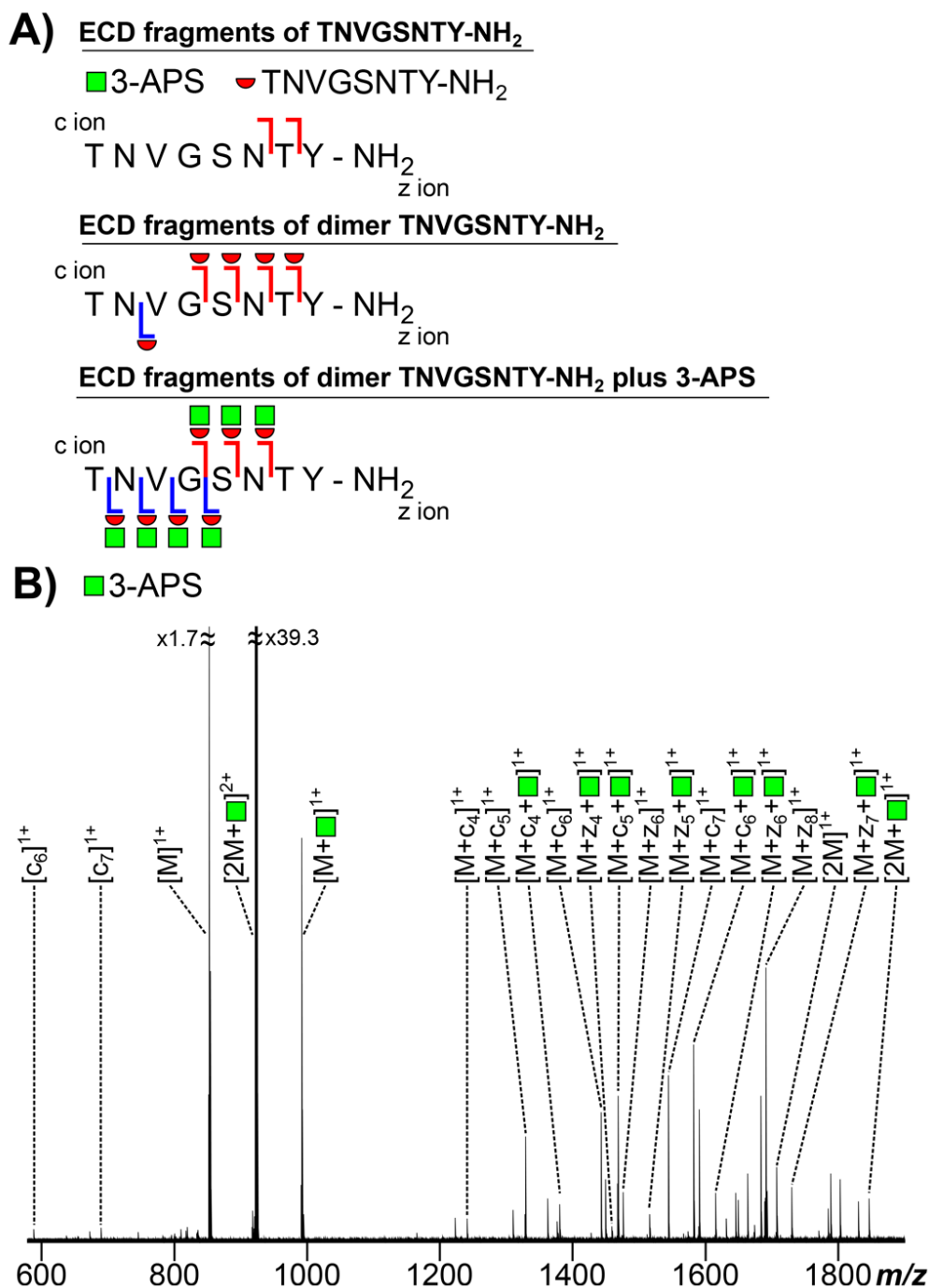
Figure S3.12. (A) The summarised ECD fragments of the 2+ charge state molecular ion of the synthetic dimer [30]TNVGSNTY[37]-NH$_2$ peptide with a 3-APS molecule. The fragments here show there is an interaction between the synthetic peptide and 3-APS molecule; however, the interaction site between [30]TNVGSNTY[37]-NH$_2$ peptide and 3-APS molecule cannot be further narrowed down using ECD MS/MS fragmentation. (B) The ECD MS/MS spectrum of the

2+ charge state molecular ion o the synthetic ^{30}TNVGSNTY^{37}-NH$_2$ peptide with a

3-APS molecule.   The assigned fragments are listed in the Supplementary

Information Table S3.6.

Assigned ECD fragments of the 2+ molecular ion of TNVG. with 3-APS

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
TNVG. only	TNVG.	M	1+	854.400272	854.400415	0.17
TNVG. only	TNVG.	c6	1+	590.289267	590.289268	0.00
TNVG. only	TNVG.	c7	1+	691.336946	691.336956	0.01
TNVG. Plus 3-APS	TNVG.	M	1+	1707.793267	1707.793256	-0.01
2 TNVG.	TNVG.	M	1+	1707.793267	1707.793256	-0.01
2 TNVG.	TNVG.	c4	1+	1241.599480	1241.599338	-0.11
2 TNVG.	TNVG.	c5	1+	1329.639338	1329.639890	0.42
2 TNVG.	TNVG.	c6	1+	1444.685021	1444.685323	0.21
2 TNVG.	TNVG.	c7	1+	1544.729945	1544.730497	0.36
2 TNVG.	TNVG.	z6	1+	1475.676112	1475.676280	0.11
2 TNVG. Plus 3-APS	TNVG.	M	1+	1846.823582	1846.823225	-0.19
2 TNVG. Plus 3-APS	TNVG.	M	2+	923.915429	923.915140	-0.31
2 TNVG. Plus 3-APS	TNVG.	c4	1+	1380.629794	1380.629143	-0.47
2 TNVG. Plus 3-APS	TNVG.	c5	1+	1468.669653	1468.670248	0.41
2 TNVG. Plus 3-APS	TNVG.	c6	1+	1582.712575	1582.713290	0.45
2 TNVG. Plus 3-APS	TNVG.	z4	1+	1459.624380	1459.623464	-0.63
2 TNVG. Plus 3-APS	TNVG.	z5	1+	1516.645844	1516.646234	0.26
2 TNVG. Plus 3-APS	TNVG.	z6	1+	1615.714258	1615.715388	0.70
2 TNVG. Plus 3-APS	TNVG.	z7	1+	1728.749354	1728.751009	0.96

Average Error (ppm): 0.12

Average Absolute Error (ppm): 0.30

Standard Deviation: 0.39

Table S3.6. List of the assigned ECD fragments of the 2+ molecular ion of the synthetic peptide ^{30}TNVGSNTY^{37}-NH$_2$ with a 3-APS molecule.
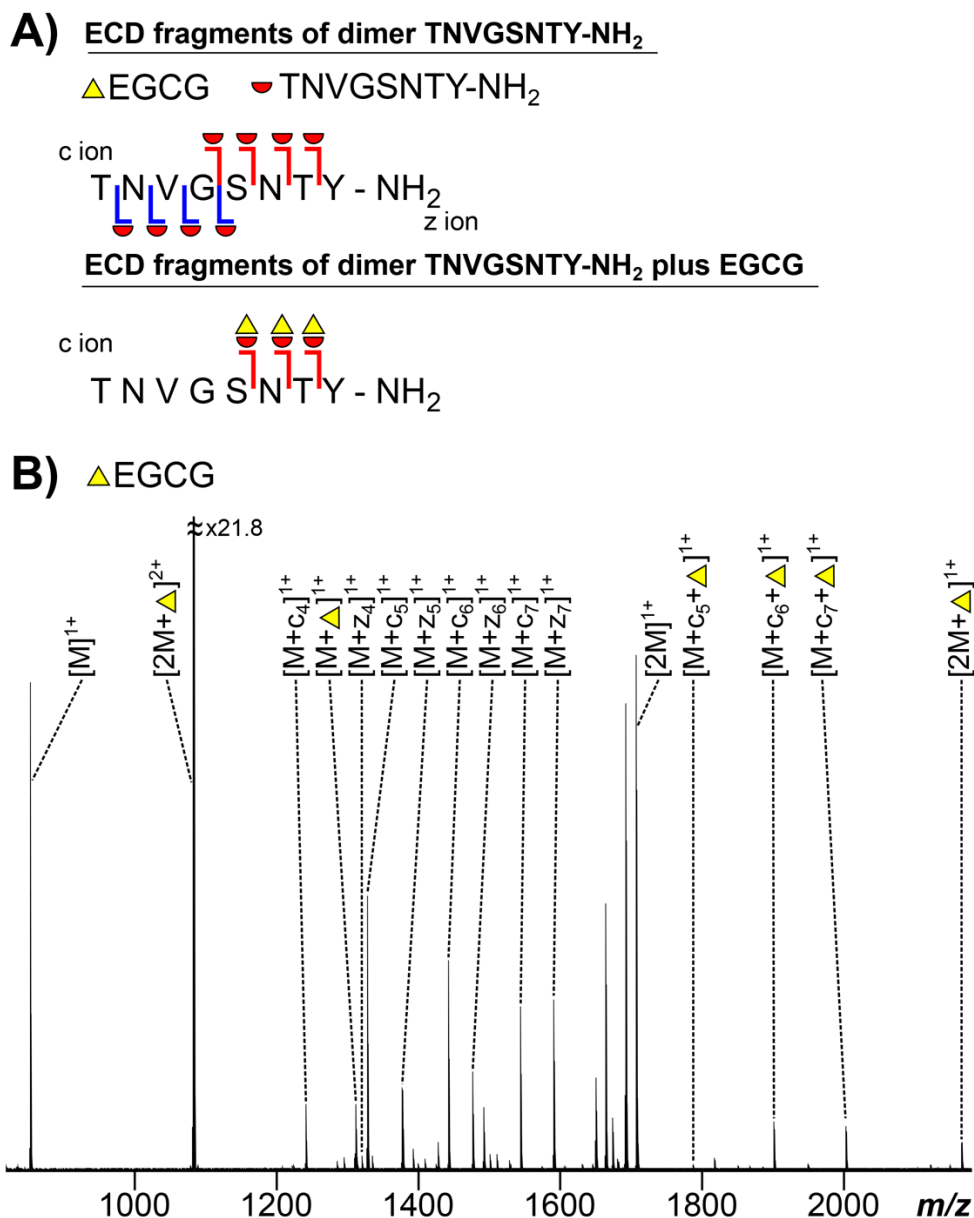
Figure S3.13. (A) The summarised ECD fragments of the 2+ charge state molecular ion of the synthetic dimer [30]TNVGSNTY[37]-NH$_2$ peptide with an EGCG molecule. The fragments here show there is an interaction between the synthetic peptide and EGCG molecule; however, the interaction site between [30]TNVGSNTY[37]-NH$_2$ peptide and EGCG molecule cannot be further narrowed down using ECD MS/MS fragmentation. (B) The ECD MS/MS spectrum of the 2+ charge state molecular ion o the synthetic [30]TNVGSNTY[37]-NH$_2$ peptide with

an EGCG molecule.   The assigned fragments are listed in the Supplementary

Information Table S3.7.

Assigned ECD fragments of the 2+ molecular ion of TNVG. with EGCG

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
TNVG. only	TNVG.	M	1+	854.400272	854.400272	0.00
TNVG. Plus EGCG	TNVG. Plus EGCG	M	1+	1312.485183	1312.486105	0.70
2 TNVG. Only	TNVG.	M	1+	1707.793267	1707.793099	-0.10
2 TNVG. Only	TNVG.	c4	1+	1242.607309	1242.607718	0.33
2 TNVG. Only	TNVG.	c5	1+	1329.639338	1329.639645	0.23
2 TNVG. Only	TNVG.	c6	1+	1443.682266	1443.682465	0.14
2 TNVG. Only	TNVG.	c7	1+	1544.729945	1544.729946	0.00
2 TNVG. Only	TNVG.	z4	1+	1321.601885	1321.601983	0.07
2 TNVG. Only	TNVG.	z5	1+	1378.623349	1378.623616	0.19
2 TNVG. Only	TNVG.	z6	1+	1477.691762	1477.692009	0.17
2 TNVG. Only	TNVG.	z7	1+	1591.734690	1591.734689	0.00
2 TNVG. Plus EGCG	2 TNVG. Plus EGCG	M	2+	1083.442728	1083.442844	0.11
2 TNVG. Plus EGCG	2 TNVG. Plus EGCG	M	1+	2166.881125	2166.880629	-0.23
2 TNVG. Plus EGCG	2 TNVG. Plus EGCG	c5	1+	1787.724253	1787.724571	0.18
2 TNVG. Plus EGCG	2 TNVG. Plus EGCG	c6	1+	1901.767181	1901.766758	-0.22
2 TNVG. Plus EGCG	2 TNVG. Plus EGCG	c7	1+	2002.814860	2002.814574	-0.14

Average Error (ppm):	0.09
Average Absolute Error (ppm):	0.18
Standard Deviation:	0.23

Table S3.7. List of the assigned ECD fragments of the 2+ molecular ion of the synthetic peptide ^{30}TNVGSNTY^{37}-NH$_2$ peptide with an EGCG molecule.

Figure S3.14. (A) The summarised ECD fragments of the 7+ charge state molecular ion of the wild-type hIAPP dimer with a 3-APS molecule. The highlight area (coloured in green) indicates the proposed non-interruption binding region, which does not interfere with the formation of the early oligomers between the wild-type hIAPP and a 3-APS molecule. (B) The ECD MS/MS spectrum of

the 7+ charge state molecular ion of the wild-type hIAPP dimer with a 3-APS molecule. The assigned fragments are listed in the Supplementary Information Table S3.8.

Assigned ECD fragments of the 7+ molecular ion of the dimer hIAPP with 3-APS

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	2+	1952.948278	1952.948270	0.00
hIAPP only	hIAPP	M	3+	1301.295128	1301.295662	0.41
hIAPP only	hIAPP	M	4+	976.473865	976.473132	-0.75
hIAPP only	hIAPP	c3	1+	362.173076	362.173080	0.01
hIAPP only	hIAPP	c7	1+	737.306908	737.306275	-0.86
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450072	-0.20
hIAPP only	hIAPP	c11	2+	597.279333	597.279307	-0.04
hIAPP only	hIAPP	c13	2+	689.339922	689.340053	0.19
hIAPP only	hIAPP	c14	2+	746.361386	746.361265	-0.16
hIAPP only	hIAPP	c15	2+	819.895593	819.895381	-0.26
hIAPP only	hIAPP	c18	2+	994.501288	994.500949	-0.34
hIAPP only	hIAPP	c19	2+	1038.518290	1038.518376	0.08
hIAPP only	hIAPP	c20	2+	1081.533317	1081.533686	0.34
hIAPP only	hIAPP	c21	2+	1139.056142	1139.056711	0.50
hIAPP only	hIAPP	c23	2+	1269.611825	1269.611172	-0.51
hIAPP only	hIAPP	c23	3+	846.743643	846.743637	-0.01
hIAPP only	hIAPP	c24	3+	865.416548	865.416448	-0.12
hIAPP only	hIAPP	c30	3+	1056.521724	1056.523134	1.33
hIAPP only	hIAPP	z2	1+	266.126109	266.126108	0.00
hIAPP plus 3-APS	hIAPP plus 3-APS	M	2+	2021.957014	2021.957522	0.25
hIAPP plus 3-APS	hIAPP plus 3-APS	M	3+	1348.306885	1348.306523	-0.27
hIAPP plus 3-APS	hIAPP	c14	2+	815.876544	815.876538	-0.01
hIAPP plus 3-APS	hIAPP	c15	2+	889.410751	889.411122	0.42
hIAPP plus 3-APS	hIAPP	c23	3+	892.752832	892.752095	-0.83
2 hIAPP	2 hIAPP	M	5+	1562.557828	1562.557362	-0.30
2 hIAPP	hIAPP	z5	3+	1476.704568	1476.703025	-1.04
2 hIAPP	hIAPP	z7	3+	1547.071595	1547.073041	0.93
2 hIAPP	hIAPP	z13	3+	1737.843926	1737.842828	-0.63
2 hIAPP	hIAPP	z14	3+	1757.187021	1757.187819	0.45
2 hIAPP	hIAPP	z16	3+	1844.558328	1844.559177	0.46
2 hIAPP	hIAPP	z16	4+	1383.418608	1383.419145	0.39
2 hIAPP	hIAPP	z22	4+	1543.249423	1543.248200	-0.79
2 hIAPP	hIAPP	z23	3+	2106.688983	2106.689746	0.36
2 hIAPP	hIAPP	z23	4+	1580.016547	1580.016915	0.23
2 hIAPP	hIAPP	z27	4+	1693.584213	1693.585304	0.64
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	M	5+	1589.962951	1589.963769	0.51
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	M	6+	1326.472079	1326.470560	-1.15
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	M	7+	1137.262236	1137.260919	-1.16
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	c32	5+	1485.117212	1485.118656	0.97
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	c33	5+	1496.723071	1496.722533	-0.36
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	c35	5+	1536.938066	1536.938162	0.06
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z7	3+	1593.080763	1593.082010	0.78
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z14	5+	1803.528660	1803.527909	-0.42
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z16	4+	1418.677268	1418.678535	0.89
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z17	4+	1446.937465	1446.936743	-0.50
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z18	4+	1468.444902	1468.444863	-0.03
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z19	4+	1489.952287	1489.953655	0.92
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z23	4+	1614.523477	1614.524319	0.52
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z24	4+	1643.284778	1643.284251	-0.32
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z26	5+	1351.453070	1351.453540	0.35
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z27	5+	1382.470406	1382.471723	0.95
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z30	5+	1442.902432	1442.902515	0.06
2 hIAPP plus 3-APS	2 hIAPP plus 3-APS	z34	5+	1517.528238	1517.527066	-0.77

Average Error (ppm):	0.02
Average Absolute Error (ppm):	0.47
Standard Deviation:	0.59

Table S3.8. List of the assigned ECD fragments of the 7+ molecular ion of the wild-type hIAPP dimer with a 3-APS molecule.

A) ECD fragments of hIAPP
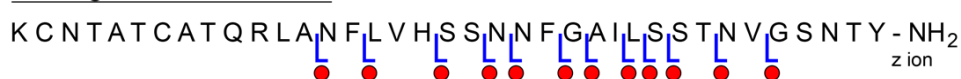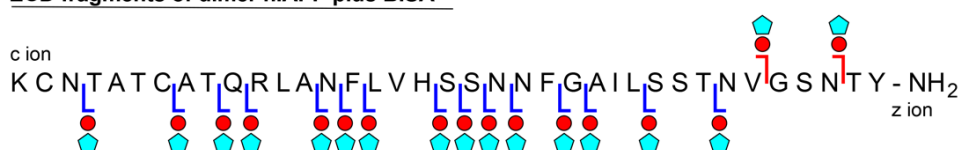
B) BISA, benzimidazole group (from BISA), sulfonic group (from BISA)
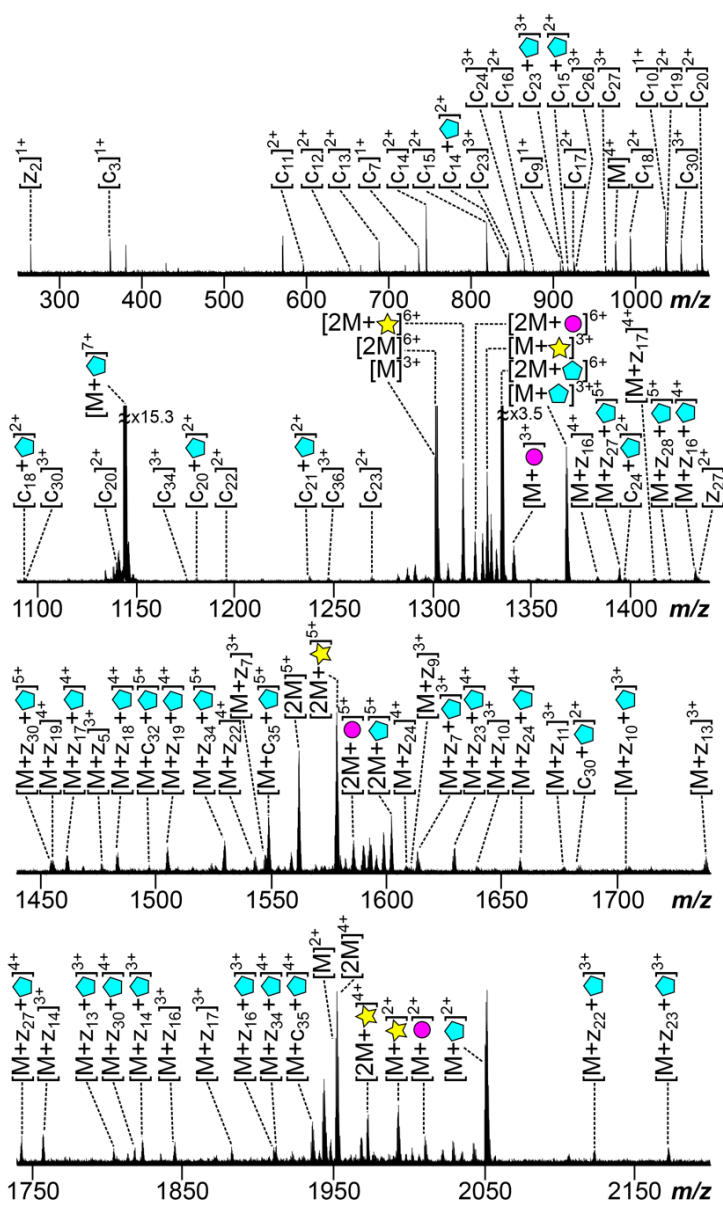
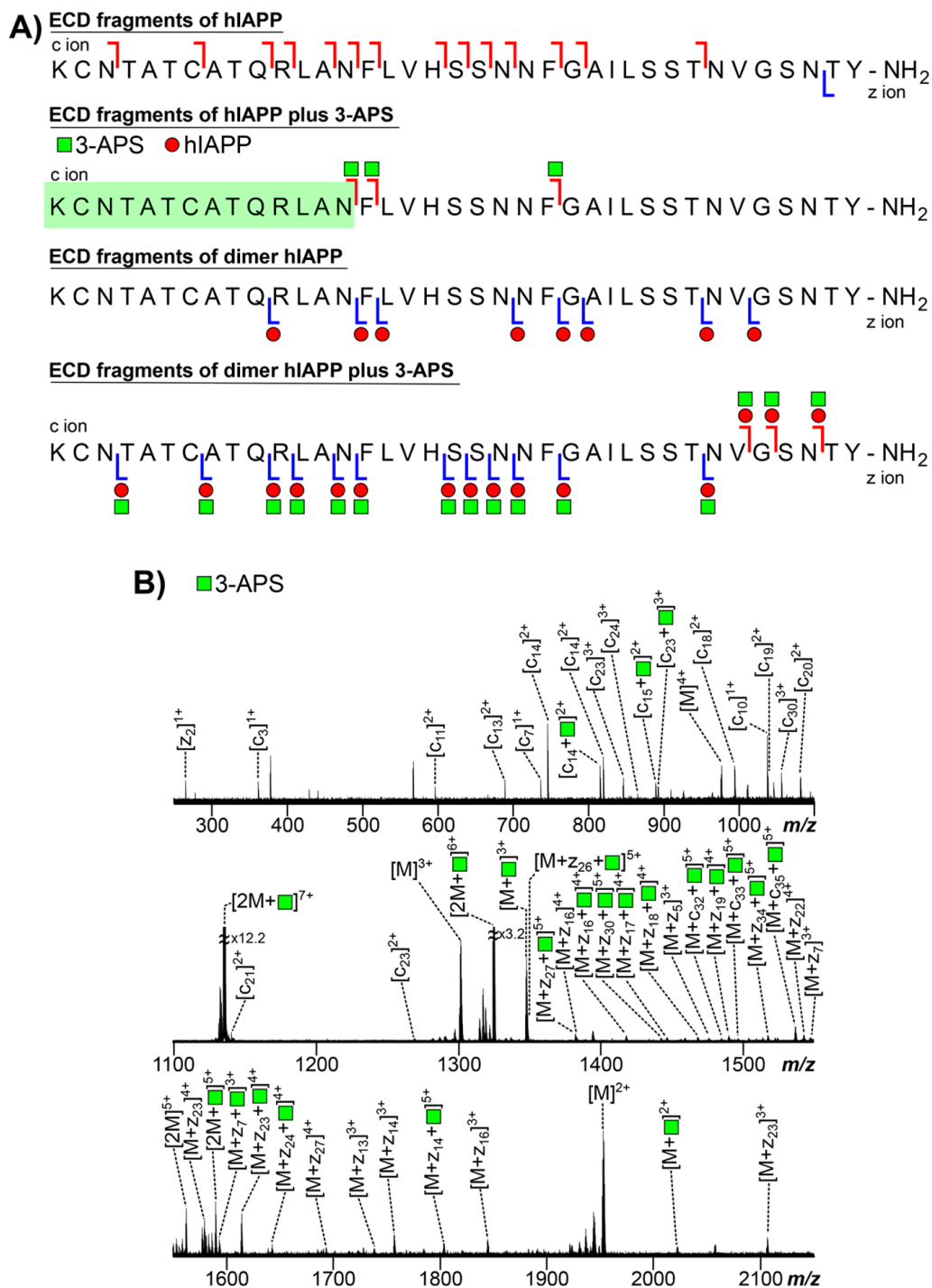**Figure S3.15.** (A) The summarised ECD fragments of the 7+ charge state molecular ion of the wild-type hIAPP dimer with a BISA molecule. The highlight area (coloured in cyan) indicates the proposed non-interruption binding region, which does not interfere with the formation of the early oligomers between the wild-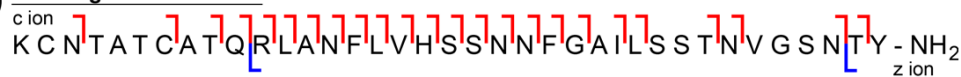type hIAPP and 3-APS molecule. (B) The ECD MS/MS spectrum of the 7+ charge state molecular ion of the wild-type hIAPP dimer with a BISA molecule. The assigned fragments are listed in the Supplementary Information Table S3.9.

Assigned ECD fragments of the 7+ molecular ion of the dimer hIAPP with BISA

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	2+	1952.444365	1952.444394	0.01
hIAPP only	hIAPP	M	3+	1301.295128	1301.295018	-0.08
hIAPP only	hIAPP	M	4+	976.724449	976.723770	-0.70
hIAPP only	hIAPP	c3	1+	362.173076	362.173065	-0.03
hIAPP only	hIAPP	c7	1+	737.306908	737.306870	-0.05
hIAPP only	hIAPP	c9	1+	909.391701	909.391878	0.19
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450354	0.07
hIAPP only	hIAPP	c11	2+	597.279333	597.279363	0.05
hIAPP only	hIAPP	c12	2+	653.821365	653.821401	0.06
hIAPP only	hIAPP	c13	2+	689.339922	689.339842	-0.12
hIAPP only	hIAPP	c14	2+	746.361386	746.361401	0.02
hIAPP only	hIAPP	c15	2+	819.895593	819.895710	0.14
hIAPP only	hIAPP	c16	2+	876.437625	876.437534	-0.10
hIAPP only	hIAPP	c17	2+	926.473206	926.473828	0.67
hIAPP only	hIAPP	c18	2+	994.501288	994.501501	0.21
hIAPP only	hIAPP	c19	2+	1038.017303	1038.017001	-0.29
hIAPP only	hIAPP	c20	2+	1081.533317	1081.533079	-0.22
hIAPP only	hIAPP	c21	2+	1138.554781	1138.555808	0.90
hIAPP only	hIAPP	c22	2+	1195.576245	1195.575874	-0.31
hIAPP only	hIAPP	c23	2+	1269.110452	1269.109694	-0.60
hIAPP only	hIAPP	c23	3+	846.409393	846.409262	-0.15
hIAPP only	hIAPP	c24	3+	865.750796	865.751018	0.26
hIAPP only	hIAPP	c26	3+	926.790274	926.789825	-0.48
hIAPP only	hIAPP	c27	3+	964.484962	964.484681	-0.29
hIAPP only	hIAPP	c30	3+	1056.188874	1056.188449	-0.40
hIAPP only	hIAPP	c31	3+	1094.535553	1094.535391	-0.15
hIAPP only	hIAPP	c34	3+	1175.578077	1175.577514	-0.48
hIAPP only	hIAPP	c36	3+	1246.940688	1246.940764	0.06
hIAPP only	hIAPP	z2	1+	266.126109	266.126112	0.01
hIAPP only	hIAPP	z27	2+	1434.222891	1434.222847	-0.03
hIAPP plus benzi- group	hIAPP	M	2+	2010.463090	2010.463422	0.17
hIAPP plus benzi- group	hIAPP	M	3+	1340.646160	1340.645969	-0.14
hIAPP plus sulf- group	hIAPP	M	2+	1993.929534	1993.928669	-0.43
hIAPP plus sulf- group	hIAPP	M	3+	1328.951654	1328.951570	-0.06
hIAPP plus BISA	hIAPP	M	2+	2050.444010	2050.442618	-0.68
hIAPP plus BISA	hIAPP	M	3+	1367.632695	1367.632413	-0.21
hIAPP plus BISA	hIAPP	c14	2+	845.366343	845.366392	0.06
hIAPP plus BISA	hIAPP	c15	2+	918.900550	918.899851	-0.76
hIAPP plus BISA	hIAPP	c18	2+	1093.506245	1093.505204	-0.95
hIAPP plus BISA	hIAPP	c20	2+	1180.538274	1180.537940	-0.28
hIAPP plus BISA	hIAPP	c21	2+	1237.559738	1237.558890	-0.69
hIAPP plus BISA	hIAPP	c23	3+	912.412698	912.412989	0.32
hIAPP plus BISA	hIAPP	c24	2+	1397.628890	1397.628477	-0.30
hIAPP plus BISA	hIAPP	c30	2+	1682.784626	1682.785326	0.42
2 hIAPP	2 hIAPP	M	4+	1952.692121	1952.692210	0.05
2 hIAPP	2 hIAPP	M	5+	1562.154698	1562.153180	-0.97
2 hIAPP	2 hIAPP	M	6+	1301.796420	1301.795906	-0.39
2 hIAPP	hIAPP	z5	3+	1476.370423	1476.370504	0.05
2 hIAPP	hIAPP	z7	3+	1547.071595	1547.071132	-0.30
2 hIAPP	hIAPP	z9	3+	1610.769064	1610.767991	-0.67
2 hIAPP	hIAPP	z10	3+	1639.445612	1639.445990	0.23
2 hIAPP	hIAPP	z11	3+	1677.140310	1677.141120	0.48
2 hIAPP	hIAPP	z13	3+	1738.179867	1738.180963	0.63
2 hIAPP	hIAPP	z14	3+	1758.187724	1758.188378	0.37
2 hIAPP	hIAPP	z16	3+	1844.892492	1844.892948	0.25
2 hIAPP	hIAPP	z16	4+	1383.418608	1383.418395	-0.15
2 hIAPP	hIAPP	z17	3+	1882.906802	1882.905923	-0.47
2 hIAPP	hIAPP	z17	4+	1412.179946	1412.178707	-0.88
2 hIAPP	hIAPP	z19	4+	1455.695966	1455.696161	0.13
2 hIAPP	hIAPP	z22	4+	1543.250787	1543.250839	0.03
2 hIAPP	hIAPP	z24	4+	1609.029828	1609.029630	-0.12
2 hIAPP plus benzi- group	hIAPP	M	5+	1585.765325	1585.765717	0.25

257

(Table S3.9. Continue)

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
2 hIAPP plus benzi- group	hIAPP	M	6+	1321.638984	1321.638004	-0.74
2 hIAPP plus sulf- group	hIAPP	M	4+	1972.934620	1972.934332	-0.15
2 hIAPP plus sulf- group	hIAPP	M	5+	1578.547987	1578.547033	-0.60
2 hIAPP plus sulf- group	hIAPP	M	6+	1314.789234	1314.788576	-0.50
2 hIAPP plus BISA	hIAPP	M	5+	1602.360185	1602.359576	-0.38
2 hIAPP plus BISA	hIAPP	M	6+	1334.464751	1334.463705	-0.78
2 hIAPP plus BISA	hIAPP	M	7+	1145.688044	1145.687153	-0.78
2 hIAPP plus BISA	hIAPP	c32	5+	1497.515521	1497.515466	-0.04
2 hIAPP plus BISA	hIAPP	c35	4+	1935.916213	1935.914436	-0.92
2 hIAPP plus BISA	hIAPP	c35	5+	1549.134820	1549.134067	-0.49
2 hIAPP plus BISA	hIAPP	z7	3+	1613.409013	1613.408341	-0.42
2 hIAPP plus BISA	hIAPP	z10	3+	1705.114750	1705.114910	0.09
2 hIAPP plus BISA	hIAPP	z13	3+	1804.515489	1804.513717	-0.98
2 hIAPP plus BISA	hIAPP	z14	3+	1823.190322	1823.190371	0.03
2 hIAPP plus BISA	hIAPP	z16	3+	1910.227438	1910.227277	-0.08
2 hIAPP plus BISA	hIAPP	z16	4+	1432.921063	1432.921432	0.26
2 hIAPP plus BISA	hIAPP	z17	4+	1461.682370	1461.682993	0.43
2 hIAPP plus BISA	hIAPP	z18	4+	1483.189804	1483.189987	0.12
2 hIAPP plus BISA	hIAPP	z19	4+	1505.198392	1505.198534	0.09
2 hIAPP plus BISA	hIAPP	z22	3+	2122.999284	2122.998777	-0.24
2 hIAPP plus BISA	hIAPP	z23	3+	2173.026381	2173.024415	-0.90
2 hIAPP plus BISA	hIAPP	z23	4+	1629.770906	1629.771747	0.52
2 hIAPP plus BISA	hIAPP	z24	4+	1658.031070	1658.031424	0.21
2 hIAPP plus BISA	hIAPP	z27	4+	1743.086650	1743.086502	-0.08
2 hIAPP plus BISA	hIAPP	z27	5+	1394.469211	1394.468215	-0.71
2 hIAPP plus BISA	hIAPP	z28	5+	1420.082006	1420.082623	0.43
2 hIAPP plus BISA	hIAPP	z30	4+	1818.373088	1818.372197	-0.49
2 hIAPP plus BISA	hIAPP	z30	5+	1454.899926	1454.899119	-0.55
2 hIAPP plus BISA	hIAPP	z34	4+	1912.156503	1912.156783	0.15
2 hIAPP plus BISA	hIAPP	z34	5+	1529.926651	1529.925915	-0.48

					Average Error (ppm):	-0.16
					Average Absolute Error (ppm):	0.34
					Standard Deviation:	0.41

Table S3.9. List of the assigned ECD fragments of the 7+ molecular ion of the wild-type hIAPP dimer with a BISA molecule.
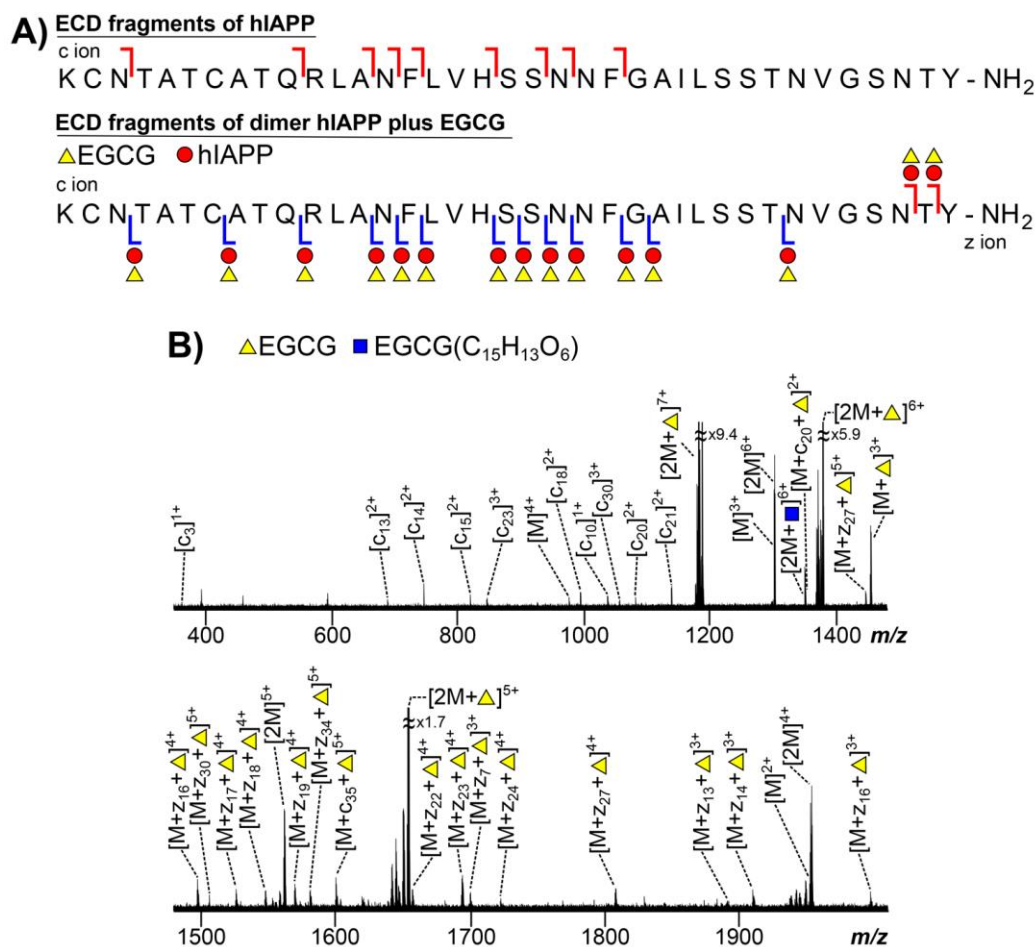
Figure S3.16. (A) The summarised ECD fragments of the 7+ charge state molecular ion of the wild-type hIAPP dimer with an EGCG molecule. (B) The ECD MS/MS spectrum of the 7+ charge state molecular ion of the wild-type hIAPP dimer with an EGCG molecule. The assigned fragments are listed in the Supplementary Information Table S3.10.

Assigned ECD fragments of the 7+ molecular ion of the dimer hIAPP with EGCG

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
hIAPP only	hIAPP	M	4+	976.974975	976.975008	-0.03
hIAPP only	hIAPP	M	3+	1301.295128	1301.295776	-0.50
hIAPP only	hIAPP	M	2+	1951.942966	1951.942591	0.19
hIAPP only	hIAPP	c3	1+	363.180901	363.180899	0.01
hIAPP only	hIAPP	c10	1+	1037.450279	1037.450212	0.06
hIAPP only	hIAPP	c13	2+	689.339922	689.339916	0.01
hIAPP only	hIAPP	c14	2+	746.361386	746.361454	-0.09
hIAPP only	hIAPP	c15	2+	819.895593	819.895655	-0.08
hIAPP only	hIAPP	c18	2+	994.501288	994.501191	0.10
hIAPP only	hIAPP	c20	2+	1081.533317	1081.533550	-0.22
hIAPP only	hIAPP	c21	2+	1138.554781	1138.554661	0.11
hIAPP only	hIAPP	c23	3+	846.743643	846.743739	-0.11
hIAPP only	hIAPP	c30	3+	1056.188872	1056.188450	0.40
hIAPP Plus EGCG	hIAPP Plus EGCG	M	3+	1454.658543	1454.658360	0.13
2 hIAPP	2 hIAPP	M	6+	1302.131432	1302.131428	0.00
2 hIAPP	2 hIAPP	M	5+	1562.556695	1562.556662	0.02
2 hIAPP	2 hIAPP	M	4+	1953.197422	1953.197217	0.10
2 hIAPP Plus EGCG ($C_{15}H_{13}O_6$)	2 hIAPP Plus EGCG ($C_{15}H_{13}O_6$)	M	6+	1350.142023	1350.142086	-0.05
2 hIAPP Plus EGCG	2 hIAPP Plus EGCG	M	7+	1181.123684	1181.123535	0.13
2 hIAPP Plus EGCG	2 hIAPP Plus EGCG	M	6+	1378.144824	1378.144461	0.26
2 hIAPP Plus EGCG	2 hIAPP Plus EGCG	M	5+	1654.174856	1654.174121	0.44
2 hIAPP Plus EGCG	hIAPP	c35	5+	1600.547484	1600.547007	0.30
2 hIAPP Plus EGCG	hIAPP	c36	6+	1351.467810	1351.467714	0.07
2 hIAPP Plus EGCG	hIAPP	z7	3+	1700.434897	1700.434369	0.31
2 hIAPP Plus EGCG	hIAPP	z13	3+	1891.207272	1891.207845	-0.30
2 hIAPP Plus EGCG	hIAPP	z14	3+	1909.880271	1909.879161	0.58
2 hIAPP Plus EGCG	hIAPP	z16	4+	1497.939865	1497.939983	-0.08
2 hIAPP Plus EGCG	hIAPP	z16	3+	1997.251561	1997.252443	-0.44
2 hIAPP Plus EGCG	hIAPP	z17	4+	1526.450596	1526.449737	0.56
2 hIAPP Plus EGCG	hIAPP	z18	4+	1548.208603	1548.208486	0.08
2 hIAPP Plus EGCG	hIAPP	z19	4+	1570.217240	1570.217123	0.07
2 hIAPP Plus EGCG	hIAPP	z22	4+	1657.268145	1657.268300	-0.09
2 hIAPP Plus EGCG	hIAPP	z23	4+	1694.537832	1694.537876	-0.03
2 hIAPP Plus EGCG	hIAPP	z24	4+	1722.048563	1722.047900	0.39
2 hIAPP Plus EGCG	hIAPP	z27	5+	1446.885254	1446.885678	-0.29
2 hIAPP Plus EGCG	hIAPP	z27	4+	1807.602877	1807.602175	0.39
2 hIAPP Plus EGCG	hIAPP	z30	5+	1506.713451	1506.712581	0.58
2 hIAPP Plus EGCG	hIAPP	z34	5+	1581.538610	1581.538951	-0.22

Average Error (ppm): 0.07
Average Absolute Error (ppm): 0.21
Standard Deviation: 0.26

Table S3.10. List of the assigned ECD fragments of the 7+ molecular ion of the wild-type hIAPP dimer with an EGCG molecule.
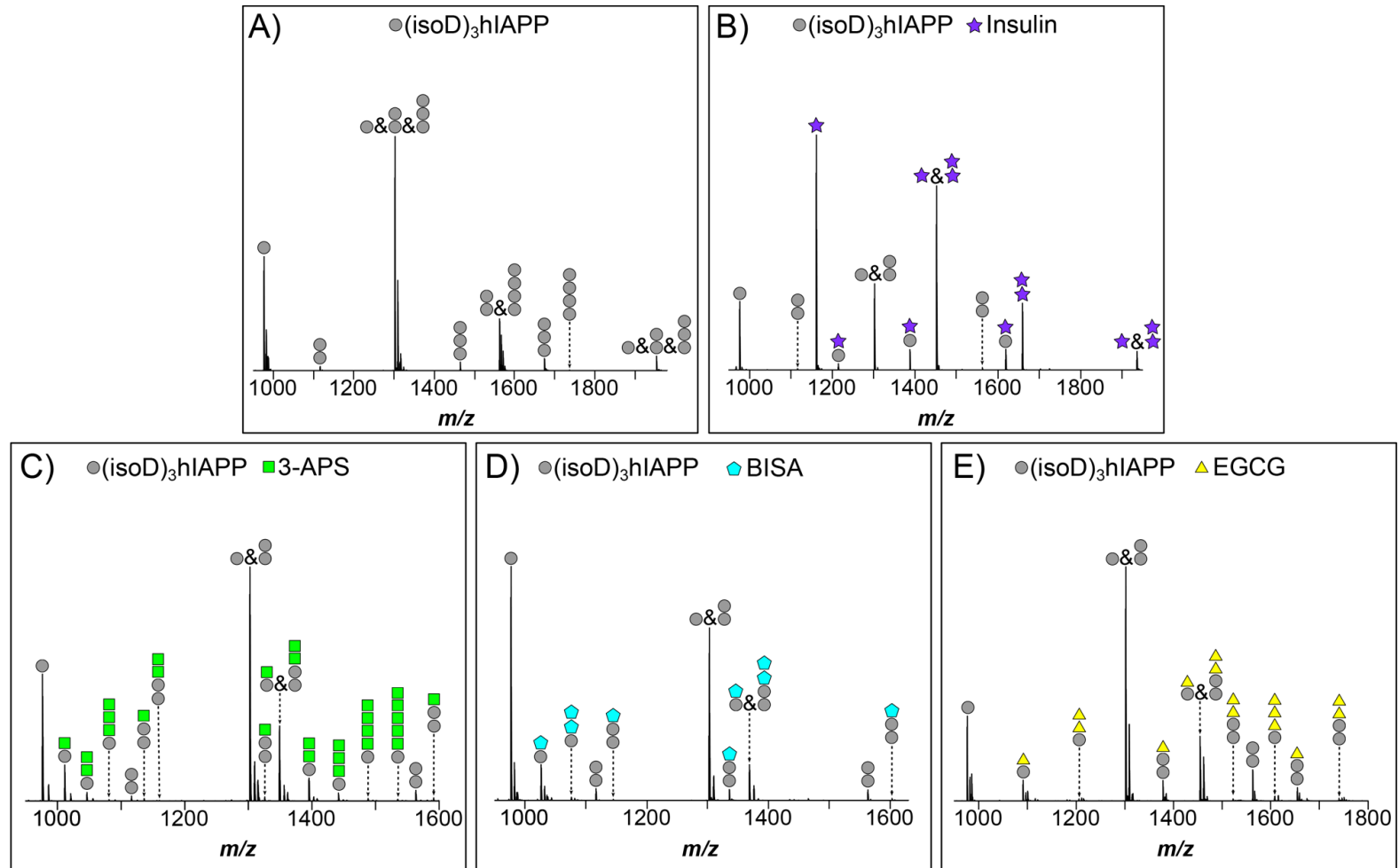
Figure S3.17.   The MS spectra of (A) the 10 µM solution of the mutant (isoD)₃hIAPP (grey circle) only, the 10 µM aqueous solutions of the mutant (isoD)₃hIAPP mixed with the (B) 10 µM insulin (purple star), (C) 10 µM 3-APS (green square), (D) 10 µM BISA (cyan pentagon), and (E) 10 µM EGCG (yellow triangle).

A)

ECD fragments of Insulin Chain A

c ion
G I V E Q C C T S I C S L Y Q L E N Y C N

ECD fragments of Insulin Chain B

c ion
F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T
z ion
y ion

ECD fragments of (isoD)₃hIAPP

c ion
K C N T A T C A T Q R L A N F L V H S S (iD) (iD) F G A I L S S T N V G S (iD) T Y - NH₂

ECD fragments of (isoD)₃hIAPP plus insulin

★ Insulin (Chain A&B) ● (isoD)₃hIAPP
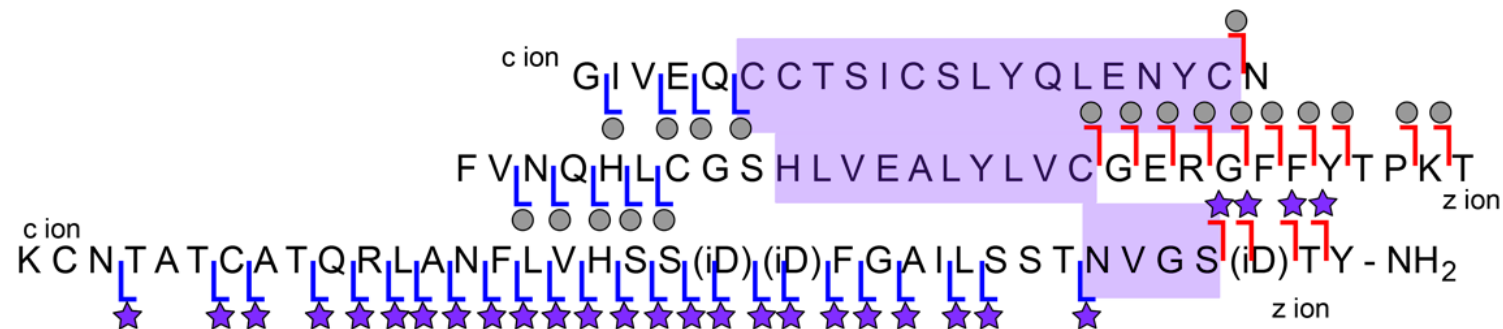
c ion G I V E Q C C T S I C S L Y Q L E N Y C N

F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T
z ion

c ion
K C N T A T C A T Q R L A N F L V H S S (iD) (iD) F G A I L S S T N V G S (iD) T Y - NH₂
z ion

(Figure S3.18. Continue)
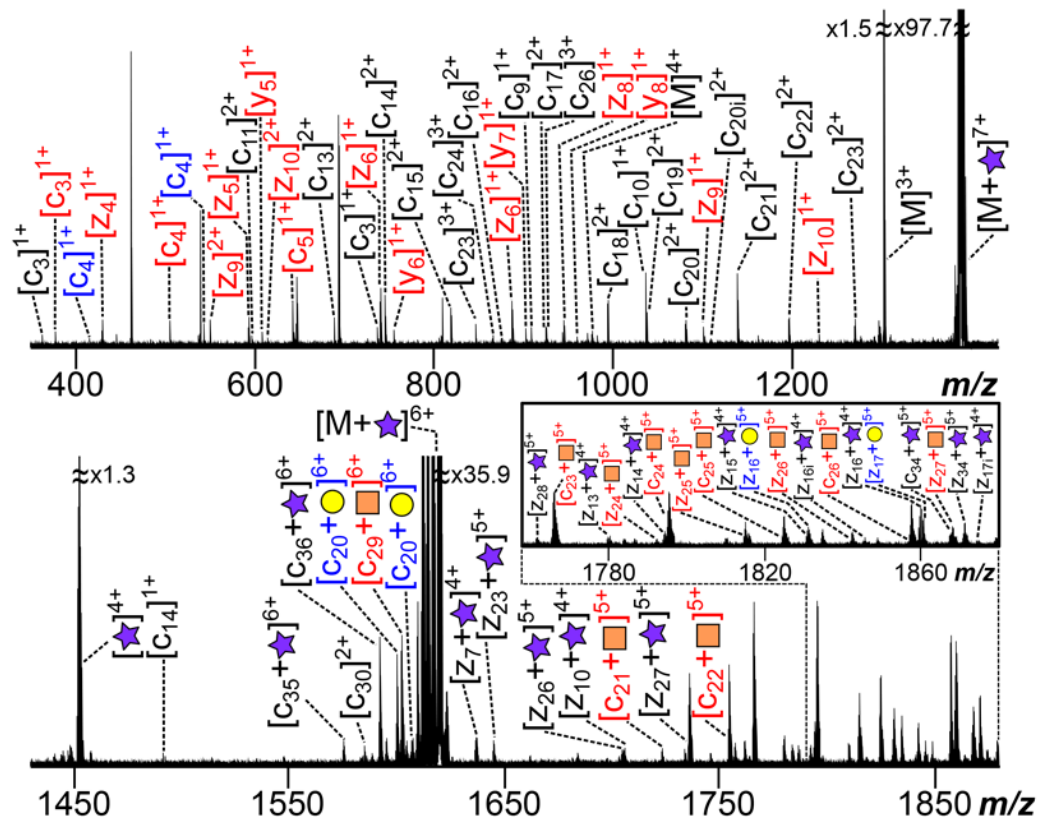
(Figure S3.18. Continue)
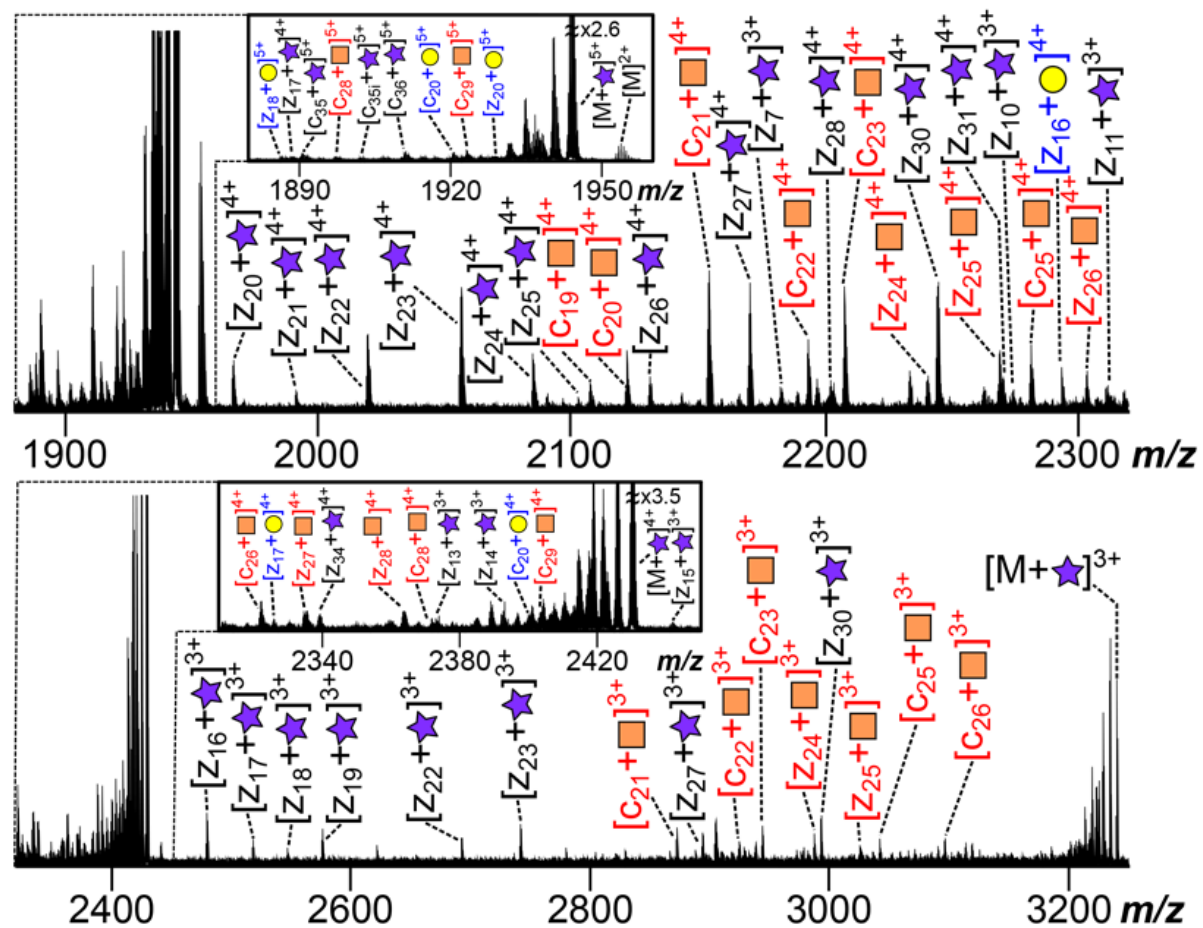
Figure S3.18. (A) The summarised ECD fragments of the 7+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and an intact insulin molecule. The highlighted area (coloured in purple) indicates the proposed interaction region between the monomer unit of the mutant (isoD)$_3$hIAPP and an insulin molecule which is between Asn-31 and Ser-34 on the mutant (isoD)$_3$hIAPP molecule. (B) The ECD MS/MS spectrum of the 7+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and an intact insulin molecule. The assigned fragments are listed in the Supplementary Information Table S3.11.

Assigned ECD fragments of the 7+ molecular ion of
(isoD)$_3$hIAPP with Insulin

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
(isoD)$_3$ only	(isoD)$_3$	M	2+	1952.915077	1952.916573	-0.77
(isoD)$_3$ only	(isoD)$_3$	M	3+	1302.947539	1302.947937	-0.31
(isoD)$_3$ only	(isoD)$_3$	M	4+	976.961177	976.961469	-0.30
(isoD)$_3$ only	(isoD)$_3$	c3	1+	363.180902	363.180901	0.00
(isoD)$_3$ only	(isoD)$_3$	c7	1+	737.306908	737.306922	-0.02
(isoD)$_3$ only	(isoD)$_3$	c9	1+	909.391701	909.392223	-0.57
(isoD)$_3$ only	(isoD)$_3$	c10	1+	1037.450279	1037.450500	-0.21
(isoD)$_3$ only	(isoD)$_3$	c11	2+	597.279333	597.279056	0.46
(isoD)$_3$ only	(isoD)$_3$	c13	2+	689.339922	689.339936	-0.02
(isoD)$_3$ only	(isoD)$_3$	c14	1+	1491.715496	1491.715986	-0.33
(isoD)$_3$ only	(isoD)$_3$	c14	2+	746.361386	746.361328	0.08
(isoD)$_3$ only	(isoD)$_3$	c15	2+	819.895593	819.895563	0.04
(isoD)$_3$ only	(isoD)$_3$	c16	2+	876.437625	876.438141	-0.59
(isoD)$_3$ only	(isoD)$_3$	c17	2+	925.971832	925.972433	-0.65
(isoD)$_3$ only	(isoD)$_3$	c18	2+	994.501288	994.501468	-0.18
(isoD)$_3$ only	(isoD)$_3$	c19	2+	1038.017303	1038.017273	0.03
(isoD)$_3$ only	(isoD)$_3$	c20	2+	1082.034683	1082.035086	-0.37
(isoD)$_3$ only	(isoD)$_3$	c20i	2+	1110.032142	1110.032510	-0.33
(isoD)$_3$ only	(isoD)$_3$	c21	2+	1139.046789	1139.046837	-0.04
(isoD)$_3$ only	(isoD)$_3$	c22	2+	1196.560261	1196.560006	0.21
(isoD)$_3$ only	(isoD)$_3$	c23	2+	1270.094468	1270.093553	0.72
(isoD)$_3$ only	(isoD)$_3$	c23	3+	847.065404	847.065136	0.32
(isoD)$_3$ only	(isoD)$_3$	c24	3+	866.072559	866.072370	0.22
(isoD)$_3$ only	(isoD)$_3$	c26	3+	927.446285	927.446898	-0.66
(isoD)$_3$ only	(isoD)$_3$	c30	2+	1585.265090	1585.265442	-0.22
Insulin only	Insulin only	M	4+	1452.668486	1452.668366	0.08
Insulin Chain A only	Insulin Chain A	c4	1+	416.250361	416.250383	-0.05
Insulin Chain A only	Insulin Chain A	c5	1+	544.308939	544.308813	0.23
Insulin Chain B only	Insulin Chain B	c3	1+	378.213581	378.213611	-0.08
Insulin Chain B only	Insulin Chain B	c4	1+	506.272159	506.272103	0.11
Insulin Chain B only	Insulin Chain B	c5	1+	643.331071	643.330997	0.12
Insulin Chain B only	Insulin Chain B	y5	1+	609.324255	609.324225	0.05
Insulin Chain B only	Insulin Chain B	y6	1+	756.392669	756.392628	0.05
Insulin Chain B only	Insulin Chain B	y7	1+	903.461083	903.461012	0.08
Insulin Chain B only	Insulin Chain B	y8	1+	960.482547	960.482506	0.04
Insulin Chain B only	Insulin Chain B	z4	1+	430.242202	430.242196	0.01
Insulin Chain B only	Insulin Chain B	z5	1+	593.305531	593.305469	0.10
Insulin Chain B only	Insulin Chain B	z6	1+	740.373945	740.373860	0.11
Insulin Chain B only	Insulin Chain B	z7	1+	887.442485	887.442444	0.05
Insulin Chain B only	Insulin Chain B	z8	1+	944.463823	944.464273	-0.48
Insulin Chain B only	Insulin Chain B	z9	1+	1101.572757	1101.572855	-0.09
Insulin Chain B only	Insulin Chain B	z9	2+	550.786105	550.786068	0.07
Insulin Chain B only	Insulin Chain B	z10	1+	1230.615350	1230.615080	0.22
Insulin Chain B only	Insulin Chain B	z10	2+	615.307402	615.307277	0.20
(isoD)$_3$ Plus Insulin	(isoD)$_3$ Plus Insulin	M	3+	3240.173471	3240.173590	-0.04
(isoD)$_3$ Plus Insulin	(isoD)$_3$ Plus Insulin	M	4+	2430.129966	2430.130028	-0.03
(isoD)$_3$ Plus Insulin	(isoD)$_3$ Plus Insulin	M	5+	1944.103863	1944.103700	0.08
(isoD)$_3$ Plus Insulin	(isoD)$_3$ Plus Insulin	M	6+	1619.919462	1619.919053	0.25
(isoD)$_3$ Plus Insulin	(isoD)$_3$ Plus Insulin	M	7+	1387.929577	1387.929056	0.38
(isoD)$_3$ Plus Insulin	Insulin Chain A	c20	4+	2399.864385	2399.863176	0.50
(isoD)$_3$ Plus Insulin	Insulin Chain A	c20	5+	1921.096063	1921.096710	-0.34
(isoD)$_3$ Plus Insulin	Insulin Chain A	c20	6+	1601.080263	1601.080577	-0.20
(isoD)$_3$ Plus Insulin	Insulin Chain A	z16	4+	2294.553419	2294.553690	-0.12
(isoD)$_3$ Plus Insulin	Insulin Chain A	z16	5+	1834.841709	1834.841128	0.32
(isoD)$_3$ Plus Insulin	Insulin Chain A	z17	4+	2326.818517	2326.819567	-0.45
(isoD)$_3$ Plus Insulin	Insulin Chain A	z17	5+	1860.452849	1860.451866	0.53
(isoD)$_3$ Plus Insulin	Insulin Chain A	z18	5+	1886.662172	1886.661199	0.52
(isoD)$_3$ Plus Insulin	Insulin Chain A	z20	5+	1928.691886	1928.692517	-0.33
(isoD)$_3$ Plus Insulin	Insulin Chain A	z20	6+	1607.745126	1607.744601	0.33
(isoD)$_3$ Plus Insulin	Insulin Chain B	c19	4+	2107.468995	2107.467783	0.58
(isoD)$_3$ Plus Insulin	Insulin Chain B	c20	4+	2122.225339	2122.225238	0.05
(isoD)$_3$ Plus Insulin	Insulin Chain B	c21	3+	2872.314220	2872.315547	-0.46
(isoD)$_3$ Plus Insulin	Insulin Chain B	c21	4+	2154.486003	2154.485220	0.36
(isoD)$_3$ Plus Insulin	Insulin Chain B	c21	5+	1723.789056	1723.788772	0.16
(isoD)$_3$ Plus Insulin	Insulin Chain B	c22	3+	2924.347928	2924.350542	-0.89
(isoD)$_3$ Plus Insulin	Insulin Chain B	c22	4+	2193.512766	2193.512357	0.19
(isoD)$_3$ Plus Insulin	Insulin Chain B	c22	5+	1754.808920	1754.808280	0.36
(isoD)$_3$ Plus Insulin	Insulin Chain B	c23	3+	2943.689059	2943.689486	-0.15
(isoD)$_3$ Plus Insulin	Insulin Chain B	c23	4+	2208.018613	2208.018861	-0.11
(isoD)$_3$ Plus Insulin	Insulin Chain B	c23	5+	1766.214397	1766.214135	0.15
(isoD)$_3$ Plus Insulin	Insulin Chain B	c24	5+	1796.028862	1796.028287	0.32
(isoD)$_3$ Plus Insulin	Insulin Chain B	c25	3+	3043.070451	3043.072454	-0.66
(isoD)$_3$ Plus Insulin	Insulin Chain B	c25	4+	2281.803352	2281.805163	-0.79
(isoD)$_3$ Plus Insulin	Insulin Chain B	c25	5+	1825.041795	1825.041970	-0.10

(Table S3.11. Continue)

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
(isoD)$_3$ Plus Insulin	Insulin Chain B	c26	3+	3096.089220	3096.088739	0.16
(isoD)$_3$ Plus Insulin	Insulin Chain B	c26	4+	2322.569218	2322.569675	-0.20
(isoD)$_3$ Plus Insulin	Insulin Chain B	c26	5+	1857.854878	1857.854278	0.32
(isoD)$_3$ Plus Insulin	Insulin Chain B	c28	4+	2372.094357	2372.093390	0.41
(isoD)$_3$ Plus Insulin	Insulin Chain B	c28	5+	1897.675376	1897.675157	0.12
(isoD)$_3$ Plus Insulin	Insulin Chain B	c29	4+	2403.615663	2403.616483	-0.34
(isoD)$_3$ Plus Insulin	Insulin Chain B	c29	5+	1923.294381	1923.294992	-0.32
(isoD)$_3$ Plus Insulin	Insulin Chain B	c29	6+	1602.913197	1602.913307	-0.07
(isoD)$_3$ Plus Insulin	Insulin Chain B	z24	3+	2988.366670	2988.367640	-0.32
(isoD)$_3$ Plus Insulin	Insulin Chain B	z24	4+	2240.773940	2240.773485	0.20
(isoD)$_3$ Plus Insulin	Insulin Chain B	z24	5+	1792.419876	1792.419161	0.40
(isoD)$_3$ Plus Insulin	Insulin Chain B	z25	3+	3026.397342	3026.396963	0.13
(isoD)$_3$ Plus Insulin	Insulin Chain B	z25	4+	2269.046508	2269.047066	-0.25
(isoD)$_3$ Plus Insulin	Insulin Chain B	z25	5+	1815.437480	1815.436826	0.36
(isoD)$_3$ Plus Insulin	Insulin Chain B	z26	4+	2303.561725	2303.562573	-0.37
(isoD)$_3$ Plus Insulin	Insulin Chain B	z26	5+	1842.448485	1842.448164	0.17
(isoD)$_3$ Plus Insulin	Insulin Chain B	z27	4+	2335.826840	2335.825940	0.39
(isoD)$_3$ Plus Insulin	Insulin Chain B	z27	5+	1868.260608	1868.260160	0.24
(isoD)$_3$ Plus Insulin	Insulin Chain B	z28	4+	2365.088898	2365.089350	-0.19
(isoD)$_3$ Plus Insulin	(isoD)$_3$	c34	5+	1867.672710	1867.672031	0.36
(isoD)$_3$ Plus Insulin	(isoD)$_3$	c35	5+	1890.878497	1890.878405	0.05
(isoD)$_3$ Plus Insulin	(isoD)$_3$	c35	6+	1575.731990	1575.731976	0.01
(isoD)$_3$ Plus Insulin	(isoD)$_3$	c35i	5+	1902.278035	1902.278605	-0.30
(isoD)$_3$ Plus Insulin	(isoD)$_3$	c36	5+	1910.887650	1910.887764	-0.06
(isoD)$_3$ Plus Insulin	(isoD)$_3$	c36	6+	1592.406283	1592.406213	0.04
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z7	3+	2182.995768	2182.994961	0.37
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z7	4+	1637.497139	1637.497312	-0.11
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z10	3+	2274.363773	2274.364133	-0.16
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z10	4+	1705.523604	1705.523105	0.29
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z11	3+	2312.058484	2312.058788	-0.13
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z13	3+	2374.100016	2374.099224	0.33
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z13	4+	1781.077386	1781.077523	-0.08
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z14	3+	2392.775341	2392.775875	-0.22
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z14	4+	1794.581369	1794.581723	-0.20
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z15	3+	2441.798178	2441.799229	-0.43
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z15	4+	1831.348497	1831.348918	-0.23
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z16-i	4+	1845.855821	1845.857090	-0.69
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z16	3+	2481.476328	2481.477078	-0.30
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z16	4+	1860.105241	1860.105782	-0.29
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z17-i	4+	1874.362038	1874.362807	-0.41
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z17	3+	2518.816823	2518.817651	-0.33
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z17	4+	1888.360880	1888.360677	0.11
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z18	3+	2547.493513	2547.491350	0.85
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z19	3+	2576.504196	2576.504353	-0.06
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z20	4+	1966.894700	1966.893921	0.40
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z21	4+	1991.409858	1991.409202	0.33
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z22	3+	2693.242048	2693.242231	-0.07
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z22	4+	2019.680886	2019.680124	0.38
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z23	3+	2742.264893	2742.265273	-0.14
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z23	4+	2056.698532	2056.698955	-0.21
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z23	5+	1645.359863	1645.359065	0.49
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z24	4+	2085.710560	2085.709268	0.62
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z25	4+	2103.219063	2103.218598	0.22
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z26	4+	2130.738494	2130.737027	0.69
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z26	5+	1704.792251	1704.792062	0.11
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z27	3+	2895.024600	2895.023753	0.29
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z27	4+	2170.516822	2170.517872	-0.48
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z27	5+	1736.212920	1736.212575	0.20
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z28	4+	2202.030375	2202.029056	0.60
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z28	5+	1762.025070	1762.025125	-0.03
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z30	3+	2994.066584	2994.066236	0.12
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z30	4+	2245.036214	2245.037564	-0.60
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z31	4+	2270.801007	2270.802349	-0.59
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z34	4+	2340.338033	2340.338894	-0.37
(isoD)$_3$ Plus Insulin	(isoD)$_3$	z34	5+	1871.468773	1871.468687	0.05

Average Error (ppm):	-0.01
Average Absolute Error (ppm):	0.27
Standard Deviation:	0.34

Table S3.11. List of the assigned ECD fragments of the 7+ molecular ion of the mutant (isoD)$_3$hIAPP monomer unit with an intact insulin molecule.

Figure S3.19. (A) The summarised ECD fragments of the 4+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and a 3-APS molecule. The highlighted area (coloured in green) indicates the proposed interaction region between the monomer unit of the mutant (isoD)$_3$hIAPP and a 3-APS molecule which is between Arg-11 and Gly-33 on the mutant (isoD)$_3$hIAPP molecule. (B) The ECD MS/MS spectrum of the 4+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and a 3-APS molecule. The assigned fragments are listed in the Supplementary Information Table S3.12.

Assigned ECD fragments of the 4+ molecular ion of (isoD)$_3$hIAPP with 3-APS

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
(isoD)$_3$ only	(isoD)$_3$	M	2+	1953.922902	1953.922894	0.00
(isoD)$_3$ only	(isoD)$_3$	M	3+	1302.615085	1302.614832	0.19
(isoD)$_3$ only	(isoD)$_3$	M	4+	976.961177	976.961816	-0.65
(isoD)$_3$ only	(isoD)$_3$	a10	1+	993.436637	993.436941	-0.31
(isoD)$_3$ only	(isoD)$_3$	c3	1+	362.173076	362.173076	0.00
(isoD)$_3$ only	(isoD)$_3$	c7	1+	737.306908	737.306885	0.03
(isoD)$_3$ only	(isoD)$_3$	c8	1+	808.344022	808.343989	0.04
(isoD)$_3$ only	(isoD)$_3$	c9	1+	909.391701	909.391743	-0.05
(isoD)$_3$ only	(isoD)$_3$	c10	1+	1037.450279	1037.450429	-0.14
(isoD)$_3$ only	(isoD)$_3$	c14	2+	746.361386	746.361484	-0.13
(isoD)$_3$ only	(isoD)$_3$	c15	2+	819.895593	819.895553	0.05
(isoD)$_3$ only	(isoD)$_3$	c17	2+	925.971832	925.972071	-0.26
(isoD)$_3$ only	(isoD)$_3$	c30	2+	1585.265090	1585.265249	-0.10
(isoD)$_3$ only	(isoD)$_3$	c31	2+	1641.785154	1641.785661	-0.31
(isoD)$_3$ only	(isoD)$_3$	c32	2+	1691.820760	1691.822215	-0.86
(isoD)$_3$ only	(isoD)$_3$	c33	2+	1720.331490	1720.332568	-0.63
(isoD)$_3$ only	(isoD)$_3$	c34	2+	1763.847504	1763.848099	-0.34
(isoD)$_3$ only	(isoD)$_3$	c34	3+	1176.234095	1176.233775	0.27
(isoD)$_3$ only	(isoD)$_3$	c34i	2+	1792.850248	1792.849775	0.26
(isoD)$_3$ only	(isoD)$_3$	c34i	3+	1195.233316	1195.234312	-0.83
(isoD)$_3$ only	(isoD)$_3$	c35	2+	1821.360977	1821.360738	0.13
(isoD)$_3$ only	(isoD)$_3$	c35	3+	1214.242145	1214.242789	-0.53
(isoD)$_3$ only	(isoD)$_3$	c36	2+	1872.388731	1872.389214	-0.26
(isoD)$_3$ only	(isoD)$_3$	c36	3+	1248.258971	1248.259106	-0.11
(isoD)$_3$ only	(isoD)$_3$	y4	1+	484.203806	484.203722	0.17
(isoD)$_3$ only	(isoD)$_3$	y5	1+	541.225270	541.225260	0.02
(isoD)$_3$ only	(isoD)$_3$	y6	1+	640.293684	640.293709	-0.04
(isoD)$_3$ only	(isoD)$_3$	z20	1+	2058.915038	2058.915667	-0.31
(isoD)$_3$ only	(isoD)$_3$	z23	1+	2417.133250	2417.134896	-0.68
(isoD)$_3$ only	(isoD)$_3$	z27	1+	2870.395527	2870.396010	-0.17
(isoD)$_3$ only	(isoD)$_3$	z27	2+	1435.197494	1435.196784	0.49
(isoD)$_3$ only	(isoD)$_3$	z28	2+	1499.226783	1499.226083	0.47
(isoD)$_3$ only	(isoD)$_3$	z34	2+	1772.836090	1772.836452	-0.20
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	M	2+	2024.440562	2024.439059	0.74
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	M	3+	1348.958523	1348.958439	0.06
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	M	4+	1011.718755	1011.718568	0.18
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	c33	3+	1194.236803	1194.236048	0.63
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	c34	3+	1221.907323	1221.907543	-0.18
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	c34i	3+	1241.242483	1241.242601	-0.10
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	c35	3+	1260.919844	1260.919014	0.66
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	c36	3+	1294.268143	1294.268256	-0.09
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	z27	2+	1505.214083	1505.214530	-0.30
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	z28	2+	1569.243368	1569.243996	-0.40
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	z29	2+	1619.767209	1619.767001	0.13
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	z30	2+	1654.784331	1654.784342	-0.01
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	z34	2+	1841.847334	1841.847989	-0.36
(isoD)$_3$ Plus 3-APS	(isoD)$_3$	z35	2+	1899.370219	1899.369759	0.24

Average Error (ppm): -0.08
Average Absolute Error (ppm): 0.28
Standard Deviation: 0.36

Table S3.12. List of the assigned ECD fragments of the 4+ molecular ion of the mutant (isoD)$_3$hIAPP monomer unit with a 3-APS molecule.

A) ECD fragments of (isoD)₃hIAPP

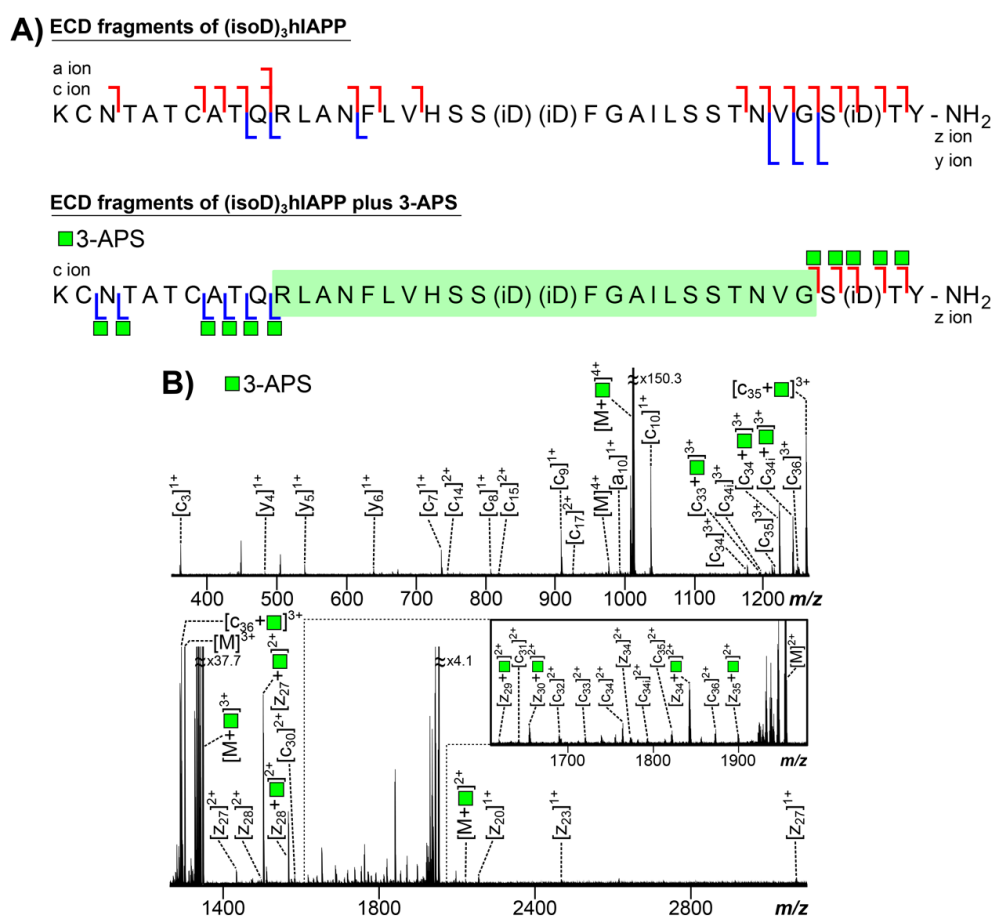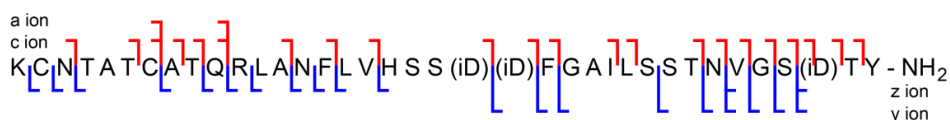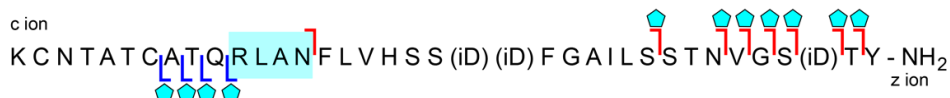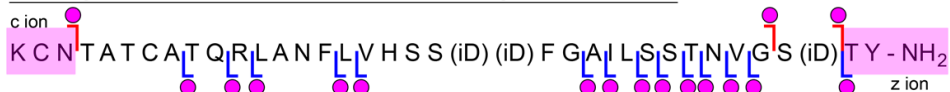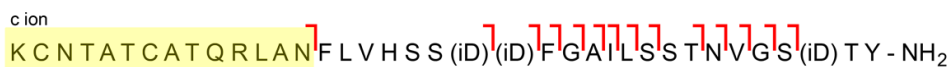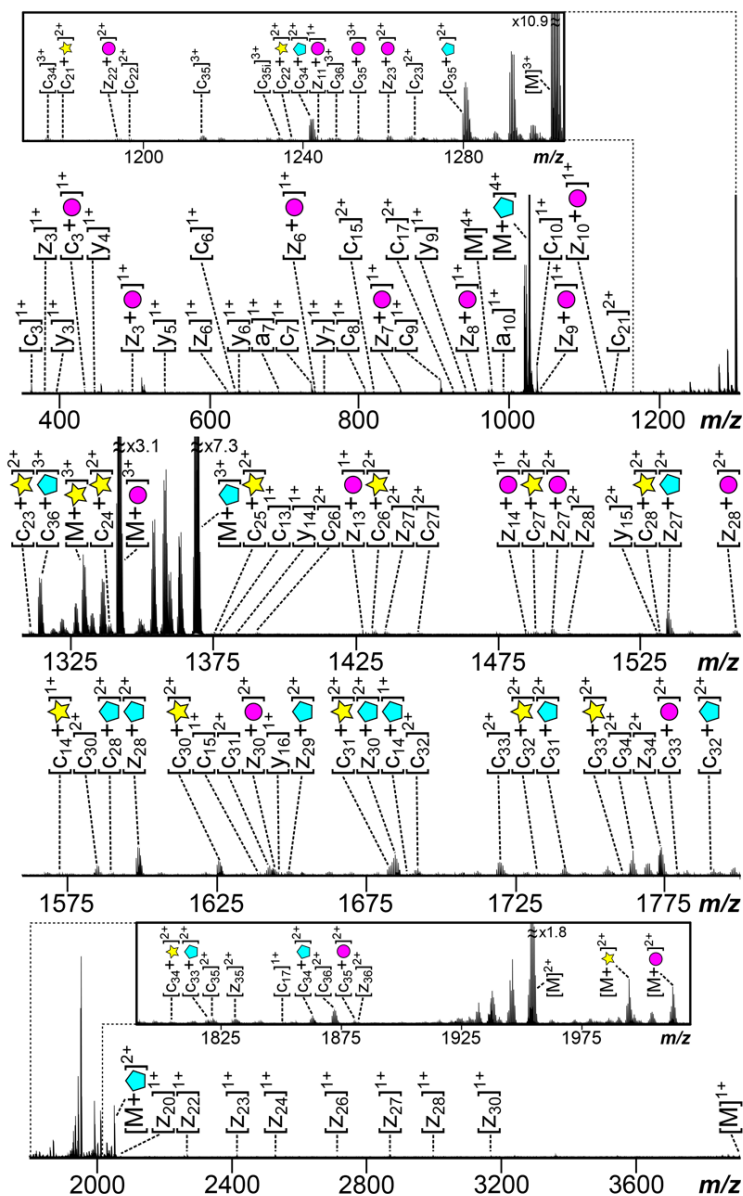**B)** BISA · benzimidazole group (from BISA) ☆ sulfonic group (from BISA)

Figure S3.20. (A) The summarised ECD fragments of the 4+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and a BISA molecule. The highlighted area (coloured in pink) indicates the proposed interaction region between the monomer unit of the mutant (isoD)$_3$hIAPP and the benzimidazole group from BISA molecule which is between Lys-1 and Asn-3, as well as, Thr-36 and Tyr-37 on the mutant (isoD)$_3$hIAPP molecule, the highlighted area (coloured in yellow) indicates the proposed interaction region between the monomer unit of the mutant (isoD)$_3$hIAPP and sulfonic acid group from BISA molecule which is between Lys-1 and Asn-15 on the mutant (isoD)$_3$hIAPP molecule, and the highlighted area (coloured in cyan) indicates the proposed interaction region between the monomer unit of the mutant (isoD)$_3$hIAPP and a BISA molecule which is between Arg-11 and Asn-15 on the mutant (isoD)$_3$hIAPP molecule. (B) The ECD MS/MS spectrum of the 4+ charge state molecular ion of the monomer unit of the mutant (isoD)$_3$hIAPP and a BISA molecule. The assigned fragments are listed in the Supplementary Information Table S3.13.

Assigned ECD fragments of the 4+ molecular ion of
(isoD)₃hIAPP with BISA

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
(isoD)₃ only	(isoD)₃	M	1+	3908.849765	3908.849451	0.08
(isoD)₃ only	(isoD)₃	M	2+	1952.915077	1952.914994	0.04
(isoD)₃ only	(isoD)₃	M	3+	1302.947539	1302.947334	0.16
(isoD)₃ only	(isoD)₃	M	4+	976.961177	976.960834	0.35
(isoD)₃ only	(isoD)₃	a7	1+	693.293268	693.293249	0.03
(isoD)₃ only	(isoD)₃	a10	1+	993.436637	993.436630	0.01
(isoD)₃ only	(isoD)₃	c3	1+	362.173076	362.173076	0.00
(isoD)₃ only	(isoD)₃	c6	1+	635.305548	635.305491	0.09
(isoD)₃ only	(isoD)₃	c7	1+	737.306908	737.306886	0.03
(isoD)₃ only	(isoD)₃	c8	1+	808.344022	808.344006	0.02
(isoD)₃ only	(isoD)₃	c9	1+	909.391701	909.391688	0.01
(isoD)₃ only	(isoD)₃	c10	1+	1037.450279	1037.450294	-0.01
(isoD)₃ only	(isoD)₃	c13	1+	1377.672568	1377.672497	0.05
(isoD)₃ only	(isoD)₃	c15	1+	1638.783910	1638.784469	-0.34
(isoD)₃ only	(isoD)₃	c15	2+	819.895593	819.895743	-0.18
(isoD)₃ only	(isoD)₃	c17	1+	1850.936388	1850.937367	-0.53
(isoD)₃ only	(isoD)₃	c17	2+	925.467918	925.468229	-0.34
(isoD)₃ only	(isoD)₃	c21	2+	1139.046789	1139.046968	-0.16
(isoD)₃ only	(isoD)₃	c22	2+	1196.560261	1196.560211	0.04
(isoD)₃ only	(isoD)₃	c23	2+	1271.097248	1271.097645	-0.31
(isoD)₃ only	(isoD)₃	c26	2+	1390.161873	1390.161275	0.43
(isoD)₃ only	(isoD)₃	c27	2+	1447.205308	1447.206014	-0.49
(isoD)₃ only	(isoD)₃	c30	2+	1584.761178	1584.761894	-0.45
(isoD)₃ only	(isoD)₃	c31	2+	1642.787956	1642.788143	-0.11
(isoD)₃ only	(isoD)₃	c32	2+	1692.322170	1692.321119	0.62
(isoD)₃ only	(isoD)₃	c33	2+	1720.832899	1720.832669	0.13
(isoD)₃ only	(isoD)₃	c34	2+	1763.346107	1763.345765	0.19
(isoD)₃ only	(isoD)₃	c34	3+	1175.898153	1175.898514	-0.31
(isoD)₃ only	(isoD)₃	c35	2+	1821.360977	1821.360961	0.01
(isoD)₃ only	(isoD)₃	c35	3+	1214.242145	1214.241947	0.16
(isoD)₃ only	(isoD)₃	c35i	3+	1233.577301	1233.577557	-0.21
(isoD)₃ only	(isoD)₃	c36	2+	1871.884818	1871.885821	-0.54
(isoD)₃ only	(isoD)₃	c36	3+	1248.258971	1248.259500	-0.42
(isoD)₃ only	(isoD)₃	y3	1+	397.171777	397.171764	0.03
(isoD)₃ only	(isoD)₃	y4	1+	484.203806	484.203791	0.03
(isoD)₃ only	(isoD)₃	y5	1+	541.225270	541.225256	0.03
(isoD)₃ only	(isoD)₃	y6	1+	640.293684	640.293672	0.02
(isoD)₃ only	(isoD)₃	y7	1+	754.336612	754.336773	-0.21
(isoD)₃ only	(isoD)₃	y9	1+	942.416320	942.416483	-0.17
(isoD)₃ only	(isoD)₃	y14	1+	1383.675055	1383.674748	0.22
(isoD)₃ only	(isoD)₃	y15	1+	1530.743535	1530.743469	0.04
(isoD)₃ only	(isoD)₃	y16	1+	1645.770413	1645.770084	0.20
(isoD)₃ only	(isoD)₃	z3	1+	381.153053	381.153030	0.06
(isoD)₃ only	(isoD)₃	z6	1+	625.282783	625.282693	0.14
(isoD)₃ only	(isoD)₃	z20	1+	2056.909419	2056.909152	0.13
(isoD)₃ only	(isoD)₃	z22	1+	2269.061897	2269.062052	-0.07
(isoD)₃ only	(isoD)₃	z23	1+	2417.133721	2417.133997	-0.11
(isoD)₃ only	(isoD)₃	z24	1+	2530.173239	2530.173905	-0.26
(isoD)₃ only	(isoD)₃	z26	1+	2715.297827	2715.296897	0.34
(isoD)₃ only	(isoD)₃	z27	1+	2871.398939	2871.398053	0.31
(isoD)₃ only	(isoD)₃	z27	2+	1435.197494	1435.197648	-0.11
(isoD)₃ only	(isoD)₃	z28	1+	2999.457516	2999.456319	0.40
(isoD)₃ only	(isoD)₃	z28	2+	1499.226783	1499.226508	0.18
(isoD)₃ only	(isoD)₃	z30	1+	3170.534484	3170.534057	0.13
(isoD)₃ only	(isoD)₃	z34	2+	1772.836096	1772.835372	0.41
(isoD)₃ only	(isoD)₃	z35	2+	1829.857553	1829.856237	0.72

Chapter 3 – The Inhibition Pathways of hIAPP

(Table S3.13 Continue)

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
(isoD)$_3$ only	(isoD)$_3$	z36	2+	1882.362527	1882.361968	0.30
(isoD)$_3$ plus benzi- group	(isoD)$_3$ plus benzi- group	M	2+	2011.941626	2011.941720	-0.05
(isoD)$_3$ plus benzi- group	(isoD)$_3$ plus benzi- group	M	3+	1341.294234	1341.294396	-0.12
(isoD)$_3$ plus benzi- group	(isoD)$_3$	c3	1+	479.218349	479.218327	0.05
(isoD)$_3$ plus benzi- group	(isoD)$_3$	c33	2+	1778.352729	1778.353663	-0.53
(isoD)$_3$ plus benzi- group	(isoD)$_3$	c35	2+	1879.883614	1879.884519	-0.48
(isoD)$_3$ plus benzi- group	(isoD)$_3$	c35	3+	1253.927444	1253.927456	-0.01
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z3	1+	498.198326	498.198325	0.00
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z6	1+	741.320233	741.320249	-0.02
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z7	1+	855.363161	855.363159	0.00
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z8	1+	956.410840	956.410280	0.59
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z9	1+	1043.442869	1043.442751	0.11
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z10	1+	1130.474898	1130.474961	-0.06
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z11	1+	1243.558962	1243.558999	-0.03
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z13	1+	1427.680140	1427.680211	-0.05
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z14	1+	1484.701604	1484.701472	0.09
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z22	2+	1193.053315	1193.053869	-0.46
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z23	2+	1266.587522	1266.587291	0.18
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z27	2+	1493.720131	1493.719647	0.32
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z28	2+	1558.250855	1558.250033	0.53
(isoD)$_3$ plus benzi- group	(isoD)$_3$	z30	2+	1643.791811	1643.791095	0.44
(isoD)$_3$ plus sulf- group	(isoD)$_3$ plus sulf- group	M	2+	1993.901309	1993.901042	0.13
(isoD)$_3$ plus sulf- group	(isoD)$_3$ plus sulf- group	M	3+	1329.601629	1329.601432	0.15
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c14	1+	1572.680137	1572.679968	0.11
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c21	2+	1180.030468	1180.030012	0.39
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c22	2+	1237.042578	1237.042287	0.24
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c23	2+	1310.576789	1310.576910	-0.09
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c24	2+	1339.087521	1339.087462	0.04
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c25	2+	1375.107454	1375.107595	-0.10
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c26	2+	1431.148110	1431.148133	-0.02
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c27	2+	1488.191533	1488.192074	-0.36
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c28	2+	1531.206156	1531.205760	0.26
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c30	2+	1625.246010	1625.245681	0.20
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c31	2+	1682.267474	1682.267173	0.18
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c32	2+	1732.303073	1732.303880	-0.47
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c33	2+	1760.312407	1760.311948	0.26
(isoD)$_3$ plus sulf- group	(isoD)$_3$	c34	2+	1803.828422	1803.828323	0.05
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	M	2+	2053.930686	2053.930374	0.15
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	M	3+	1368.282448	1368.282142	0.22
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	M	4+	1026.463655	1026.463443	0.21
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c14	1+	1689.725410	1689.726242	-0.49
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c28	2+	1590.731589	1590.731698	-0.07
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c31	2+	1741.792905	1741.791955	0.55
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c32	2+	1790.324318	1790.323223	0.61
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c33	2+	1819.837848	1819.837221	0.34
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c34	2+	1862.351058	1862.350961	0.05
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c34	3+	1241.901454	1241.901314	0.11
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c35	3+	1279.909504	1279.909256	0.19
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	c36	3+	1313.928009	1313.927791	0.17
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	z27	2+	1534.202451	1534.202150	0.20
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	z28	2+	1598.231740	1598.231544	0.12
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	z29	2+	1649.257008	1649.256773	0.14
(isoD)$_3$ plus BISA	(isoD)$_3$ plus BISA	z30	2+	1685.276986	1685.276681	0.18

Average Error (ppm):	0.05
Average Absolute Error (ppm):	0.21
Standard Deviation:	0.27

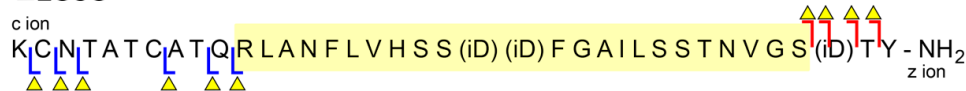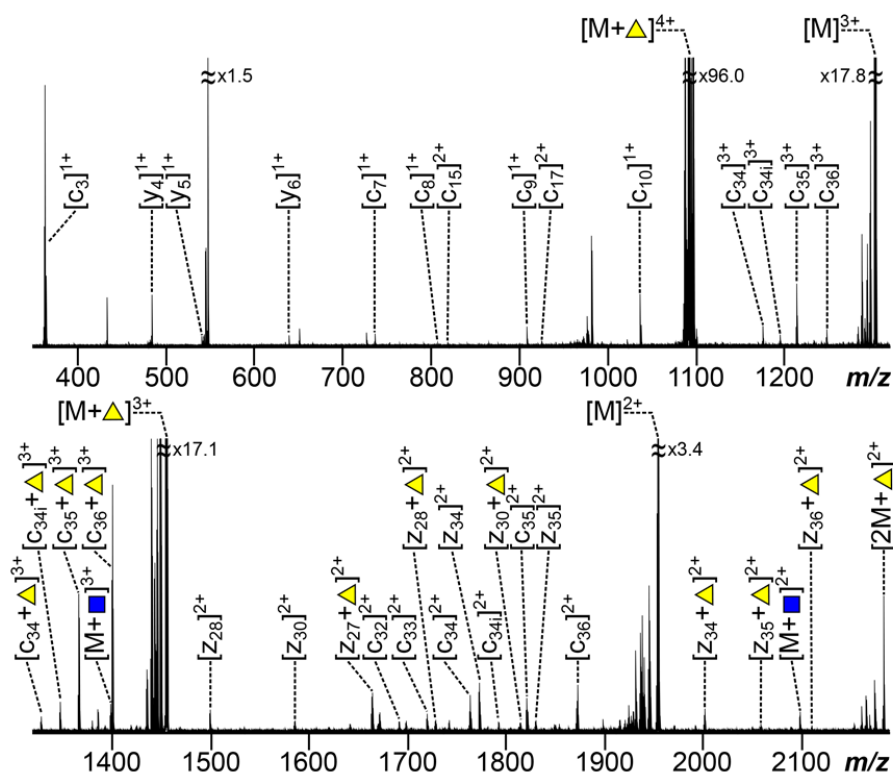**Table S3.13.** List of the assigned ECD fragments of the 4+ molecular ion of the mutant (isoD)$_3$hIAPP monomer unit with a BISA molecule.
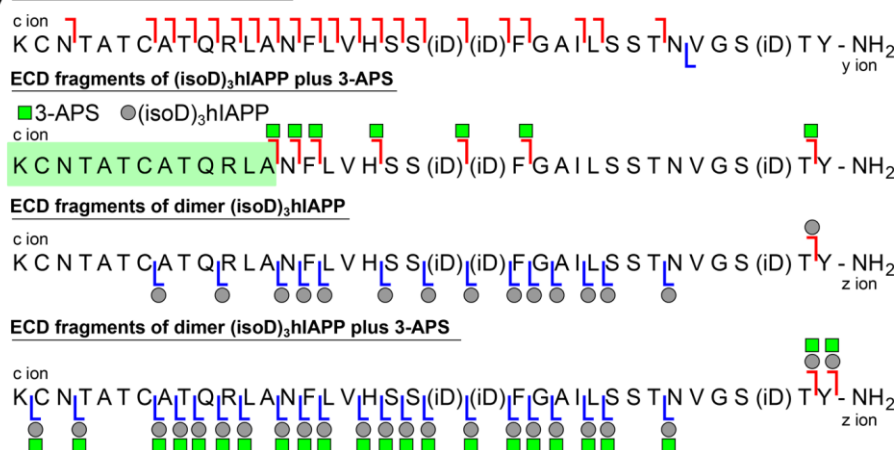
275
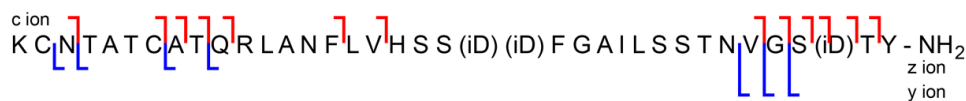
Figure S3.21. (A) The summarised ECD fragments of the 4+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and an EGCG molecule. The highlighted area (coloured in yellow) indicates the proposed interaction region between the monomer unit of the mutant (isoD)$_3$hIAPP and an EGCG molecule which is between Arg-11 and Ser-34 residues on the mutant (isoD)$_3$hIAPP molecule. (B) The ECD MS/MS spectrum of the 4+ charge state molecular ion of the mutant (isoD)$_3$hIAPP monomer unit and an EGCG molecule. The assigned fragments are listed in the Supplementary Information Table S3.14.

Assigned ECD fragments of the 4+ molecular ion of (isoD)$_3$hIAPP with EGCG

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
(isoD)$_3$ only	(isoD)$_3$	M	2+	1954.925496	1954.924942	0.28
(isoD)$_3$ only	(isoD)$_3$	M	3+	1302.279143	1302.279059	0.06
(isoD)$_3$ only	(isoD)$_3$	c3	1+	363.180901	363.180898	0.01
(isoD)$_3$ only	(isoD)$_3$	c7	1+	737.306908	737.306854	0.07
(isoD)$_3$ only	(isoD)$_3$	c8	1+	808.344022	808.344160	-0.17
(isoD)$_3$ only	(isoD)$_3$	c9	1+	909.391701	909.391865	-0.18
(isoD)$_3$ only	(isoD)$_3$	c10	1+	1037.450279	1037.450652	-0.36
(isoD)$_3$ only	(isoD)$_3$	c15	2+	819.895593	819.895799	-0.25
(isoD)$_3$ only	(isoD)$_3$	c17	2+	925.971832	925.972020	-0.20
(isoD)$_3$ only	(isoD)$_3$	c32	2+	1692.825783	1692.825730	0.03
(isoD)$_3$ only	(isoD)$_3$	c33	2+	1719.830087	1719.829403	0.40
(isoD)$_3$ only	(isoD)$_3$	c34	2+	1763.847504	1763.847939	-0.25
(isoD)$_3$ only	(isoD)$_3$	c34	3+	1176.234095	1176.234008	0.07
(isoD)$_3$ only	(isoD)$_3$	c34i	2+	1793.852388	1793.851868	0.29
(isoD)$_3$ only	(isoD)$_3$	c34i	3+	1195.569257	1195.569163	0.08
(isoD)$_3$ only	(isoD)$_3$	c35	2+	1821.864890	1821.863500	0.76
(isoD)$_3$ only	(isoD)$_3$	c35	3+	1214.242141	1214.241984	0.13
(isoD)$_3$ only	(isoD)$_3$	c36	2+	1872.388731	1872.388124	0.32
(isoD)$_3$ only	(isoD)$_3$	c36	3+	1248.927090	1248.927464	-0.30
(isoD)$_3$ only	(isoD)$_3$	y4	1+	484.203806	484.203809	-0.01
(isoD)$_3$ only	(isoD)$_3$	y5	1+	541.225270	541.225288	-0.03
(isoD)$_3$ only	(isoD)$_3$	y6	1+	640.293684	640.293720	-0.06
(isoD)$_3$ only	(isoD)$_3$	z28	2+	1500.232133	1500.232022	0.07
(isoD)$_3$ only	(isoD)$_3$	z30	2+	1586.775917	1586.774655	0.80
(isoD)$_3$ only	(isoD)$_3$	z34	2+	1772.836090	1772.836105	-0.01
(isoD)$_3$ only	(isoD)$_3$	z35	2+	1830.358982	1830.358487	0.27
(isoD)$_3$ plus EGCG ($C_{15}H_{13}O_6$)	(isoD)$_3$ plus EGCG ($C_{15}H_{13}O_6$)	M	2+	2098.957258	2098.957106	0.07
(isoD)$_3$ plus EGCG ($C_{15}H_{13}O_6$)	(isoD)$_3$ plus EGCG ($C_{15}H_{13}O_6$)	M	3+	1398.970505	1398.969616	0.64
(isoD)$_3$ plus EGCG	(isoD)$_3$ plus EGCG	M	2+	2184.469182	2184.468035	0.53
(isoD)$_3$ plus EGCG	(isoD)$_3$ plus EGCG	M	3+	1455.978516	1455.978501	0.01
(isoD)$_3$ plus EGCG	(isoD)$_3$ plus EGCG	M	4+	1091.482404	1091.482042	0.33
(isoD)$_3$ plus EGCG	(isoD)$_3$	c34	3+	1328.927284	1328.927366	-0.06
(isoD)$_3$ plus EGCG	(isoD)$_3$	c34i	3+	1348.262449	1348.262382	0.05
(isoD)$_3$ plus EGCG	(isoD)$_3$	c35	3+	1366.935462	1366.935379	0.06
(isoD)$_3$ plus EGCG	(isoD)$_3$	c36	3+	1401.288112	1401.288199	-0.06
(isoD)$_3$ plus EGCG	(isoD)$_3$	z27	2+	1664.239945	1664.239592	0.21
(isoD)$_3$ plus EGCG	(isoD)$_3$	z28	2+	1729.274622	1729.273583	0.60
(isoD)$_3$ plus EGCG	(isoD)$_3$	z30	2+	1814.311630	1814.311018	0.34
(isoD)$_3$ plus EGCG	(isoD)$_3$	z34	2+	2002.377420	2002.377122	0.15
(isoD)$_3$ plus EGCG	(isoD)$_3$	z35	2+	2060.401324	2060.400919	0.20
(isoD)$_3$ plus EGCG	(isoD)$_3$	z36	2+	2110.903378	2110.902434	0.45

Average Error (ppm):	0.13
Average Absolute Error (ppm):	0.22
Standard Deviation:	0.28

Table S3.14.   List of the assigned ECD fragments of the 4+ molecular ion of the mutant (isoD)$_3$hIAPP monomer unit with an EGCG molecule.
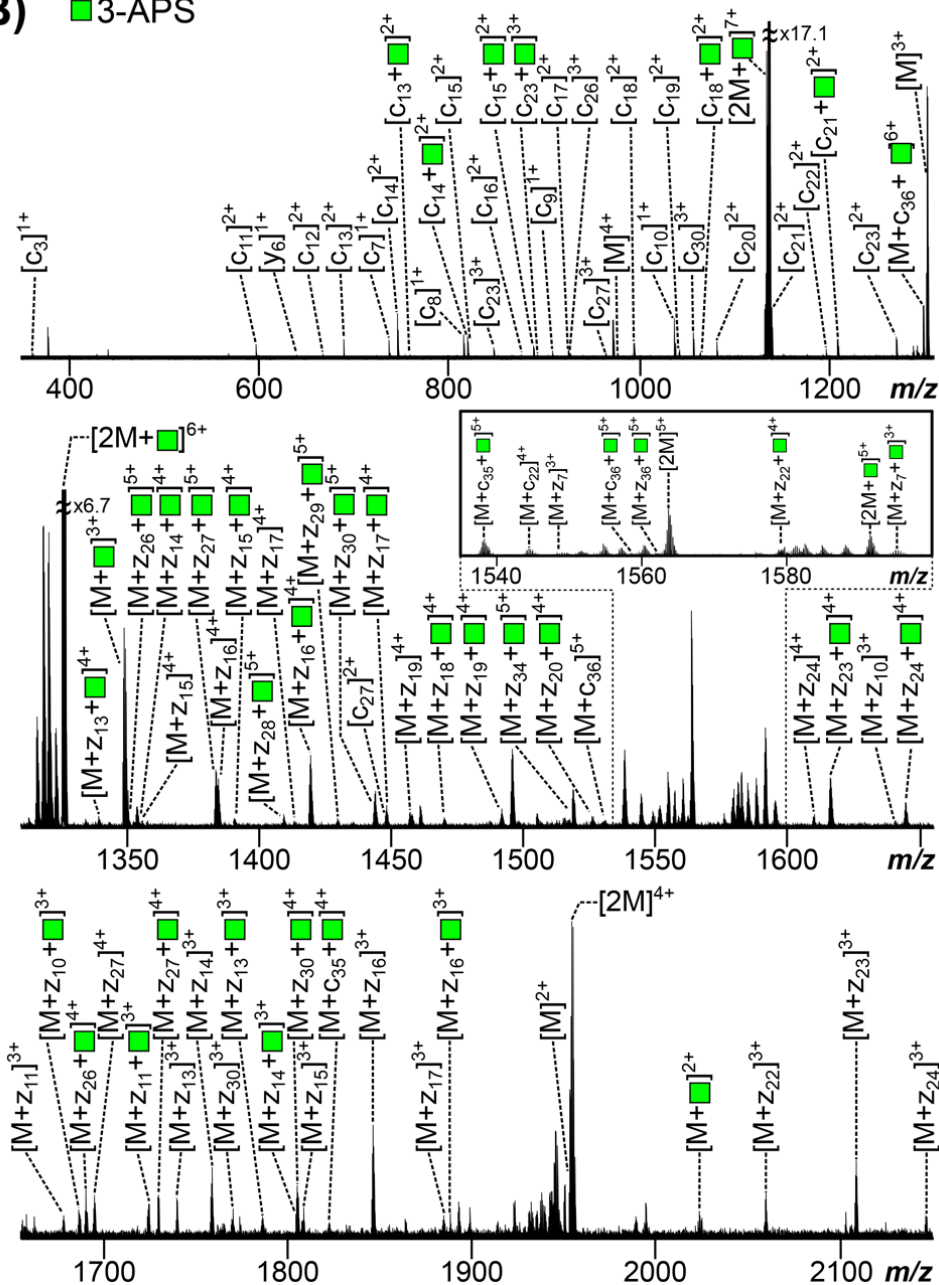
Figure S3.22. (A) The summarised ECD fragments of the 7+ charge state molecular ion of the mutant (isoD)$_3$hIAPP dimer with a 3-APS molecule. The highlight area (coloured in green) indicates the proposed non-interruption binding region, which does not interfere with the formation of the early oligomers between the mutant (isoD)$_3$hIAPP and a 3-APS molecule. (B) The ECD MS/MS spectrum of the 7+ charge state molecular ion of the mutant (isoD)$_3$hIAPP dimer with a 3-APS molecule. The assigned fragments are listed in the Supplementary Information Table S3.15.

Assigned ECD fragments of the 7+ molecular ion of the dimer $(isoD)_3$hIAPP with 3-APS

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
$(isoD)_3$ only	$(isoD)_3$	M	2+	1952.915077	1952.914286	0.41
$(isoD)_3$ only	$(isoD)_3$	M	3+	1302.279143	1302.279094	0.04
$(isoD)_3$ only	$(isoD)_3$	M	4+	976.961177	976.960828	0.36
$(isoD)_3$ only	$(isoD)_3$	c3	1+	362.173076	362.173076	0.00
$(isoD)_3$ only	$(isoD)_3$	c7	1+	737.306908	737.307001	-0.13
$(isoD)_3$ only	$(isoD)_3$	c8	1+	808.344022	808.344094	-0.09
$(isoD)_3$ only	$(isoD)_3$	c9	1+	909.391701	909.391927	-0.25
$(isoD)_3$ only	$(isoD)_3$	c10	1+	1037.450279	1037.450648	-0.36
$(isoD)_3$ only	$(isoD)_3$	c11	2+	597.279333	597.279446	-0.19
$(isoD)_3$ only	$(isoD)_3$	c12	2+	653.821365	653.821510	-0.22
$(isoD)_3$ only	$(isoD)_3$	c13	2+	689.339922	689.340098	-0.26
$(isoD)_3$ only	$(isoD)_3$	c14	2+	746.361386	746.361545	-0.21
$(isoD)_3$ only	$(isoD)_3$	c15	2+	819.895593	819.895700	-0.13
$(isoD)_3$ only	$(isoD)_3$	c16	2+	876.437625	876.437420	0.23
$(isoD)_3$ only	$(isoD)_3$	c17	2+	925.971832	925.972043	-0.23
$(isoD)_3$ only	$(isoD)_3$	c18	2+	994.501288	994.501674	-0.39
$(isoD)_3$ only	$(isoD)_3$	c19	2+	1038.017303	1038.016800	0.48
$(isoD)_3$ only	$(isoD)_3$	c20	2+	1081.533317	1081.533610	-0.27
$(isoD)_3$ only	$(isoD)_3$	c21	2+	1139.046789	1139.046754	0.03
$(isoD)_3$ only	$(isoD)_3$	c22	2+	1197.061634	1197.062459	-0.69
$(isoD)_3$ only	$(isoD)_3$	c23	2+	1270.094468	1270.095113	-0.51
$(isoD)_3$ only	$(isoD)_3$	c23	3+	847.065404	847.065535	-0.15
$(isoD)_3$ only	$(isoD)_3$	c26	3+	927.446285	927.446769	-0.52
$(isoD)_3$ only	$(isoD)_3$	c27	2+	1447.709221	1447.709981	-0.52
$(isoD)_3$ only	$(isoD)_3$	c27	3+	965.140973	965.141127	-0.16
$(isoD)_3$ only	$(isoD)_3$	c30	3+	1056.844885	1056.844954	-0.07
$(isoD)_3$ only	$(isoD)_3$	y6	1+	640.293684	640.293708	-0.04
$(isoD)_3$ Plus 3-APS	$(isoD)_3$ Plus 3-APS	M	2+	2023.432737	2023.432507	0.11
$(isoD)_3$ Plus 3-APS	$(isoD)_3$ Plus 3-APS	M	3+	1349.290917	1349.291263	-0.26
$(isoD)_3$ Plus 3-APS	$(isoD)_3$	c13	2+	758.855080	758.855677	-0.79
$(isoD)_3$ Plus 3-APS	$(isoD)_3$	c14	2+	815.876650	815.876544	0.13
$(isoD)_3$ Plus 3-APS	$(isoD)_3$	c15	2+	889.410751	889.410799	-0.05
$(isoD)_3$ Plus 3-APS	$(isoD)_3$	c18	2+	1064.016446	1064.016311	0.13
$(isoD)_3$ Plus 3-APS	$(isoD)_3$	c21	2+	1208.561947	1208.562326	-0.31
$(isoD)_3$ Plus 3-APS	$(isoD)_3$	c23	3+	893.408842	893.409180	-0.38
2 $(isoD)_3$	2 $(isoD)_3$	M	4+	1955.174545	1955.172652	0.97
2 $(isoD)_3$	2 $(isoD)_3$	M	5+	1563.738668	1563.737907	0.49
2 $(isoD)_3$	$(isoD)_3$	c36	5+	1530.924421	1530.924653	-0.15
2 $(isoD)_3$	$(isoD)_3$	z7	3+	1549.387825	1549.387353	0.30
2 $(isoD)_3$	$(isoD)_3$	z10	3+	1640.421710	1640.421490	0.13
2 $(isoD)_3$	$(isoD)_3$	z11	3+	1678.450526	1678.450445	0.05
2 $(isoD)_3$	$(isoD)_3$	z13	3+	1740.494307	1740.493867	0.25
2 $(isoD)_3$	$(isoD)_3$	z14	3+	1758.499052	1758.499582	-0.30
2 $(isoD)_3$	$(isoD)_3$	z15	3+	1807.856058	1807.856680	-0.34
2 $(isoD)_3$	$(isoD)_3$	z15	4+	1355.891906	1355.892142	-0.17
2 $(isoD)_3$	$(isoD)_3$	z16	3+	1845.864176	1845.864266	-0.05
2 $(isoD)_3$	$(isoD)_3$	z16	4+	1384.899255	1384.898978	0.20

(Table S3.15. Continue)

Sample	Fragment	Ion	Charge	Theoretical m/z	Experimental m/z	Error (ppm)
2 (isoD)$_3$	(isoD)$_3$	z17	3+	1884.540692	1884.540169	0.28
2 (isoD)$_3$	(isoD)$_3$	z17	4+	1413.404039	1413.404368	-0.23
2 (isoD)$_3$	(isoD)$_3$	z19	4+	1457.172015	1457.172096	-0.06
2 (isoD)$_3$	(isoD)$_3$	z22	3+	2059.300020	2059.299752	0.13
2 (isoD)$_3$	(isoD)$_3$	z22	4+	1544.474878	1544.474617	0.17
2 (isoD)$_3$	(isoD)$_3$	z23	3+	2108.656545	2108.656983	-0.21
2 (isoD)$_3$	(isoD)$_3$	z24	3+	2146.337153	2146.337462	-0.14
2 (isoD)$_3$	(isoD)$_3$	z24	4+	1609.502095	1609.501215	0.55
2 (isoD)$_3$	(isoD)$_3$	z27	4+	1694.558953	1694.558773	0.11
2 (isoD)$_3$	(isoD)$_3$	z30	4+	1770.346724	1770.346968	-0.14
2 (isoD)$_3$ Plus 3-APS	2 (isoD)$_3$ Plus 3-APS	M	5+	1591.745099	1591.744241	0.54
2 (isoD)$_3$ Plus 3-APS	2 (isoD)$_3$ Plus 3-APS	M	6+	1326.120118	1326.119521	0.45
2 (isoD)$_3$ Plus 3-APS	2 (isoD)$_3$ Plus 3-APS	M	7+	1136.817475	1136.817131	0.30
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	c35	4+	1823.400301	1823.399628	0.37
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	c35	5+	1538.519733	1538.519106	0.41
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	c36	5+	1558.528872	1558.528209	0.43
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	c36	6+	1298.774916	1298.774685	0.18
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z7	3+	1595.061180	1595.060915	0.17
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z10	3+	1686.765105	1686.765120	-0.01
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z11	3+	1725.129811	1725.128623	0.69
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z13	3+	1786.503563	1786.503257	0.17
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z13	4+	1339.626976	1339.626167	0.60
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z14	3+	1805.174778	1805.174410	0.20
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z14	4+	1354.132903	1354.132860	0.03
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z15	4+	1390.649457	1390.649362	0.07
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z16	3+	1892.541778	1892.540359	0.75
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z16	4+	1419.656770	1419.656658	0.08
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z17	4+	1448.413513	1448.413450	0.04
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z18	4+	1470.171524	1470.171512	0.01
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z19	4+	1491.678952	1491.678571	0.26
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z20	4+	1526.444813	1526.443930	0.58
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z22	4+	1579.232403	1579.232101	0.19
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z23	4+	1615.999522	1615.998613	0.56
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z24	4+	1644.510255	1644.509304	0.58
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z26	4+	1690.039299	1690.040025	-0.43
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z26	5+	1352.633904	1352.633246	0.49
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z27	4+	1729.567181	1729.567898	-0.41
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z27	5+	1384.054585	1384.054103	0.35
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z28	5+	1409.465843	1409.464757	0.77
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z29	5+	1429.875846	1429.875514	0.23
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z30	4+	1805.354790	1805.354166	0.35
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z30	5+	1444.484162	1444.484438	-0.19
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z34	5+	1518.508553	1518.508742	-0.12
2 (isoD)$_3$ Plus 3-APS	(isoD)$_3$	z36	5+	1562.618813	1562.619037	-0.14

Average Error (ppm):	0.06
Average Absolute Error (ppm):	0.28
Standard Deviation:	0.35

Table S3.15.   List of the assigned ECD fragments of the 7+ molecular ion of the mutant (isoD)$_3$hIAPP dimer with a 3-APS molecule.

A)

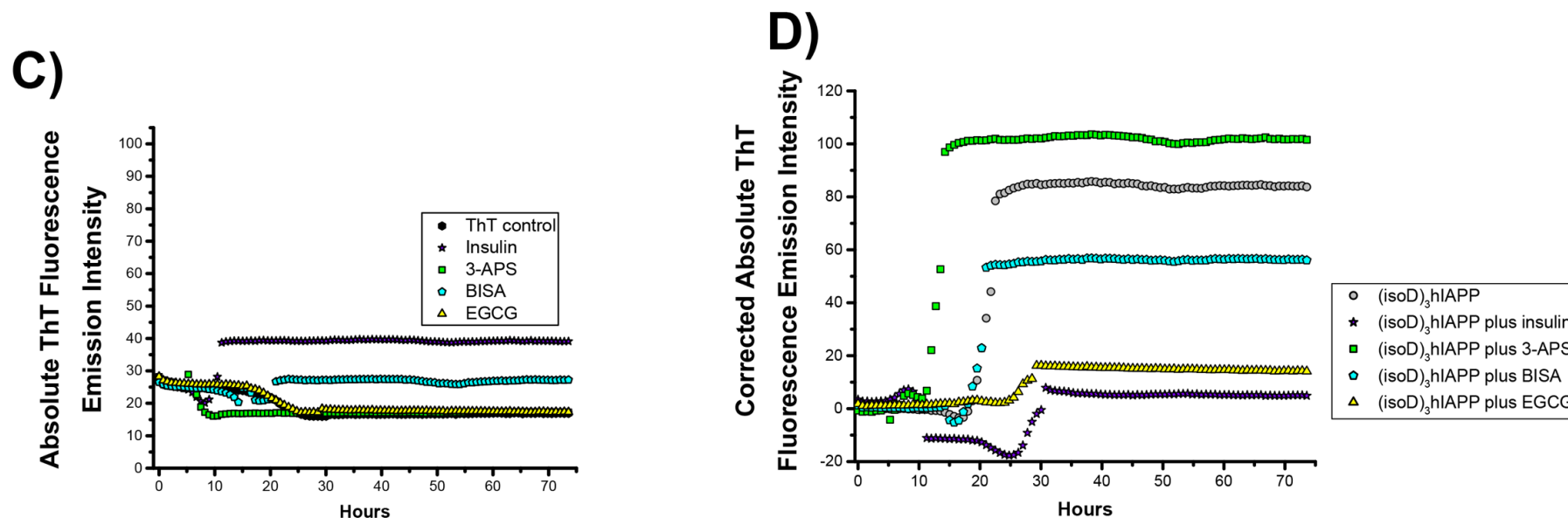| Model	Boltzmann Fitting				
Sample	(isoD)$_3$hIAPP only	(isoD)$_3$hIAPP + Insulin	(isoD)$_3$hIAPP + 3-APS	(isoD)$_3$hIAPP + BISA	(isoD)$_3$hIAPP + EGCG
Calculated lag phase	17.9	24.5	11.0	18.5	28.1

B)
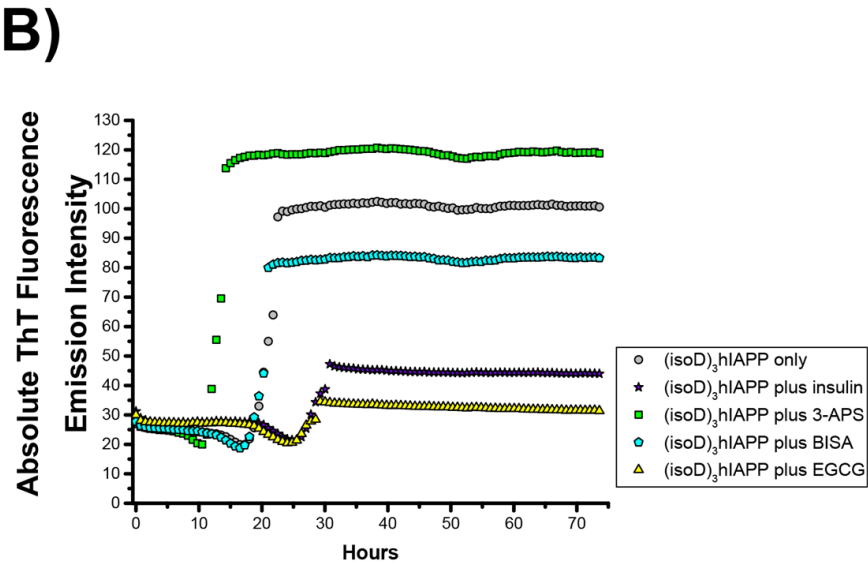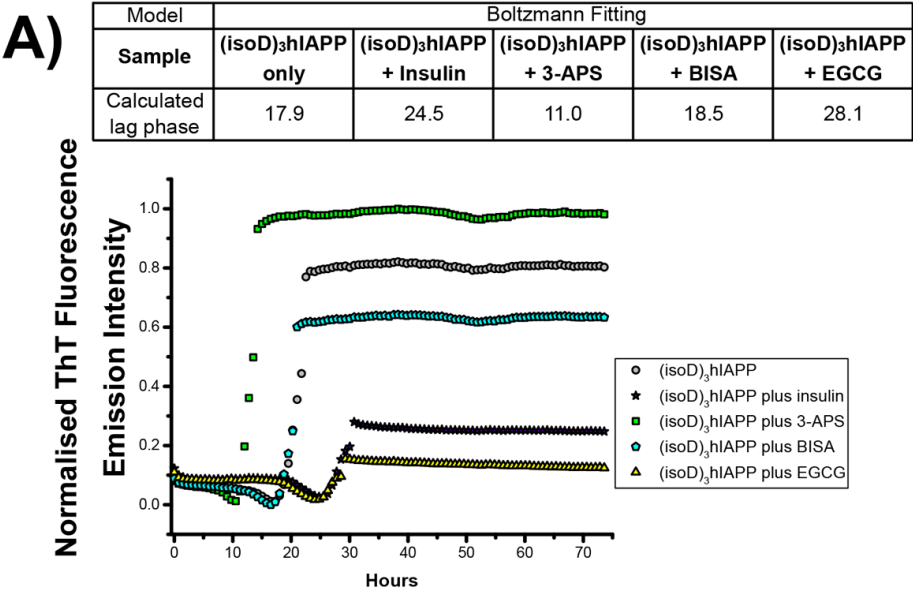
(Figure S3.23. Continue)

Figure S3.23. (A) The normalised ThT fluorescence emission of the 50 µM solution of the mutant $(isoD)_3hIAPP$ only (grey circle), and mixtures of the 50 µM solutions of mutant $(isoD)_3hIAPP$ with the addition of equal molar of insulin (purple star), 3-APS (green rectangle), BISA (cyan pentagon), and EGCG (yellow triangle) solutions. Each emission was fitted with a Boltzmann curve and the lag phase of each solution was calculated.[68-69] (B) The absolute ThT fluorescence emission of the mutant $(isoD)_3hIAPP$ (grey circles), and the mutant $(isoD)_3hIAPP$ mixed with equal molar of insulin (purple star), 3-APS (green rectangle), BISA (cyan pentagon), and EGCG (yellow triangle) solutions. (C) The absolute

ThT fluorescence emission of the ThT control (black hexagon), insulin, 3-APS, BISA, and EGCG solutions.   (D) The corrected absolute ThT fluorescence emission of the mutant $(isoD)_3$hIAPP only, and the mutant $(isoD)_3$hIAPP mixed with equal molar of insulin, 3-APS, BISA, and EGCG solutions.   The corrected absolute ThT was calculated by using the absolute ThT fluorescence emission of the mutant $(isoD)_3$hIAPP mixed with various binding molecules (see Figure S3.21A) minus the absolute ThT fluorescence emission of the binding molecules alone (see Figure S3.21B).

	(isoD)₃hIAPP only	(isoD)₃hIAPP + Insulin	(isoD)₃hIAPP + 3-APS	(isoD)₃hIAPP + BISA	(isoD)₃hIAPP + EGCG	Insulin in (isoD)₃hIAPP with insulin
% Change	-53.84	-36.49	-28.29	-39.77	-63.66	-13.93
p Value	1.22E-06	1.36E-07	5.27E-05	3.68E-08	3.12E-08	1.32E-05
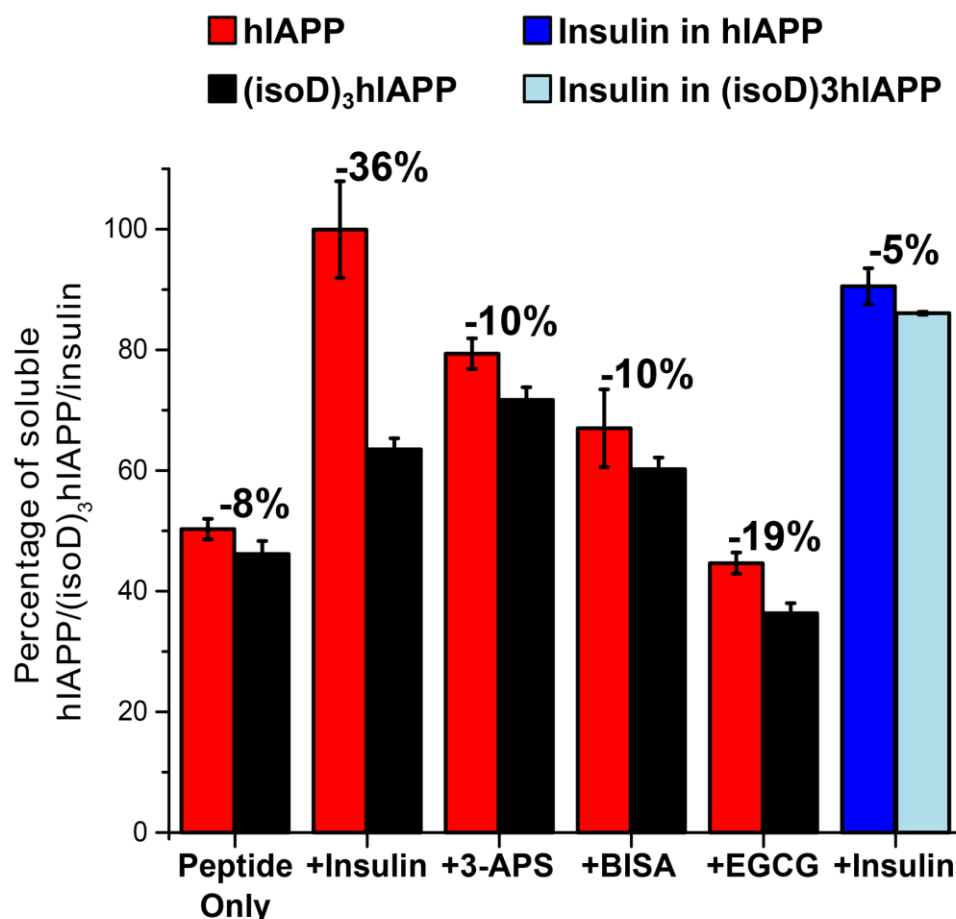
Figure S3.24. The percentage of the soluble hIAPP (red bar), (isoD)₃hIAPP (black bar), and insulin (blue bar) present in the 10 µM incubated solutions which has been mixed with the binding molecules in equal molar and incubated for one week. The percentage of the soluble mutant (isoD)₃hIAPP and insulin in the fresh and incubated solutions were tested against two-tailed t-tests and the t-test p value less than 0.05 (red font) which indicates there is a significant different the fresh and incubated solutions. The percentage different of the soluble amyloid proteins and insulin between the wild-type hIAPP and the mutant (isoD)₃hIAPP solutions were calculated and labelled in the bar chart.
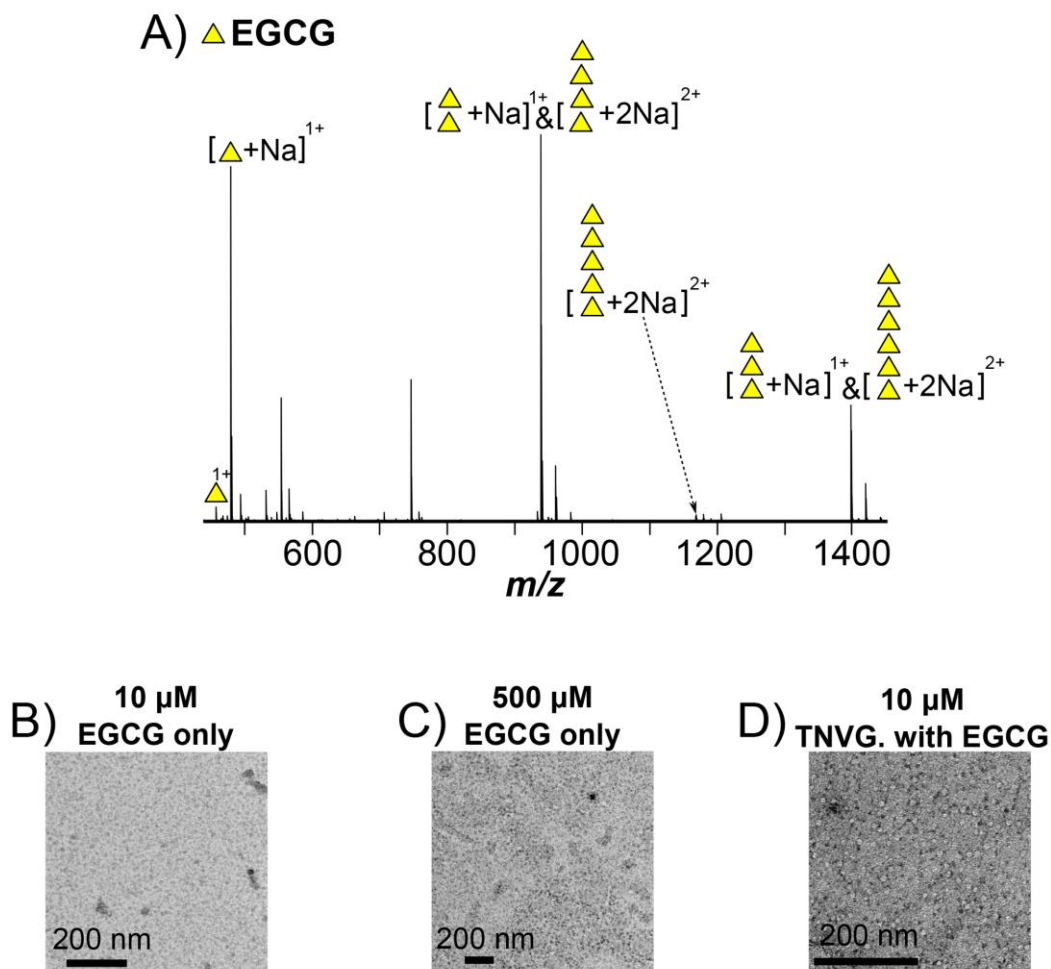
Figure S3.25. (A) The MS spectrum of the 10 µM aqueous solution of EGCG only. The EGCG solution was diluted to 50 µM and vortexed for half an hour, then the solution was further diluted to 10 µM and centrifuged at 14,000 rpm before MS analysis. Early oligomers, from monomer to hexamer, of EGCG were observed in the MS spectrum which indicates EGCG can rapidly aggregate in an aqueous solution. The TEM images of one week incubated solutions of (B) 10 µM EGCG solution only, (C) 500 µM EGCG solution only, and (D) 10 µM ^{30}TNVGSNTY^{37}-NH$_2$ peptide mixed with equal molar of EGCG solution. Circular spots are observed in the TEM images with the addition of EGCG solution, the average diameter of the circular spots in the 10 µM EGCG solution is 09 nm, the average diameter of the spots in the 500 µM EGCG solution is 12 nm, and the average diameter of the circular spots in the TEM image of the mixture of equal

molar of ^{30}TNVGSNTY^{37}-NH$_2$ peptide and EGCG solution is 16 nm. The TEM

images here demonstrate EGCG aggregates in the aqueous solution.

Chapter 4 The Next Dimension in Proteomics

This chapter demonstrates the application of two dimensional mass spectrometry (2DMS) to proteomic studies ranging from standard protein digests to a whole cell lysate digest without prior separation using online liquid chromatography (LC). The development of 2DMS aims to provide an alternative method for proteomic studies, and hopefully it can further apply to study fragile interaction between biomolecules in a complex proteomic sample through eliminating the limitations in online LC system.

The results presented herein were predominantly carried out by the thesis Author, aside from the in-house software was written by Dr. Christopher A. Wootton, Tomos E. Morgan, and Remy Gavard from Prof. Peter O'Connor and Dr. Mark P. Barrow groups.

A manuscript entitled "The Next Dimension in Proteomics" by Yuko P. Y. Lam, Christopher A. Wootton, Cookson K. C. Chiu, Tomos E. Morgan, Remy Gavard, Maria A. van Agthoven, M. Li, Mark P. Barrow, and Peter B. O'Connor has been prepared for submission to Nature Methods based on the results presented in this chapter.

4.1 Abstract

Reversed phase liquid chromatography (RP-LC) is an essential tool for the proteomic studies due to its powerful separation of peptides and proteins prior to MS analysis; however, LC separation of extremely complex samples with widely varying properties suffers from the bias of the chromatography columns applied in LC separation and incompatibility with certain separation conditions such as solvent compatibility, pH, and column chemistry. An alternative analytical method which is able to handle complex proteomic samples and high flexibility in sample conditions is required. Herein, we apply two dimensional mass spectrometry (2DMS) to study proteomic samples ranging from a standard protein digest to a whole cell lysate digest. In the 2DMS experiments, we show that it is possible to obtain detailed sequence information from peptides without online LC separation. We then compared the results obtained from the 2DMS to the LC tandem MS data and show that 2DMS can achieve similar results to LC MS/MS for identified proteins in the proteomic samples shown. Furthermore, our data show that the 2DMS experiments tend to be more adept at analysing and identifying particularly hydrophilic, basic, or short peptides, which is complementary to those being assigned in LC tandem MS. These results demonstrate that 2DMS provides an alternative technique for future proteomics.

4.2 Introduction

Proteomic studies focus on analysing the proteins generated by living organisms. Mass spectrometry (MS) is the leading analytical technique for large scale proteomic studies due to its high sensitivity, throughput, acquisition speed, and wide dynamic range.[1-3] In proteomic studies, MS is used to measure the masses of proteins and peptides, then fragment them using serial tandem mass spectrometry to provide sequence information and identify the biomolecules present.[2-4] Proteomic studies are currently involved in searching for biomarkers of diseases,[5] quantifying critical proteins and relative PTMs changes,[6, 7] observing structural changes of proteins,[8, 9] as well as monitoring the interactions between potential inhibitors and proteins in living organisms.[2, 10] These studies contribute to prevention and treatment of the diseases, such as cancer,[11] neurodegenerative diseases,[12] bacterial infected diseases.[13] Currently, the majority of proteomics is conducted under the bottom-up regime where proteins are chemically or enzymatically digested from proteins and is the most common method applied in MS-based proteomics.[2, 14]

MS can detect many ions simultaneously, but it is challenging to handle tens-of-thousands of peptides from a digested sample in a single injection. Thus, peptides are normally separated using high performance liquid chromatography (HPLC) or nano-liquid chromatography (nLC) prior to online MS and MS/MS analysis.[2, 3] Yeast from *Saccharomyces cerevisiae* is commonly used as a model system to evaluate the capabilities of newly developed proteomic systems.[15-20] With early methods, the yeast proteome was first separated using two dimensional polyacrylamide gels (2D-PAGE), the proteins were then digested into peptides and further separated by a C18 reversed phase liquid chromatography (RP-LC) column.[16] The separated peptides were analysed using tandem MS (MS/MS) and 90 proteins were assigned in the experiment. Several years later, Washburn *et al.* demonstrated the application of a two dimensional

liquid chromatography (2D-LC) system to study yeast proteome, which assigned more than 1,400 yeast proteins in the experiment.[17]   With the use of multiple orthogonal LC columns, the number of identified proteins, as well as the dynamic range of the proteome can be enhanced.[20]   Recently, Hebert *et al.* showed over 4,000 proteins were assigned in a 70-minute C18 RP-LC MS/MS experiment, showing a significant improvement in the rate of protein assignments per minute.[18, 21]   Rapid development of multiple dimensional LC MS/MS has led to faster and deeper analysis of proteomes, but LC separation, especially low pH RP-LC, becomes unavoidable in current proteomic studies which results in limiting the study of species retained tightly in RP columns (very hydrophobic peptides), species unable to be retained (very hydrophilic peptides), as well as fragile biomolecules and peptide-ligand species due to the conditions of the solvents used in RP-LC systems.

Study of non-covalent interactions between biomolecules is a critical research area for development of therapeutics.[22, 23]   Non-covalent interactions between biomolecules are generally studied by *in vitro* experiments using direct MS infusion with suitable physiological buffers, such as ammonium acetate.[24-26]   Previous studies showed that hydrogen-deuterium exchange (HDX) experiments could preserve the non-covalent interactions between biomolecules while using RP-LC separation prior to MS analysis.[27, 28]   The generation of complicated MS spectra, however, limits the application of HDX experiments in *in vivo* studies.   Thus, a new analytical technique that can provide a higher flexibility in the choices of LC column and the solvent system is required for future proteomics, in which the non-covalent interactions between biomolecules can be preserved.

Two dimensional mass spectrometry (2DMS) was first introduced by Pfandler *et al.* in 1987 and further developed and applied by Ross *et al.* in 1993 using an FTICR mass analyzer.[29, 30]   In 2DMS, a modulating excitation or de-excitation pulse is used

to alter the positions of ions inside the FTICR cell, ions that are modulated to overlap with the position of the fragmentation zone are dissociated with the applied radius-dependent in-cell MS/MS technique, such as infrared multiphoton dissociation (IRMPD) or electron capture dissociation (ECD).[31, 32] Since the fragment ions are modulated with the same frequency as the precursor ion, the modulation frequency can be used to correlate fragments with precursors. In 2DMS, samples are directly infused into the MS which allows a broader selection of sample preparation methods, especially the selection of sample buffers. 2DMS is a data-independent acquisition technique in that all precursor ions are fragmented during the experiment, and a two dimensional mass spectrum, showing all precursor and fragment ions, is obtained (see Supplementary Information, Figure S4.1).[31] Compared to the traditional MS/MS, 2DMS does not require isolation of a precursor ion prior to fragmentation and detection, instead utilising the frequencies of ions in the second dimension to separate the peaks which provides an advantage in the application to complex proteomic samples as discussed below.

2DMS is newly applied to proteomic studies, including both top-down and bottom-up proteomics.[33] Previous studies showed that similar top-down and bottom-up fragments were obtained from both quadrupole precursor isolated MS/MS and 2DMS.[34-36] Peptides from digested collagen showed that 2DMS can differentiate and fragment two precursor ions with similar mass-to-charge (*m/z*) ratios, which, in general, cannot be isolated and fragmented individually by quadrupole mass isolation.[37] Peptide assignment from the 2DMS spectra of digested proteins, such as cytochrome c, calmodulin, and collagen, also showed that 2DMS has the potential for complex proteomic studies without LC separation.[33, 34, 37, 38]

Herein, we have applied a mixture of digested standard proteins and a digested yeast lysate sample to evaluate the performance of 2DMS by comparing the data to

proteomic results obtained from a standard nano-LC (nLC) MS/MS approach. Various sample preparation and 2DMS data acquisition methods were used in the mixture of digested standard proteins in order to determine the most suitable workflow for a complex proteomic study. The optimized method was further applied to the digested yeast sample to demonstrate the ability of 2DMS in proteomic studies. From the results, we showed that 2DMS could identify a significant number of peptides in a digested proteomic sample, in which the number of assigned peptides was comparable to the nLC MS/MS data. Furthermore, the identities of proteins and peptides observed in the 2DMS experiments were complementary to the results from nLC MS/MS experiments, thus a better coverage in proteome as well as protein sequence can be obtained through a combination of methods. The solvent and column constraints in LC systems mentioned in above were also shown to be eliminated in our 2DMS experiments without affecting the quality of data. The peptides that were exclusively assigned in the 2DMS experiments suggested this newly developed method is suitable for complex proteomic studies and associating the data obtained from 2DMS with nLC MS/MS results provides a new capability in future proteomic studies.

4.3 Experimental Section

Sample Preparation

Standard proteins digestion.   Standard proteins, including bovine serum albumin (BSA; Sigma A7030), chicken lysozyme (CLZ; Sigma L6876), human haemoglobin (HHA; Sigma H7379), human lysozyme (HLZ; Sigma L1667), and human serum albumin (HSA; A3782), were purchased from Sigma-Aldrich Company Ltd (Dorset, UK).   The proteins were individually dissolved in 100 mM ammonium bicarbonate (ABC) solution to 1 µg/µL.   Protein disulphide bonds were then reduced using 50 mM dithiothreitol (DTT) for 30 minutes at 60°C, followed by alkylation using 100 mM iodoacetamide (IAA) in the dark at room temperature for 60 minutes.   Bovine trypsin solution was added into the alkylated solutions at 25:1 ratio (protein : trypsin) and the mixtures were incubated at 37°C for 16 hours.   500 µL was taken from each digested sample and combined to generate the "5 digested proteins" (5CP) solution.

Yeast cell lysis and digestion.   0.1 g of yeast from *Saccharomyces cerevisiae* (Sigma-Aldrich Company Ltd; Dorset, U.K.) was dissolved in 50 mL sodium dodecyl sulfate (SDS) lysis buffer containing 50 mM TRIS, 1% SDS, 2% glycerol, 50 mM DTT, and 5 mM ethylenediaminetetraacetic acid (EDTA), the pH of the lysis buffer was then adjusted to pH 7.5 using hydrochloric acid (HCl).   The solution was incubated at 95°C for 5 minutes and centrifuged at 5,000 rpm, 4°C for 30 minutes.   The supernatant was collected, mixed with 5-fold of iced acetone solution, and kept at -20°C in a freezer for 2 hours to isolate the proteins from the lysate.   The iced solution was then centrifuged at 5,000 rpm, 4°C for 30 minutes to separate the supernatant from the precipitated proteins (in pellet form), the protein pellet was then isolated.   6 M urea was used to dissolve the pellet, the amount of protein (µg/µL) in the solution was quantified by Bradford reagent assay and then diluted to 1 µg/µL using 100 mM ABC.   The sample was then digested using the same protocol mentioned above for standard proteins.

Peptide purification using solid phase extraction (SPE) cartridges. SPE cartridges were used to enrich peptides with various physical properties based on the principles of LC separation.[39] C18 RP (Thermo Fisher Scientific; Loughborough, UK), hydrophilic interaction chromatography (HILIC; Macherey-Nagel GmbH & Co. KG; Düren, Germany), and strong anion exchange (SAX; Thermo Fisher Scientific; Loughborough, UK) SPE cartridges were used to enrich the digested 5CP and yeast samples prior to MS analysis. 1 mg of 5CP and yeast samples were loaded onto each SPE cartridge. The pH conditions for the experiments, as well as the loading and elution buffers for each SPE cartridge are summarized in the Supplementary Information Table S4.1 and S4.2. After passing through the SPE cartridges, all samples were dried down using a Savant SPD121D SpeedVac concentrator (Thermo Fisher Scientific; Loughborough, UK) and then re-dissolved in purified H_2O from a Milli-Q purification system (Direct-Q® 3 UV System, Millipore Corporation, USA) to 1 µg/µL.

Peptide fractionation using solid phase extraction (SPE) cartridges. 2.5 mg of 5CP was loaded onto a C18 RP and HILIC SPE cartridges at pH 2 and pH 6.8 respectively. 9 fractions were collected from each SPE cartridge extraction by varying the percentage of acetonitrile (ACN) in the elution buffers (see Supplementary Information Figure S4.2A and S4.2B). Similarly, 5 mg of yeast was loaded onto a C18 RP SPE cartridge at pH 2 and 16 fractions were collected by increasing the percentage of ACN in the elution buffers (see Supplementary Information, Figure S4.2C). All individual fractions of 5CP and yeast were then dried down using a Savant SPD121D SpeedVac concentrator and re-dissolved in 300 µL of H_2O (~ 1 µg/µL with an assumption of peptides were evenly distributed across fractions).

Data Acquisition.

Data Acquisition of FTICR MS and 2DMS spectra.   All samples were diluted 4-fold to 0.25 µg/µL to a final solvent ratio of 75:25:0.1 water:acetonitrile:formic acid (H_2O:ACN:FA).   The samples were then centrifuged at 13,000 rpm for 10 minutes prior to MS and 2DMS analysis.   For FTICR MS and 2DMS experiments, samples were acquired using nano-electrospray ionization (nESI) with a capillary voltage of 0.6 – 1.0 kV and a source temperature of 180°C.   nESI glass capillaries (World Precision Instruments, England) were pulled to ~ 1 µm internal diameter (i.d.) using a Sutter P-97 capillary puller instrument (One Digital Drive, Novato, USA).   Samples were loaded onto the glass capillaries and a nickel chromium wire (Jacobs, USA) was inserted to provide the electric potential connection for ionization.

All mass spectra were acquired on a 12 tesla solariX FTICR MS (Bruker Daltonics, Bremen, Germany) in positive ionization mode.   Sample ions were first accumulated from 0.02 – 0.1 seconds in a hexapole collision cell before being directed to an infinity ICR cell for excitation and detection.[40]

Standard FTICR MS spectra were acquired with 4 mega-word (M) data-points (32 bits) over a mass range of *m/z* 147 – 3,000 to produce 1.67 second transients.   Different acquisition parameters were applied to the 2DMS spectra in order to optimize analytical performance.   In brief, 2DMS spectra with fragment *m/z* dimension of 128 k/256 k/1 M/2 M data-points over a mass range of *m/z* 147 – 3,000 and precursor *m/z* dimension (increments of t_1) of 4 k/ 8 k scan lines over a mass range of *m/z* 328 – 3,000 were acquired; the acquisition details of each 2DMS spectrum are summarized in Supplementary Information Table S4.1 and S4.2.   A 25 W continuous-wave CO_2 infrared laser was used for infrared multiphoton dissociation (IRMPD) in the 2DMS experiments, the pulse length and the power of the laser was set to 0.45 µs and 12.5 W respectively.
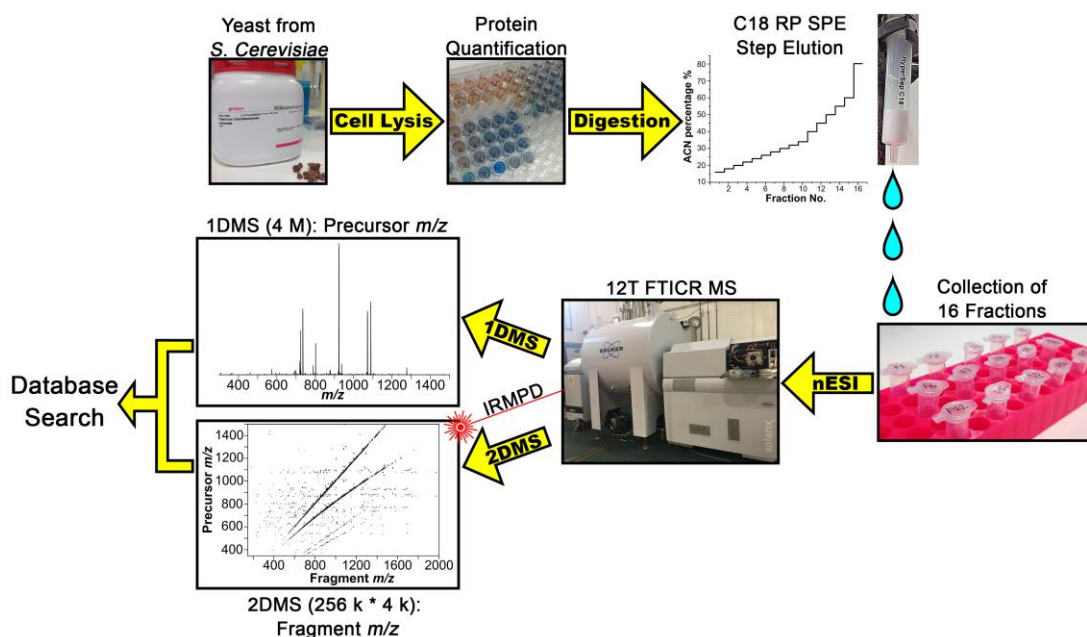
Figure 4.1. A summarised workflow of the sample preparation as well as 2DMS data acquisition of the offline fractionated yeast samples from *S. Cerevisiae*.

Data Acquisition of nano-Liquid Chromatography tandem MS (nLC MS/MS). An 18 cm in-house packed C18 RP nLC column (75 µm internal diameter) was connected to an EASY-nLC II system (Thermo Fisher Scientific; Loughborough, UK) to separate the peptides of BSA, 5CP, and yeast samples. 0.06 µg and 0.2 µg of BSA and 5CP samples respectively were loaded onto the C18 RP column, and a 30-minute effective gradient from 5% to 30% ACN in a 90-minute experiment was used (see Supplementary Information, Figure S4.3A). 1.5 µg of yeast sample was loaded onto the C18 RP column, and an extended 120-minute effective gradient from 5% to 30% ACN in a 180-minute runtime experiment was used (see Supplementary Information, Figure S4.3B).

The MS spectra of nLC MS/MS were acquired with 1 M data-points to 0.42 second transients. The eluted peptides from nLC were sprayed into the FTICR MS, the ions were accumulated in the hexapole for 0.3 s and transferred to the infinity cell for

detection.   Auto MS/MS functionality was used to fragment the ions which contained charge states ranging from 1+ to 5+ as well as peak intensities with threshold higher than 1E6.   The energies applied in collisionally activated dissociation (CAD) for peptide fragmentation were automatically determined according to the masses and charges of the ions (see Supplementary Information, Table S4.3).   Ions that contained the same *m/z* ratio (±10 ppm) were excluded from CAD MS/MS automatically for 1 minute during the experiment to avoid redundant multiple MS/MS events in a similar fashion to AL Rocha *et al.*[41]

Data Processing.

One dimensional (1D) MS spectra.   The 1D FTICR MS spectra collected for determination of precursor ions were analysed using Bruker DataAnalysis 4.2 software (Bruker Daltonics, Bremen, Germany).   Four peptides with known *m/z* ratio (for BSA and 5CP samples) were used to internal calibrate each spectrum.   The Fourier transform mass spectrometry (FTMS) peak picking algorithm provided in DataAnalysis 4.2 software was applied to generate the precursor ion peak list which included the peaks with a signal-to-noise (S/N) ratio higher than 10 as well as the charge state determination of each peak.

2DMS spectra.   The 2DMS spectra were first processed with the Spectrometry Processing Innovative Kernel (SPIKE) software on the high performing computers cluster, Tinis, at the University of Warwick and the urQRD denoising rank was set at 15.[42]   The processed data were then loaded onto an in-house T2D software to visualize and process the 2DMS spectra, the entire 2DMS spectrum was then inversely Fourier transformed and processed, line-by-line to produce transients for export to binary file format for Bruker programs (.ser files).   The transients (for the fragment *m/z* dimension) were reduced by half and zero-filled twice using FTMS Processing software

(Bruker Daltonics, Bremen, Germany) in order to improve the peak shape.[43]   The processed 2DMS data were imported into Bruker DataAnalysis 4.2 software and a scan with known *m/*z fragments was internally calibrated, the calibration constants were then applied to the other scans in the same 2DMS spectrum.   The peaks with S/N ratio over 5 in each scan line (precursor *m/z* dimension) were picked and a calibrated peak list was generated.

Mascot database search.   The peaks generated from 1DMS spectra (precursor ions) were combined with the 2DMS peak list (fragment ions) to produce the files in .mgf format for mascot database search.   The files were submitted to MS/MS Ions Search on MASCOT server with carbamidomethyl (C) set as fixed modification, and deamidated (NQ), oxidation (HMW), as well as phospho (STY) set as variable modifications.   In the MS/MS search, the missed cleavage was set to 1, the peptide tolerance was set to 1 ppm (for 2DMS BSA and 5CP)/ 8 ppm (for 2DMS yeast and nLC MS/MS), and the MS/MS tolerance was set to 8 ppm.   The 2DMS spectra were then manually matched with the shortlisted peptides assigned by MASCOT as the general peptide scoring system used in MASCOT could not be fully applied to 2DMS data (see Supplementary Information, Figure S4.4).   The hydrophobicity and isoelectric point (pI) of peptides were calculated using the Sequence Specific Retention Calculator[44] and the Compute pI/MW tool.[45]

4.4 Results and Discussion

Comparison of 2DMS and nLC MS/MS using digested BSA. Digested BSA provides a standard proteomic sample to evaluate the performance of different systems.[46, 47]   Herein, digested BSA was used to evaluate 2DMS for proteomic applications by comparing with the results obtained from a standard nLC MS/MS obtained on the same instrument.   A 2DMS spectrum, with 256 k * 4 k data-points (fragment m/z dimension * precursor m/z dimension), was obtained (see Figure 4.2A) and displayed as a contour plot of all peaks with an absolute intensity greater than 1E7 which is ~ 300 S/N ratio at m/z 600.   The resolving power (full width half maximum) of the fragment m/z and the precursor m/z dimensions were ~ 25 k and ~ 1 k at m/z 449.8 respectively (see Figure 4.2B).   The 1 k resolving power of the precursor m/z dimension was not sufficient to baseline-resolve the 2+ charge state ion at m/z 449.8 in a single dimension; however, with an additional fragment m/z dimension the contour peaks in Figure 4.2B were completely resolved.   The results indicate that the total resolving power in 2DMS is higher than a 1DMS dimension, demonstrating the capability of 2DMS to resolve peaks with a small m/z ratio difference in a complex proteomic sample which can't be isolated in a single MS dimension.

In the 2DMS spectrum, the auto-correlation line (m/z precursor = m/z fragment) represents all the precursor ions fragmented in the 2DMS experiment (see Figure 4.2C). The precursor ions (on the auto-correlation line) of digested BSA in the 2DMS spectrum were similar to its 4 M data-points 1DMS spectrum (see Supplementary Information, Figure S4.5).   Thus, the peaks in the 1DMS spectrum could be used for the precursor ion masses in database searches to improve calibration and 1D resolving power.
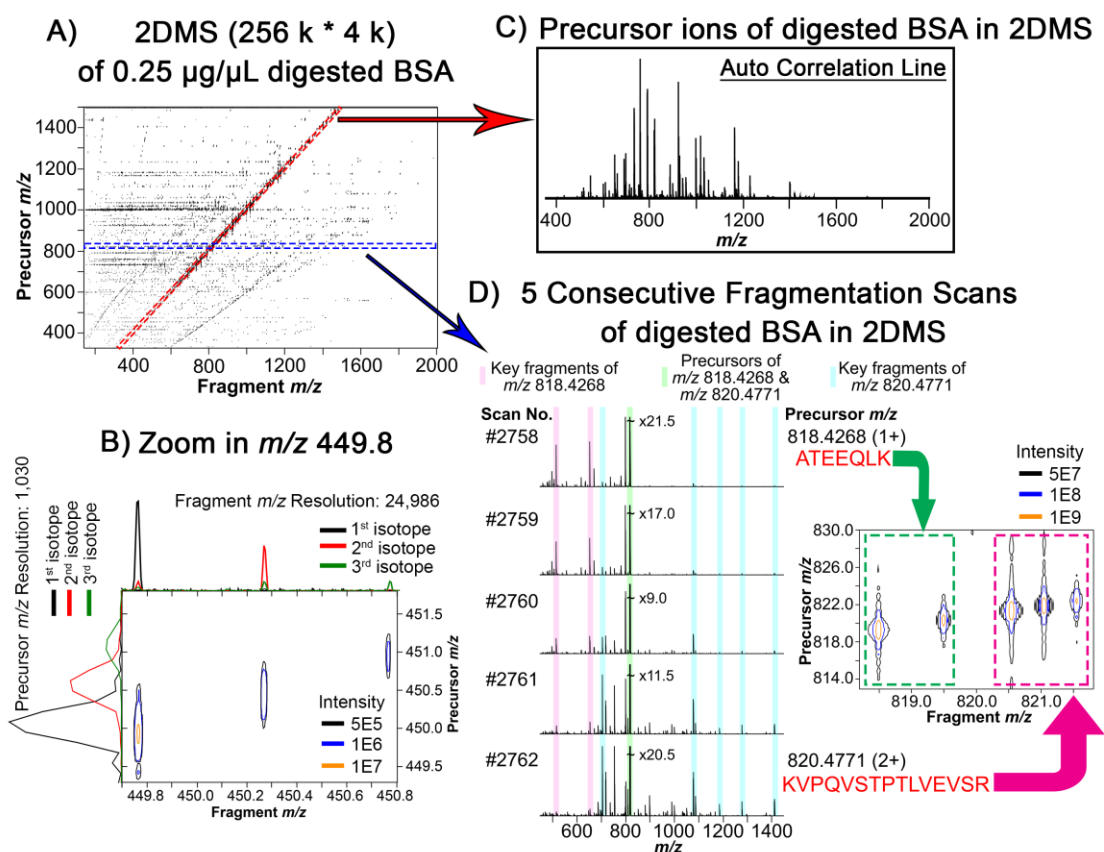
Figure 4.2. 2DMS spectrum of BSA tryptic digest. (A) The 2DMS spectrum of digested BSA acquired with 256 k * 4 k data-points (fragment *m/z* dimension * precursor *m/z* dimension). (B) Zoom-in of the 2DMS spectrum at *m/z* 449.8 region. Extractions of the (C) auto-correlation line (red arrow) and (D) 5 consecutive fragmentation scans (blue arrow) from the 2DMS spectrum in Figure 4.2A.

2DMS is a truly data independent acquisition method in which all detectable ions are fragmented simultaneously. Figure 4.2D shows the five consecutive fragmentation scans (from line number 2,758 to 2,762) in the 2DMS spectrum of the digested BSA which involved the fragmentation of two peptides with an *m/z* difference of ~ 2. In the 2,758th scan, the peptide ATEEQLK with an *m/z* of 818.4268 was assigned with the presence of the y_4 and y_5 fragments. The intensity of the precursor ion decreased ~ 2.4-fold with the scan number increased from 2,758 to 2,760. In the 2,760th scan, the 2DMS spectrum contained two precursor ions which represented the

peptides of ATEEQLK and KVPQVSTPTLVEVSR were co-fragmenting, the key fragments from both peptides, therefore, were observed in the same 2DMS spectrum. From the scans of 2,760 to 2,762, the intensity of the precursor ion increased again as the scanning precursor ion mass was approaching m/z 820.4771 which was the m/z ratio of the KVPQVSTPTLVEVSR peptide. In the $2,762^{th}$ scan, y_6, y_{11}, y_{12}, as well as y_{13} fragments from KVPQVSTPTLVEVSR were observed while the fragments from ATEEQLK peptide were not detected. Thus, we can baseline-differentiate the fragments of a peptide at m/z 818.43 from another of m/z 820.48, which demonstrated the advantage of using 2DMS over quadrupole isolation to separate ions with close m/z ratios as the ions were entirely isolated in two 2DMS spectra using an automatic pulse program without prior separation in a true data independent fashion, and the same effect is observed automatically throughout the 2DMS, demonstrating the potential of applying 2DMS to the study of complex proteomic samples with minimal or no LC separation.

Using the nLC FTICR CAD MS/MS, 42 peptides of the digested BSA were assigned, the sequence and cleavage coverages of BSA were 62% and 49% respectively (see Supplementary Information, Figure S4.6). The 2DMS of digested BSA with IRMPD fragmentation assigned 31 peptides which covered 51% of the sequence as well as 34% of the cleavage sites (see Supplementary Information, Figure S4.7). Combining the results of 2DMS with nLC MS/MS, 47 peptides in total were identified which covered 68% and 57% of the sequence and the potential cleavage sites (see Figure 4.3A). 26 out of 47 peptides (55%) were commonly identified in the 2DMS and nLC MS/MS experiments (see Figure 4.3B) and the fragments obtained from the 2DMS were comparable with the nLC MS/MS results (see Figure 4.3D). Complementary fragments were observed in 2DMS compared to nLC CAD MS/MS. The data showed the capability of 2DMS to analysis proteomic samples as nLC MS/MS.

A) Fragmentation map of digested BSA using 2DMS and nLC MS/MS
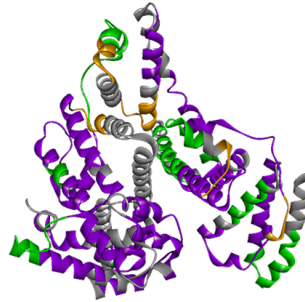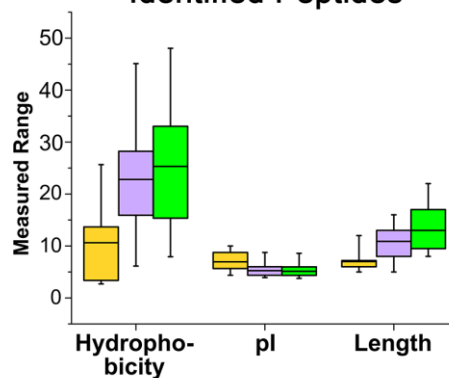

Total Sequence Coverage: 68%   Total Cleavage Coverage: 57%
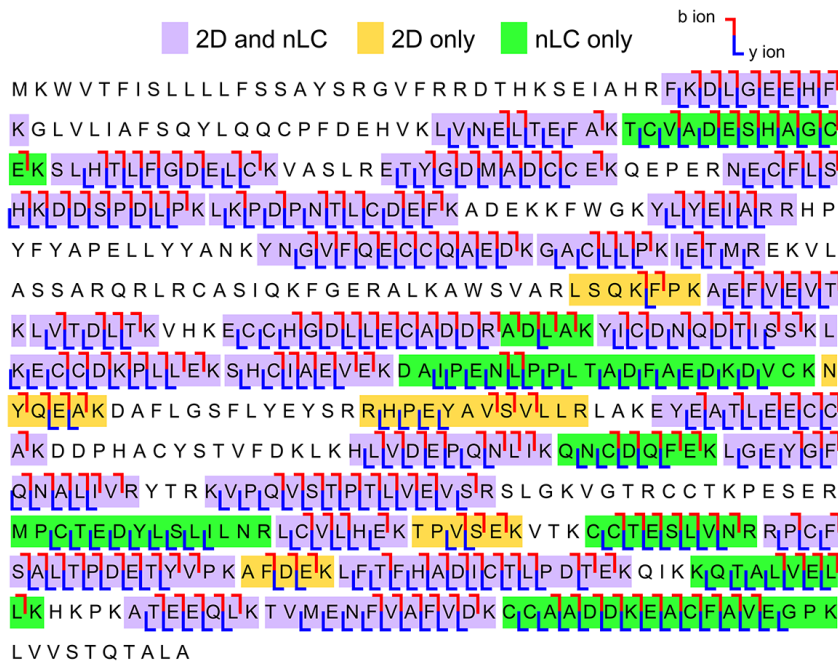
B) Venn Diagram of Identified Peptides

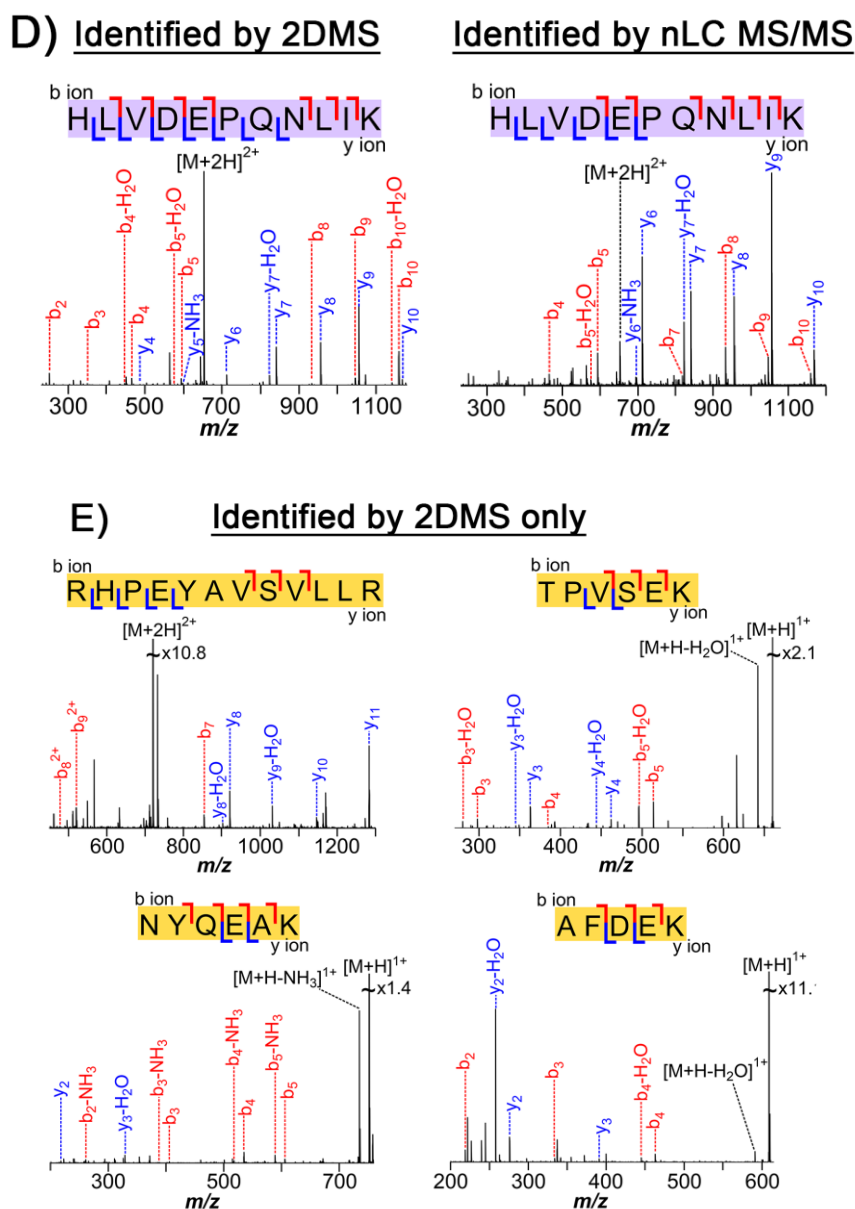# C) Physical Properties of Identified Peptides

303
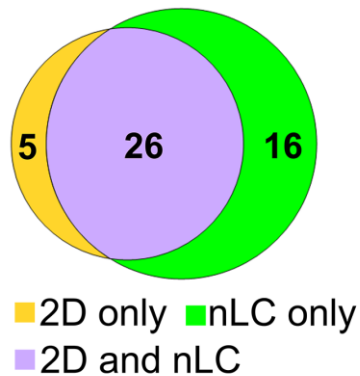
Figure 4.3. Proteomic results of BSA tryptic digest using 2DMS and nLC MS/MS. (A) Fragmentation map of digested BSA using 2DMS and nLC MS/MS. (B) Venn diagram showing the overlapping of the identified peptides from 2DMS and nLC MS/MS experiments. (C) Box chart showing the range of hydrophobicity, pI, and length covered by 2DMS only (yellow), 2DMS and nLC MS/MS (purple), and nLC MS/Ms only (green). (D) Spectra showing example peptides assigned by 2DMS (left) and nLC MS/MS (right) experiments. (E) Spectra showing peptides exclusively assigned by the 2DMS experiment.

304

Five peptides (11%) were exclusively detected by the 2DMS experiment with good sequence coverage that could not be assigned in the nLC MS/MS experiment (see Figure 4.3B & Figure 4.3E). The average hydrophobicity and the length of the peptides that were exclusively assigned by 2DMS were lower than those being assigned by the nLC MS/MS experiment, suggesting 2DMS has the advantage in detecting hydrophilic and/or short peptides over nLC MS/MS (see Figure 4.3C). The exclusive peptides being identified by 2DMS also contained higher pI value which implied those peptides were composed of more basic amino acid residues. The results herein showed peptides that were hydrophilic, short, and basic could be exclusively assigned by a 2DMS experiment, while these peptides were likely not retained on the column of RP-LC. These data suggested the results obtained from the 2DMS and nLC MS/MS experiments were partially complementary and the sequence as well as cleavage coverages were significantly enhanced when both methods were applied.

Comparison of 2DMS and nLC MS/MS using a mixture of standard digested proteins. A mixture of 5 digested proteins (5CP) was used to further evaluate the performance of 2DMS compared to nLC MS/MS. 75 peptides of 5CP were assigned in a nLC MS/MS experiment and the average sequence as well as cleavage coverages per protein were 60% and 47% respectively (see Supplementary Information, Figure S4.8). Using the same sample in the 2DMS experiment (256 k * 4 k data-points), only 30 peptides were assigned without prior separation, the average sequence and cleavage coverages per protein were 35% and 19% respectively (see Supplementary Information, Figure S4.9). This result showed that the performance of 2DMS was limited compared to nLC MS/MS for a more complex proteomic sample, suggesting either prior sample preparation or higher resolution of 2DMS experiment is required, so partial fractionation using SPE cartridge was explored in the first instance.
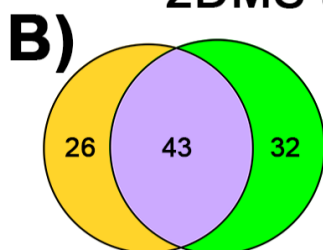
Three different SPE cartridges, including RP, HILIC, and SAX, were used to enrich peptides with various physical properties,[39] to vary the types of peptides observable in the 2DMS experiments.   A 2DMS spectrum (256 k * 4 k data-points) was acquired for each of the SPE enriched sample (see Supplementary Information, Figure S4.10 – S4.14), the numbers of peptide assigned from each experiment were 51 (RP pH 2), 35 (HILIC pH 6.8), 37 (RP pH 10), and 33 (SAX pH 10) (see Figure 4.4A). The Venn diagram of the peptides identified by various SPE cartridges showed that most of the peptides were able to be observed in the conditions of C18 RP at pH 2 and HILIC at pH 6.8, suggesting further experiments could focus on these two enrichment samples as they could provide the highest sequence coverage of the digested 5CP sample.

A) Peptides identified by various SPE cartridges

SPE	RP (pH 2)	HILIC (pH 6.8)	RP (pH 10)	SAX (pH 10)
No. of Peptide	51	35	37	33
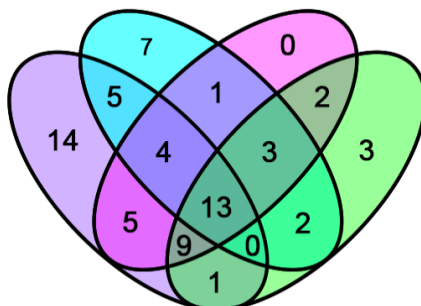
Peptides identified between 2DMS and nLC MS/MS

B)

2D and nLC
2D only
nLC only

C)

Measured Range

Hydrophobicity   pI   Length

D) Peptides identified by 2DMS only

b ion
$[F L A S V S T V L T S K]^{2+}$
y ion

$[M+2H]^{2+}$

y_3 y_4 b_5-H_2O y_5 y_6 y_8 b_9-H_2O $y_{10}-H_2O$ y_{10} $b_{10}-H_2O$ y_9 $b_{11}-H_2O$

400 600 800 1000
m/z

b ion
$[L V N E L T E F A K]^{2+}$
y ion

$[M+2H]^{2+}$

y_4 y_5 b_6-H_2O y_6-H_2O y_6 b_7-H_2O y_7 y_8-H_2O y_8 b_9-H_2O b_9

600 800 1000
m/z

b ion
$[N Y A E A K]^{1+}$
y ion

$[M+H]^{1+}$
$[M+H-NH_3]^{1+}$ x4.4

y_2 b_2 y_3-H_2O b_3-NH_3 b_3 b_4-NH_3 b_4 b_5-NH_3 b_5

250 450 650
m/z

b ion
$[A A C L L P K]^{1+}$
y ion

$[M+H]^{1+}$
x3.8

y_2 b_3 y_3 b_4 y_4 b_5 y_5

250 450 650
m/z
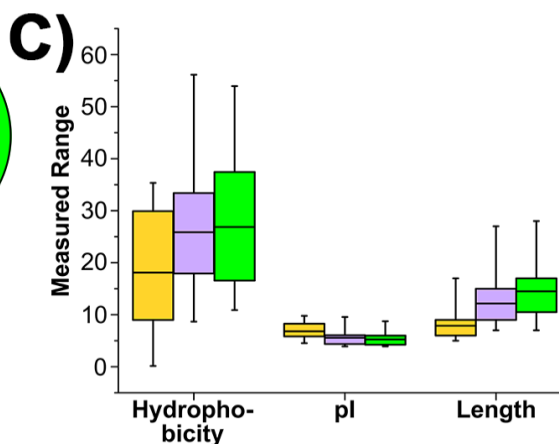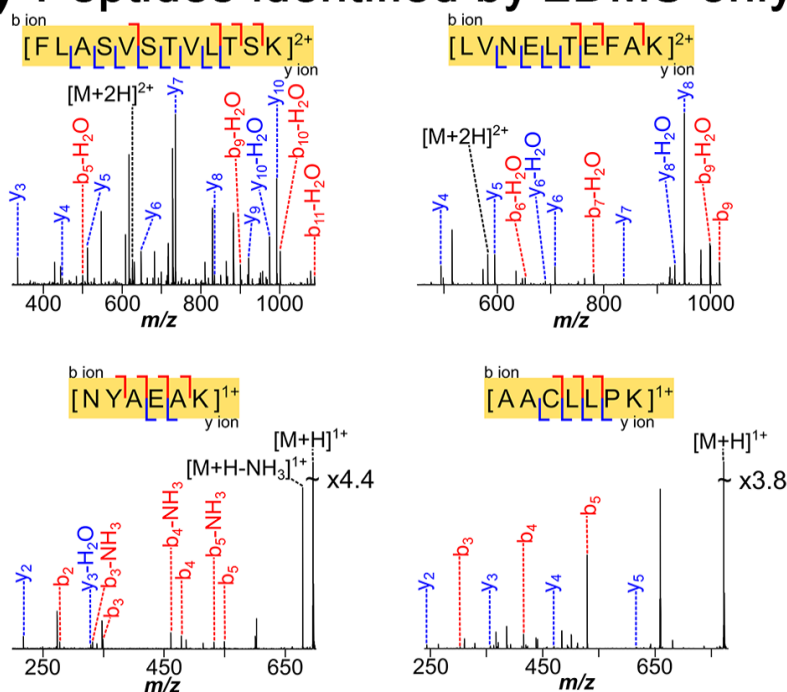
(Figure 4.4. Continue)

E) Fragmentation map of digested 5CP with 2DMS (SPEs enrichment) and nLC MS/MS
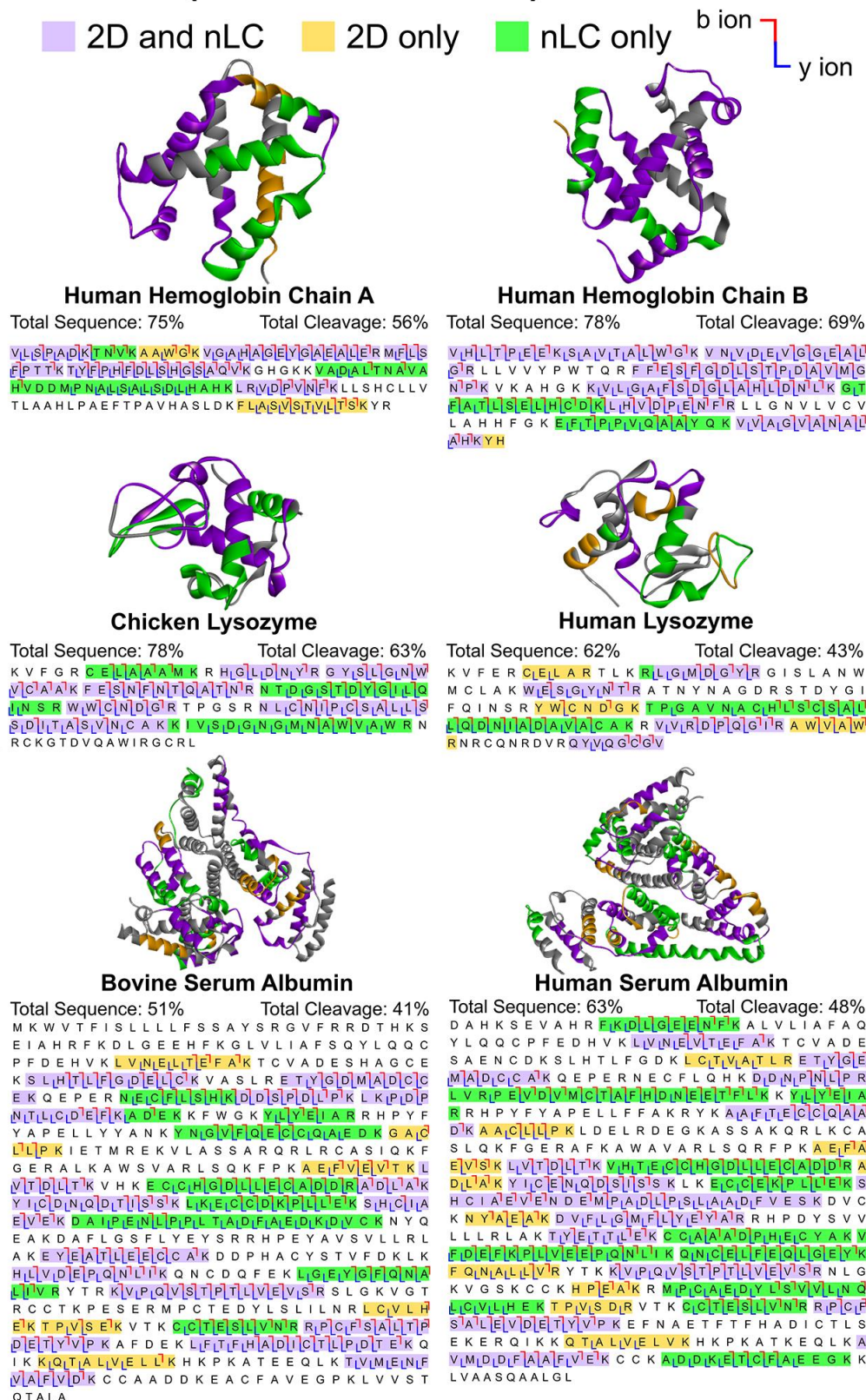
Figure 4.4. Proteomic results of 5CP tryptic digest using 2DMS with SPEs enrichment and nLC MS/MS. (A) The number and Venn diagram of peptides from 5CP identified by C18 RP at pH 2 (purple), HILIC at pH 6.8 (cyan), C18 RP at pH 10 (pink), and SAX at pH 10 (green). (B) Venn diagram of peptides being identified by 2DMS via SPE cartridge enrichment and nLC MS/MS. (C) Box chart diagram to show the hydrophobicity, pI, and length of peptides being detected in 2DMS and nLC MS/MS. (D) Spectra showing the exclusive peptides being assigned by 2DMS experiments. (E) Fragmentation maps of the digested 5CP showing the complementary results of 2DMS and nLC MS/MS experiments.

From the 5CP sample, 69 peptides in total were assigned from the 2DMS spectra of the SPE enriched samples and the average sequence as well as cleavage coverages per protein were 45% and 29% respectively (see Supplementary Information, Figure S4.15). Compared to the non-enriched 2DMS result, 10% additional coverage was observed in both average sequence as well as cleavage coverages. 26 out of 101 peptides (26%) were exclusively detected in the 2DMS spectra with good sequence coverage while 43% of the peptides were mutually observed in both 2DMS and nLC MS/MS spectra (see Figure 4.4B & 4.4D). The average sequence and cleavage coverages per protein were enhanced by 8% and 6% respectively (see Figure 4.4E) by combining the results to the nLC MS/MS data (see Supplementary Information, Figure S4.8). As seen previously, the more hydrophilic, basic, and short peptides were exclusively identified in the 2DMS experiment; the range of physical properties, however, were enhanced after the SPE cartridges enrichment (see Figure 4.4C). The results here demonstrated the peptides' properties were diversified through SPE enrichment, more exclusive peptides were detected in the experiments of 2DMS which resulted in enhancing the sequence and cleavage coverages of the proteins when the

data of 2DMS and nLC MS/MS were combined.   More importantly, the solvents used

in each 2DMS experiment could be varied due to the different buffers required for SPE

cartridge enrichment; but the variations in solvents did not affect the performance of

the 2DMS experiments, suggesting this method could remove some solvent limitations
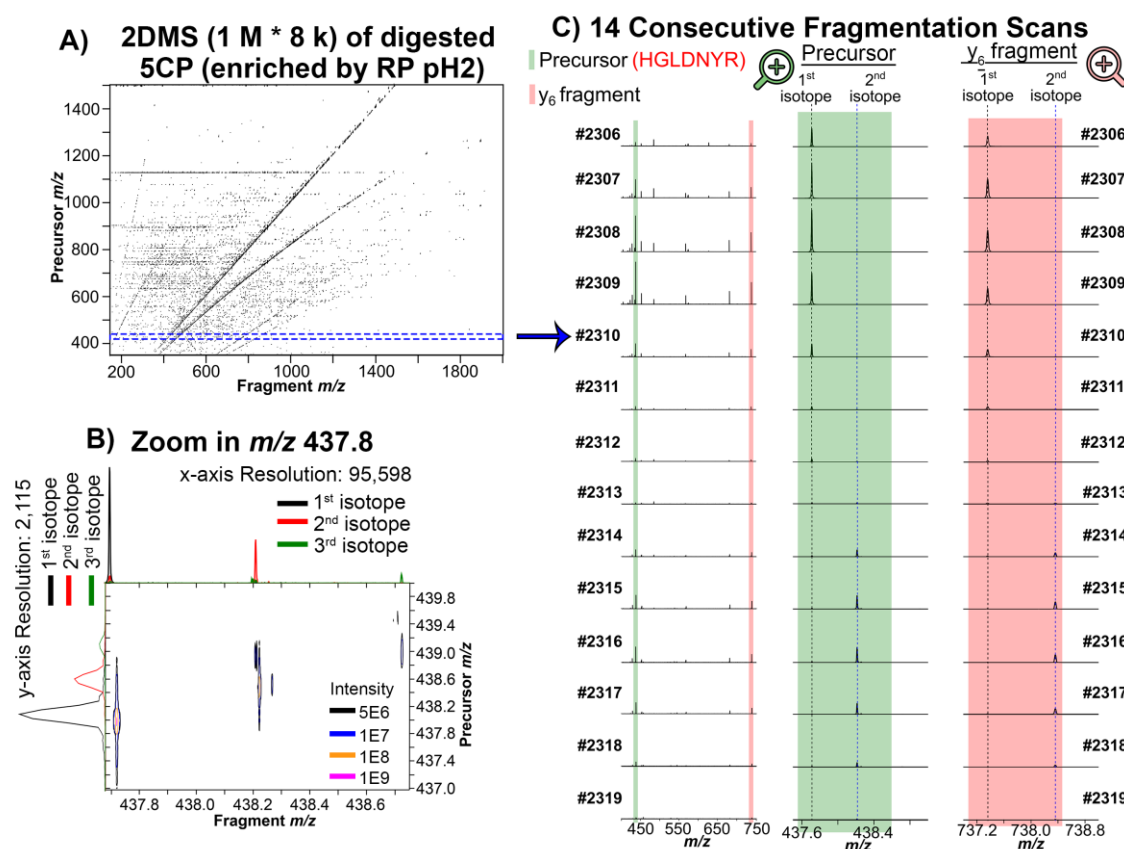
and retention bias.

Figure 4.5.   High resolution 2DMS spectrum of 5CP tryptic digest.   (A) 2DMS

spectrum with resolution 1 M * 8 k data-points (fragment *m/z* * precursor *m/z*) of

digested 5CP which was enriched by C18 RP SPE cartridge at pH 2.   (B) Zoom-in of

the 2DMS spectrum at the *m/z* 437.8 region.   (C) The 14 consecutive fragmentation

scans (from scan number 2,306 to 2,319) extracted from the 2DMS spectrum which

correspond to the HGLDYR peptide.

    Optimisation of acquisition parameters was also explored to improve 2DMS

proteomic using 5CP (see Supplementary Information, Figure S4.16 – S4.21).

Unsurprisingly using a higher resolution in both fragment and precursor m/z dimensions (1 M * 8 k data-points) increased detectable fragment intensity, allowing more fragments to be observed and assigned in the 2DMS spectrum of the digested 5CP with SPE enrichment (see Figure 4.5A), and the 2DMS experiment time expanded from ~ 55 minutes to ~ 167 minutes (see Supplementary Information, Table S4.1).   The peak resolutions (at m/z 437.8) in the fragment and precursor m/z dimension were ~ 96 k and ~ 2 k respectively (see Figure 4.5B), additionally these resolutions improved peak shapes in the 2DMS spectrum through resolving closely spaced peaks and enhanced peak S/N ratios which assisted in peptide assignment (see Supplementary Information, Figure S4.22).

Increasing the resolution in the fragment and precursor m/z dimensions did not only enhance the number of peptides identified, but also improved separation in the mass of precursor ion.   In Figure 4.5C, 14 consecutive fragmentation scans were extracted from the 2DMS spectrum (1 M * 8 k data-points) of the digested 5CP (see Figure 4.5A), these 14 scans all show the HGLDNYR peptide at m/z 437.7119 (2+). The intensity of the precursor ion (green region) increased twice within these spectra which were between the 2,306[th] and 2,309[th] scans and the 2,313[th] and 2,316[th] scans, similar intensity variation was observed in the key fragment (y_6) ion (pink region). Zooming-in the m/z region of the precursor ion clearly shows the first increment in intensity was due to the rise of the 1st isotope of the precursor ion (between the 2,306[th] and 2,309[th] scans) and the second increment in the intensity was caused by the rise of the 2nd isotope of the precursor ion (between the 2,313[th] and 2,316[th] scans); as a result, a similar observation was also obtained for the y_6 fragment.   Thus, 2DMS has baseline separated precursor and fragments at the individual isotope level which is difficult to achieve using quadrupole or even in-cell isolations in a data independent fashion. This capability will be particularly useful for separation of fragments from nearby

precursors such as the peaks between deamidated and non-deamidated species (Δ mass = 0.984 Da).[48, 49]
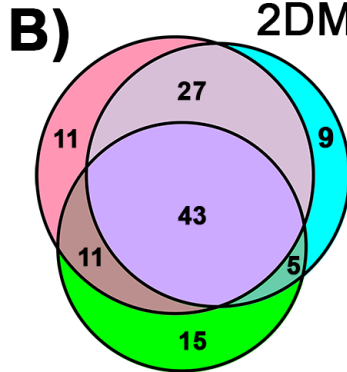
Mass range fractionation using several wide mass isolation windows in the quadrupole (see Supplementary Information, Figure S4.23 – S4.24) as well as SPE offline fractionation (see Supplementary Information, Figure S4.25 – 4.26) were applied in combination to reduce the number of ions in the FTICR cell and minimize ion suppression effects, so that the number of peptides being identified by 2DMS could be enhanced further. The results of the mass range fractionation showed that the number of peptides could be significantly enhanced (RP identified: 64 and HILIC identified: 85), as a result, a better sequence and cleavage coverages were achieved (see Supplementary Figure S4.27 – S4.28). The data showed that the number of peptides being identified by 2DMS could reach a similar, or even better than that of standard nLC MS/MS without online LC separation and extra sample preparation, suggesting this method is suitable for the study of complex proteomic samples as limited sample preparation and separation was required; while still able to achieve similar results as standard nLC MS/MS.

A) Peptides identified by various SPE in each fraction

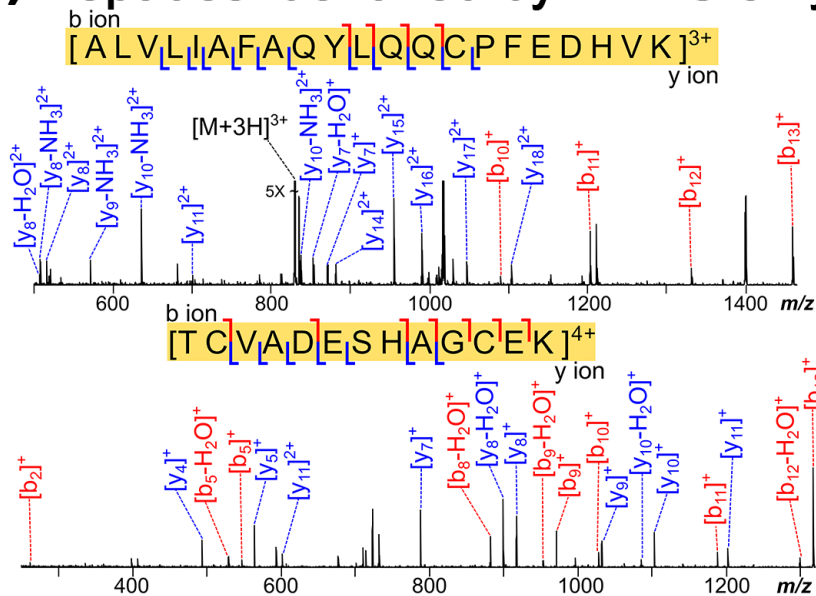| SPE	RP (pH 2)	HILIC (pH 6.8)
No. of Peptide	92	85
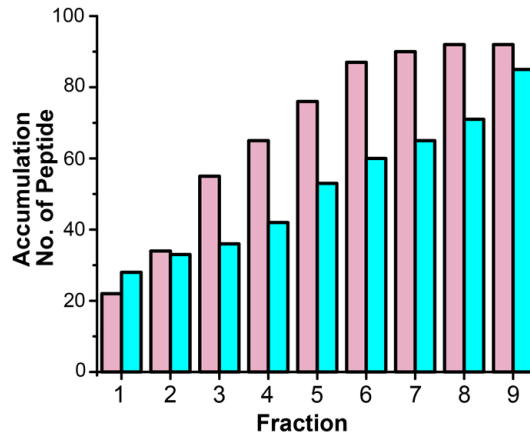
Peptides identified between 2DMS and nLC MS/MS

B)

- RP 2D
- HILIC 2D
- nLC MS/MS

C)

- 2D and nLC
- 2D only
- nLC only

D) Peptides identified by 2DMS only

$[A\ L\ V\ L\ I\ A\ F\ A\ Q\ Y\ L\ Q\ Q\ C\ P\ F\ E\ D\ H\ V\ K]^{3+}$

$[T\ C\ V\ A\ D\ E\ S\ H\ A\ G\ C\ E\ K]^{4+}$

(Figure 4.6. Continue)

E) Fragmentation map of digested 5CP using 2DMS (fractionated) and nLC MS/MS
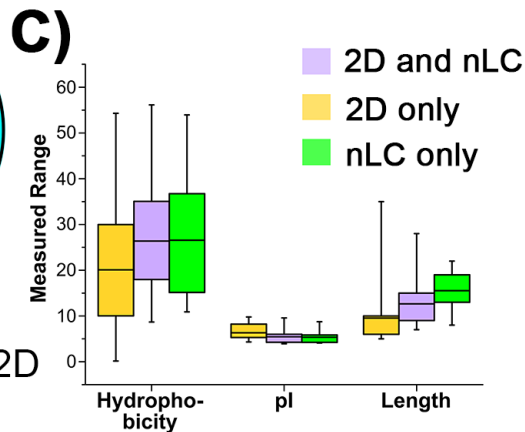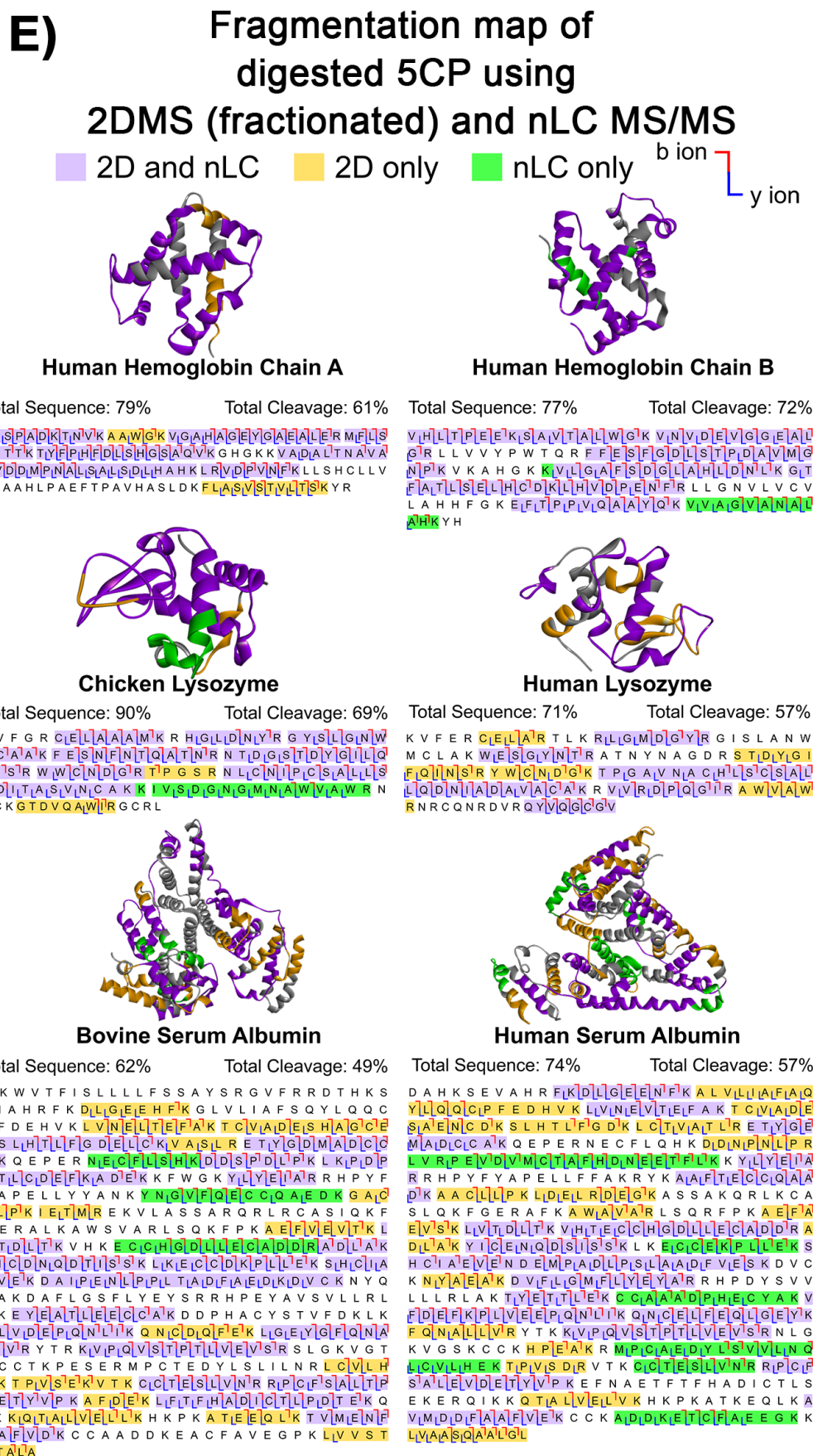
Figure 4.6. Proteomic results of 5CP tryptic digest using 2DMS with SPE offline fractionation and nLC MS/MS. (A) Peptides identified in each offline fraction using RP (pink) and HILIC (cyan) SPE technique. (B) Venn diagram showing the overlapping of peptides that identified in 2DMS (fractionated) and nLC MS/MS. (C) Box chart showing the physical properties of the peptides that identified in 2DMS (fractioned) and nLC MS/MS. (D) Spectra of the exclusive peptides identified using 2DMS. (E) Fragmentation maps of digested 5CP using offline fractionation 2DMS and nLC MS/MS.

The offline SPE fractionation using RP and HILIC cartridges provided the highest sequence and cleavage coverages for the 5CP among all the methods tested (see Supplementary Information, Figure S4.29 and S4.30). 92 and 85 peptides in total were assigned by the RP and HILIC fractionation 2DMS, respectively (see Figure 4.6A), and more peptides were assigned using either one of these methods compared to nLC MS/MS. Thus, the offline SPE fractionation provided a similar or better analytical ability than nLC MS/MS in proteomic study.

121 peptides were assigned using the offline fractionation 2DMS and nLC MS/MS results (see Figure 4.6B). 47 out of 121 peptides (39%) were exclusively identified using 2DMS. The average sequence and cleavage coverages per protein were 76% and 61%, respectively, (see Figure 4.6E) which has been increased by 27% and 30% compared to purely nLC MS/MS (see Supplementary Information, Figure S4.8). These results, once again, showed that the data obtained from 2DMS and nLC MS/MS were complementary, thus applying both 2DMS and nLC MS/MS could provide the highest sequence and cleavage coverages of proteins. The box chart diagram in Figure 4.6C shows the peptides exclusively assigned by 2DMS were more hydrophilic, basic, as well as shorter in general than nLC MS/MS as previously noted. However, the

ranges of the physical properties were the widest compared to the previous 2DMS results, suggesting the offline SPE fractionation was a better method to diversify the types of peptides observable in 2DMS. The most significant improvement in using offline fractionation was relatively long peptides with good sequence coverage were able to be detected and assigned in the 2DMS spectra (see Figure 4.6D).

Comparison of 2DMS and nLC MS/MS using digested yeast. From the 2DMS results of the digested 5CP, the high resolution (1 M * 8 k data-points) and C18 RP SPE cartridge offline fractionation 2DMS experiments yielded the highest number of identified peptides per sample. Thus, these two acquisition methods were combined and applied to the digested yeast sample to evaluate the performance of 2DMS in complex proteomic studies. The digested yeast sample was first analysed using nLC MS/MS (see Supplementary Information, Figure S4.31), 153 proteins and 332 peptides were identified at 1% false discovery rate (FDR). The number of proteins and peptides being identified in this nLC MS/MS experiment was not as high as the previous literature,[21] mainly because the scan rate in FTICR MS limited the number of tandem MS/MS spectra acquired per injection.

The digested yeast sample was then analysed using 2DMS with 1 M * 8 k data-points (see Supplementary Information, Figure S4.32A). 28 peptides corresponding to 20 yeast proteins were manually assigned in the 2DMS spectra, and 16 peptides (57%) were exclusively observed in the 2DMS (see Supplementary Information, Figure S4.32B – S4.32D). The number of peptides being identified in the 2DMS without fractionation was lower than expected, the main reason was due to the complexity of the digested yeast sample (see Supplementary Information, Figure S4.33). In the 5CP and the yeast 2DMS experiments, samples were both diluted into 0.25 µg/µL and sprayed into the MS. In the 5CP sample, there were ~ 400 peptides (including 1

missed cleavage) and the dynamic range of the sample was around an order of magnitude; however in the digested yeast, it contained more than 577,000 tryptic digested peptides (no missed cleavage included) and the dynamic range of the proteins was at least 6 orders of magnitude.[50]   The concentration per peptide in the yeast solution, therefore, was much lower than the peptide concentration in the 5CP sample. Since the initial intensity as well as the fragmentation efficiency were the two factors that influenced the intensities of fragments, thus the fragments obtained in the 2DMS spectrum of yeast were extremely low that the fragment peaks could not picked by the software used.   Further optimisation of the loading of sample for yeast 2D MS analysis is ongoing.

A) No. of yeast protein & peptide

	2DMS	nLC MS/MS
No. of Protein	147	153
No. of Peptide	243	332

B) Yeast Proteins

Yeast Peptides

C) Physical properties of the identified peptides
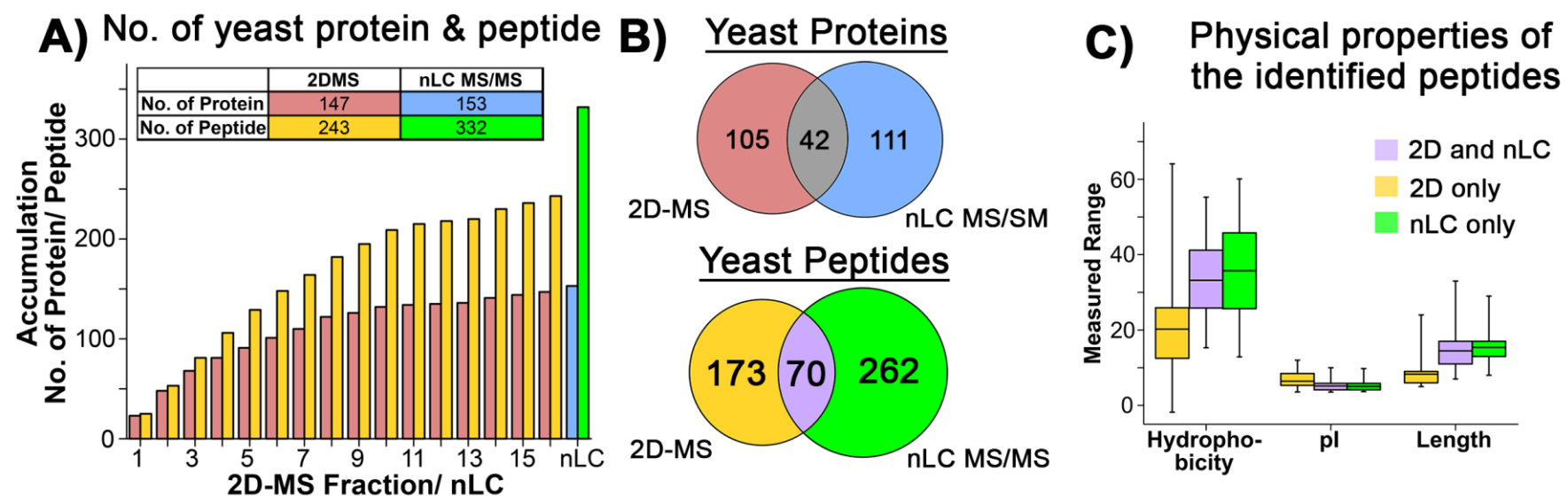
(Figure 4.7. Continue)
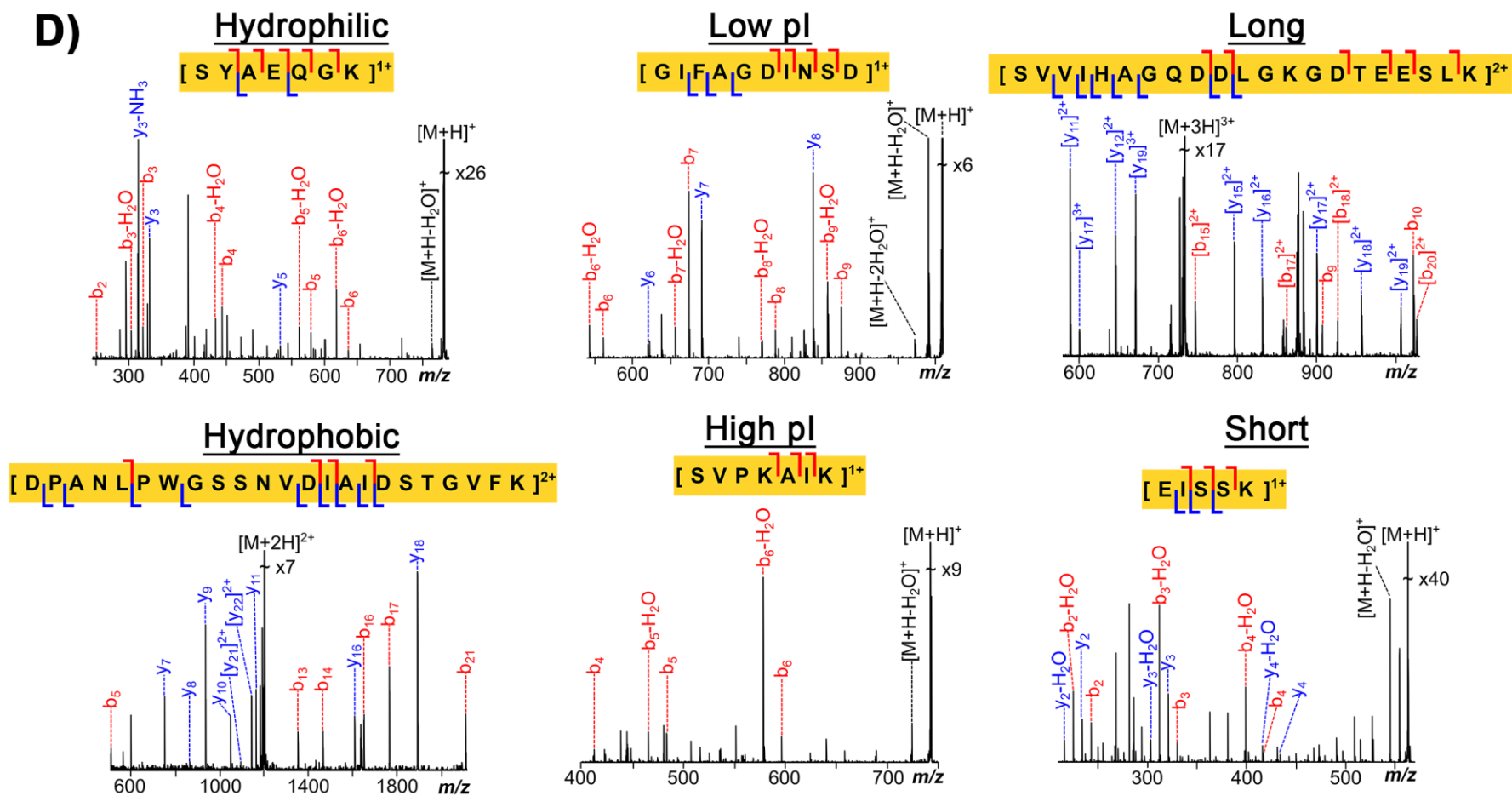
Figure 4.7. Comparison of SPE fractionated 2DMS and nLC MS/MS results for yeast tryptic digest. (A) Graph showing the accumulated

number of proteins and peptides being identified in each offline fraction using 2DMS as well as nLC MS/MS.   (B) Venn diagram showing the complementary proteins and peptides assigned in 2DMS and nLC MS/MS experiments.   (C) Box chart showing the physical properties of the peptides identified in 2DMS only (yellow), 2D and nLC (purple), and nLC only (green).   (D) 2DMS spectra showing peptides exclusively assigned by 2DMS.

The yeast sample was then separated into 16 fractions using a C18 RP SPE cartridge, shown above to achieve a higher number of peptide assignments (see Figure 4.6). 16 2DMS spectra in total with 256 k * 4 k data-points each were then acquired (see Supplementary Information, Figure S4.34 – S4.35). 147 proteins and 243 peptides in total were manually assigned in the experiment (see Figure 4.7A). The results here show that the number of proteins being identified in 2DMS was similar to the result obtained from nLC MS/MS (147 vs 153), but the assigned peptides in the 2DMS experiment were ~ 25% fewer than the nLC MS/MS experiment (243 vs 332). Combining the results of 2DMS with nLC MS/MS, 258 proteins and 505 peptides in total were identified; however, only 42 proteins (16%) and 70 peptides (14%) were commonly assigned in both experiments (see Figure 4.7B). The numbers of protein and peptides were doubled compared to purely 2DMS or nLC MS/MS result, indicating the data provided by 2DMS and nLC MS/MS experiments were largely complementary, which is useful to enhance the analysis coverage of proteomes. The 243 yeast peptides which were assigned in the 2DMS experiment were further compared with the previous literature that has been recognized as the benchmark of protein and peptide assignments, for yeast, being assigned in a single experiment (see Supplementary Information, Figure S4.36).[18] 126 out of 243 peptides (52%) from 2DMS experiment were commonly assigned in both experiments, which indicated 48% of the peptides were not detected previously, showing certain areas of the proteome may be more easily accessed or studied using 2DMS than standard nLC MS/MS.

The average value in the box chart of assignments (see Figure 4.7C) also showed the peptides that were exclusively assigned by 2DMS tends to be more hydrophilic, basic, and short as shown previously. Example peptides that represented different physical properties and those exclusively assigned by 2DMS are shown in Figure 4.7D. The results of the yeast analysis effectively demonstrated the capability of 2DMS in

proteomic studies of complex systems as the number of proteins and peptides being assigned in the experiments were similar to that of nLC MS/MS experiment. However, the peptides assigned in 2DMS were also frequently complementary to nLC MS/MS results. Combining the results of 2DMS with nLC MS/MS enhanced the coverage of the yeast proteome.

4.5 Conclusion

   The results presented demonstrate that 2DMS is a viable alternative technique for proteomic studies, showing online LC MS/MS is no long an unavoidable tool in proteomics. Therefore, many proteomic analyses, such as biomolecular non-covalent interactions, which are usually very limited by available LC solvent systems and the column chemistry of chromatography, are able to be performed using 2DMS. Furthermore, complementary protein and peptide assignments from the results of 2DMS and nLC MS/MS suggests a deeper proteome coverage can be achieved by combing the data obtained from 2DMS and nLC MS/MS, which also indicates the proteomes that exclusively covered by 2DMS are currently under-studied by standard nLC MS/MS techniques. Thus 2DMS is not only an alternative but may be considered as a required complementary technique for deeper proteome analysis.   2DMS also shows to have a higher efficiency in observing hydrophilic, basic, or short peptides, suggesting 2DMS is also suitable for the application in PTMs analysis as most of the important PTMs in disease-related issue, such as phosphopeptides and glycopeptides, are extremely hydrophilic which is challenging to analysis using standard nLC MS/MS technique.

4.6 References

1.   Larance, M.; Lamond, A. I., Multidimensional proteomics for cell biology. *Nat. Rev. Mol. Cell Biol.* **2015,** *16* (5), 269.

2.   Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.

3.   Aebersold, R.; Mann, M., Mass-spectrometric exploration of proteome structure and function. *Nature* **2016,** *537* (7620), 347-355.

4.   Yates, J. R., Mass spectrometry and the age of the proteome. *J. Mass Spectrom.* **1998,** *33* (1), 1-19.

5.   McDonald, W. H.; Yates, J. R., Shotgun proteomics and biomarker discovery. *Dis. Markers* **2002,** *18* (2), 99-105.

6.   Bennett, E. J.; Shaler, T. A.; Woodman, B.; Ryu, K.-Y.; Zaitseva, T. S.; Becker, C. H.; Bates, G. P.; Schulman, H.; Kopito, R. R., Global changes to the ubiquitin system in Huntington's disease. *Nature* **2007,** *448* (7154), 704.

7.   Gillette, M. A.; Carr, S. A., Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. *Nat. Methods* **2013,** *10* (1), 28.

8.   Kellie, J. F.; Tran, J. C.; Lee, J. E.; Ahlf, D. R.; Thomas, H. M.; Ntai, I.; Catherman, A. D.; Durbin, K. R.; Zamdborg, L.; Vellaichamy, A., The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Mol. Biosyst.* **2010,** *6* (9), 1532-1539.

9.   Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.-H.; Begley, T. P.; McLafferty, F. W., Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. *J. Am. Chem. Soc.* **2002,** *124* (4), 672-678.

10.   Wootton, C. A.; Lam, Y. P.; Willetts, M.; van Agthoven, M. A.; Barrow, M. P.; Sadler, P. J.; Peter, B., Automatic assignment of metal-containing peptides in proteomic LC-MS and MS/MS data sets. *Analyst* **2017,** *142* (11), 2029-2037.

11.   Zhang, P.; Chiu, C. K.; Huang, H.; Lam, Y. P.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; Sadler, P. J., Organoiridium photosensitizers induce specific oxidative attack on proteins within cancer cells. *Angew. Chem. Int. Ed.* **2017,** *56* (47), 14898-14902.

12.   Adav, S. S.; Qian, J. R.; Ang, Y. L.; Kalaria, R. N.; Lai, M. K. P.; Chen, C. P.; Sze, S. K., iTRAQ Quantitative Clinical Proteomics Revealed Role of Na+K+-ATPase and Its Correlation with Deamidation in Vascular Dementia. *J. Proteome Res.* **2014,** *13* (11), 4635-4646.

13.   Gravett, M. G.; Novy, M. J.; Rosenfeld, R. G.; Reddy, A. P.; Jacob, T.; Turner, M.; McCormack, A.; Lapidus, J. A.; Hitti, J.; Eschenbach, D. A., Diagnosis of intra-amniotic infection by proteomic profiling and identification of novel biomarkers. *JAMA* **2004,** *292* (4), 462-469.

14.   Yates, J. R.; Ruse, C. I.; Nakorchevsky, A., Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **2009,** *11*, 49-79.

15.   Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P., Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC−MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2003,** *2* (1), 43-50.

16.   Ducret, A.; Oostveen, I. V.; Eng, J. K.; Yates, J. R.; Aebersold, R., High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein Sci.* **1998,** *7* (3), 706-719.

17.   Washburn, M. P.; Wolters, D.; Yates III, J. R., Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001,** *19* (3), 242.

18.   Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J., The one hour yeast proteome. *Mol. Cell. Proteomics* **2014,** *13* (1), 339-347.

19.   Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Færgeman, N. J.; Mann, M.; Jensen, O. N., Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **2005,** *4* (3), 310-327.

20.   Motoyama, A.; Yates, J. R., Multidimensional LC Separations in Shotgun Proteomics. *Anal. Chem.* **2008,** *80* (19), 7187-7193.

21.   Richards, A. L.; Hebert, A. S.; Ulbrich, A.; Bailey, D. J.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J., One-hour proteome analysis in yeast. *Nat. Protoc.* **2015,** *10* (5), 701.

22.   Bleiholder, C.; Do, T. D.; Wu, C.; Economou, N. J.; Bernstein, S. S.; Buratto, S. K.; Shea, J.-E.; Bowers, M. T., Ion Mobility Spectrometry Reveals the Mechanism of Amyloid Formation of A beta(25-35) and Its Modulation by Inhibitors at the Molecular Level: Epigallocatechin Gallate and Scyllo-inositol. *J. Am. Chem. Soc.* **2013,** *135* (45), 16926-16937.

23.   Woods, L.; Radford, S.; Ashcroft, A., Advances in ion mobility spectrometry–mass spectrometry reveal key insights into amyloid assembly. *Biochim. Biophys. Acta, Proteins Proteomics* **2013,** *1834* (6), 1257-1268.

24.   Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E., Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nat. Chem.* **2015,** *7* (1), 73-81.

25.   Landreh, M.; Alvelius, G.; Johansson, J.; Jornvall, H., Insulin, islet amyloid polypeptide and C-peptide interactions evaluated by mass spectrometric analysis.

Rapid Commun. Mass Spectrom. **2014,** *28* (2), 178-184.

26.   Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T., Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to beta-sheet in amyloid fibril formation. *Nat. Chem.* **2011,** *3* (2), 172-177.

27.   Pan, J. X.; Han, J.; Borchers, C. H.; Konermann, L., Conformer-Specific Hydrogen Exchange Analysis of A beta(1-42) Oligomers by Top-Down Electron Capture Dissociation Mass Spectrometry. *Anal. Chem.* **2011,** *83* (13), 5386-5393.

28.   Woods, V. L.; Hamuro, Y., High resolution, high-throughput amide deuterium exchange-mass spectrometry (DXMS) determination of protein binding site structure and dynamics: Utility in pharmaceutical design. *J. Cell. Biochem.* **2001,** *84* (S37), 89-98.

29.   Pfändler, P.; Bodenhausen, G.; Rapin, J.; Houriet, R.; Gäumann, T., Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *Chem. Phys. Lett.* **1987,** *138* (2-3), 195-200.

30.   Ross III, C. W.; Guan, S.; Grosshans, P. B.; Ricca, T. L.; Marshall, A. G., Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry/mass spectrometry with stored-waveform ion radius modulation. *J. Am. Chem. Soc.* **1993,** *115* (17), 7854-7861.

31.   van Agthoven, M. A.; Barrow, M. P.; Chiron, L.; Coutouly, M.-A.; Kilgour, D.; Wootton, C. A.; Wei, J.; Soulby, A.; Delsuc, M.-A.; Rolando, C., Differentiating fragmentation pathways of cholesterol by two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (12), 2105-2114.

32.   van Agthoven, M. A.; Lynch, A. M.; Morgan, T. E.; Wootton, C. A.; Lam, Y. P. Y.; Chiron, L.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Can Two-Dimensional IR-ECD Mass Spectrometry Improve Peptide de Novo Sequencing? *Anal. Chem.* **2018,** *90* (5), 3496-3504.

33.   Floris, F.; van Agthoven, M.; Chiron, L.; Soulby, A. J.; Wootton, C. A.; Lam, Y. P.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., 2D FT-ICR MS of calmodulin: a top-down and bottom-up approach. *J. Am. Soc. Mass. Spectrom.* **2016,** *27* (9), 1531-1538.

34.   Floris, F.; van Agthoven, M. A.; Chiron, L.; Wootton, C. A.; Lam, P. Y. Y.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Bottom-Up Two-Dimensional Electron-Capture Dissociation Mass Spectrometry of Calmodulin. *J. Am. Soc. Mass. Spectrom.* **2018,** *29* (1), 207-210.

35.   Floris, F.; Chiron, L.; Lynch, A. M.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Application of Tandem Two-Dimensional Mass Spectrometry for Top-Down Deep Sequencing of Calmodulin. *J. Am. Soc. Mass. Spectrom.* **2018**.

36.   Floris, F.; Chiron, L.; Lynch, A. M.; Barrow, M. P.; Delsuc, M.-A.; O'Connor, P. B., Top-Down Deep Sequencing of Ubiquitin Using Two-Dimensional Mass Spectrometry. *Anal. Chem.* **2018,** *90* (12), 7302-7309.

37.   Simon, H.; van Agthoven, M.; Lam, P.; Floris, F.; Chiron, L.; Delsuc, M.-A.; Rolando, C.; Barrow, M.; O'Connor, P., Uncoiling collagen: a multidimensional mass spectrometry study. *Analyst* **2016,** *141* (1), 157-165.

38.   Maria A. van Agthoven, C. A. W., Lionel Chiron, Marie-Aude Coutouly, Andrew Soulby, Juan Wei, Mark. P. Barrow, Marc-Andre Delsuc, Christian Rolando, Peter B. O'Connor, Two-Dimensional Mass Spectrometry for Proteomics, a Comparative Study with Cytochrome c. *J. Anal. Chem.* **2016**.

39.   Manadas, B.; Mendes, V. M.; English, J.; Dunn, M. J., Peptide fractionation in proteomics approaches. *Expert review of proteomics* **2010,** *7* (5), 655-663.

40.   Caravatti, P.; Allemann, M., The 'Infinity Cell': a New Trapped-ion Cell With Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991,** *26* (5), 514-518.

41.   AL Rocha, V.; N Maeda, R.; Pereira, N.; F Kern, M.; Elias, L.; Simister, R.; Steele-King, C.; Gómez, L. D.; McQueen-Mason, S. J., Characterization of the cellulolytic secretome of Trichoderma harzianum during growth on sugarcane bagasse and analysis of the activity boosting effects of swollenin. *Biotechnol. Progr.* **2016,** *32* (2), 327-336.

42.   Chiron, L.; van Agthoven, M. A.; Kieffer, B.; Rolando, C.; Delsuc, M.-A., Efficient denoising algorithms for large experimental datasets and their applications in Fourier transform ion cyclotron resonance mass spectrometry. *Proc. Natl. Acad. Sci. USA* **2014,** *111* (4), 1385-1390.

43.   Scigelova, M.; Hornshaw, M.; Giannakopulos, A.; Makarov, A., Fourier transform mass spectrometry. *Mol. Cell. Proteomics* **2011,** *10* (7), M111. 009431.

44.   Krokhin, O. V.; Spicer, V., Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides. *Anal. Chem.* **2009,** *81* (22), 9522-9530.

45.   Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M. R.; Appel, R. D.; Bairoch, A., Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*, Springer: New Jersey, 2005; pp 571-607.

46.   Geiger, T.; Cox, J.; Mann, M., Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell. Proteomics* **2010,** *9* (10), 2252-2261.

47.   Liu, Q.; Cobb, J. S.; Johnson, J. L.; Wang, Q.; Agar, J. N., Performance Comparisons of Nano-LC Systems, Electrospray Sources and LC–MS-MS Platforms. *J. Chromatogr. Sci.* **2013,** *52* (2), 120-127.

48.   Robinson, A. B.; Rudd, C. J., Deamidation of glutaminyl and asparaginyl residues in peptides and proteins. *Curr. Top. Cell. Regul.* **1974,** *8* (0), 247-95.

49.   Robinson, N. E.; Robinson, A. B., Deamidation of human proteins. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (22), 12409-12413.

50.   Picotti, P.; Bodenmiller, B.; Mueller, L. N.; Domon, B.; Aebersold, R., Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell* **2009,** *138* (4), 795-806.

4.7 Supplementary Information
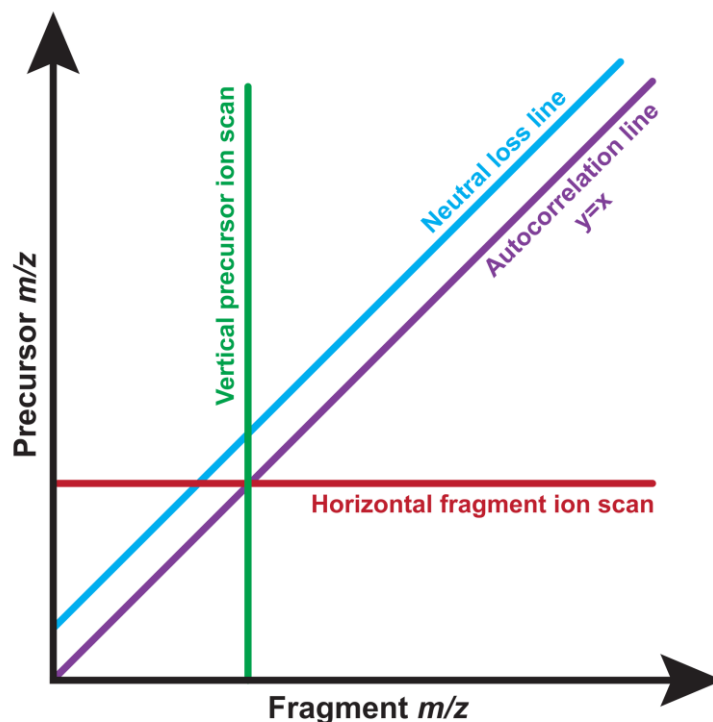
Figure S4.1. Four key 2DMS lines; autocorrelation (purple), neutral loss (cyan), vertical precursor ion scan (green), and horizontal fragment ion scan (red) lines, in a standard two dimensional mass spectrometry (2DMS) spectrum. The auto-correlation line contains the same mass-to-charge (*m/z*) ratio in the x-axis (fragment *m/z* dimension) and the y-axis (precursor *m/z* dimension), representing the fragmented precursor ions in the spectrum. The neutral loss line is produced by precursor ions (autocorrelation line) losing the same side chain group or molecule, such as water (mass difference of 18 Da in the x-axis from the autocorrelation line). The horizontal fragment ion scan line contains fragments from different precursor ions (represented by the number on the y-axis).

Table S4.1 to summarising experimental parameters of 5 combined proteins (5CP) analysis in 2DMS experiments

File Name	Buffer for loading & Elution	Final Solvent Composition for nESI	Acquisition File size (Fragment * Precursor)	Acquisition Length	File size	Tinis Process Time
non-desalted samples (BSA & 5 CP)						
2D_BSA (non-desalt)	N/A	H_2O:ACN:FA 75%:25%:0.1%	256 k * 4 k	55 - 59 minutes (depending on ion accumulation time)	11.7 GB	~ 30 minutes
2D_5CP (non-desalt)						

(Table S4.1. Continue)

File Name	Buffer for loading & Elution	Final Solvent Composition for nESI	Acquisition File size (Fragment * Precursor)	Acquisition Length	File size	Tinis Process Time
Pass through SPE column - one step elution (5 CPs)						
2D_5CP C18 RP pH 2	Load: H_2O/0.1% FA Elute:80% ACN/0.1%FA	H_2O:ACN:FA 75%:25%:0.1%	256 k * 4 k	55 - 59 minutes (based on ion accumulation time)	11.7 GB	~ 30 minutes
2D_5CP C18 RP pH 10	Load: H_2O/NH_4OAc Elute:80% ACN/NH_4OAc	H_2O:ACN:FA 75%:25%:0.1%				
2D_5CP HILIC pH 6.8	Load: ACN/$HCOONH_4$ Elute:80% H_2O/$HCOONH_4$	H_2O:ACN:FA 75%:25%:0.1%				
2D_5CP SAX pH 10	Load: H_2O/ACN/NH_4OH(pH10) Elute: H_2O/ACN/FA(pH 2)	H_2O:ACN:FA 75%:25%:0.1%				

File Name	Buffer for loading & Elution	Final Solvent Composition for nESI	Acquisition File size (Fragment * Precursor)	Acquisition Length	File size	Tinis Process Time
Low Resolution and High Resolution (5 CPs)						
2D_5CP C18 RP pH 2_8192*128k	Load: H$_2$O/0.1% FA Elute: 80%ACN/0.1%FA	H$_2$O:ACN:FA 75%:25%:0.1%	128 k * 8 k	106 minutes	11.7 GB	~ 30 minutes
2D_5CP C18 RP pH 2_8192*1M			1 M * 8 k	167 minutes	93.1 GB	~ 5 hours
2D_5CP C18 RP pH 2_4096*2M			2 M * 4 k	110 minutes	92.9 GB	~ 5 hours
2D_5CP HILIC pH 6.8_8192*128k	Load: ACN/HCOONH$_4$ Elute: 80%H$_2$O/HCOONH$_4$	H$_2$O:ACN:FA 75%:25%:0.1%	128 k * 8 k	106 minutes	11.7 GB	~ 30 minutes
2D_5CP HILIC pH 6.8_8192*1M			1 M * 8 k	167 minutes	93.1 GB	~ 5 hours
2D_5CP HILIC pH 6.8_4096*2M			2 M * 4 k	110 minutes	92.9 GB	~ 5 hours

(Table S4.1 Continue)

File Name	Buffer for loading & Elution	Final Solvent Composition for nESI	Acquisition File size (Fragment * Precursor)	Acquisition Length	File size	Process Time
spanning all	*m/z* **Range Fractionation (5 CPs)**					
2D_5CP C18 RP pH 2_*m/z* fraction - *m/z* 400/600/800 ± 200 - m/z 1200 ± 600	Load: H₂O/0.1% FA Elute: 80%ACN/0.1%FA	H₂O:ACN:FA 75%:25%:0.1%	256 k * 4 k	55 - 59 minutes * 4 (based on ion accumulation time) = ~ 228 minutes	11.7 GB * 4 =46.8 GB	~ 30 minutes (processed in parallel)
2D_5CP HILIC pH 6.8_*m/z* fraction - *m/z* 400/600/800 ± 200 - m/z 1200 ± 600	Load: ACN/HCOONH₄ Elute: 80%H₂O/HCOONH₄	H₂O:ACN:FA 75%:25%:0.1%		55 - 59 minutes * 4 = ~ 228 minutes	11.7 GB * 4 =46.8 GB	~ 30 minutes (processed in parallel)

(Table S4.1 Continue)

File Name	Buffer for loading & Elution	Final Solvent Composition for nESI	Acquisition File size (Fragment * Precursor)	Acquisition Length	File size	Process Time
Offline Fractionation (5 CPs)						
2D_5CP C18 RP pH 2_offline fractionation (9 fractions)	Gradient: see RP gradient (see Figure S4.2)	H$_2$O:ACN:FA 75%:25%:0.1%	256 k * 4 k	55 - 59 minutes * 9 (based on ion accumulation time) = ~ 513 minutes	11.7 GB * 9 =105.3 GB	~ 30 minutes (processed in parallel)
2D_5CP HILIC pH 6.8_offline fractionation (9 fractions)	Gradient: see HILIC gradient (see Figure S4.2)	H$_2$O:ACN:FA 75%:25%:0.1%		55 - 59 minutes * 9 = ~ 513 minutes	11.7 GB * 9 =105.3 GB	~ 30 minutes (processed in parallel)

Table S4.1. Summary of the experimental details of 5 combined proteins (5CP) in the 2DMS experiments. The offline fractionation gradients applied in C18 reserved phase (RP) and hydrophilic interaction chromatography (HILIC) solid phase extraction (SPE) cartridges are shown in the Supplementary Information Figure S4.2A and S4.2B.

Table to summarise all experimental parameters of yeast in 2DMS experiments

File Name	Buffer for loading & Elution	Final Solvent Composition for nESI	Acquisition File size (Fragment * Precursor)	Acquisition Length	File size	Tinis Process Time
Pass through SPE column - one step elution (Yeast)						
2D_Yeast C18 RP pH 2	Load: H_2O/0.1% FA Elute: 80%ACN/0.1%FA	H_2O:ACN:FA 75%:25%:0.1%	1 M * 8 k	167 minutes	93.1 GB	~ 5 hours
Offline Fractionation (Yeast)						
2D_Yeast C18 RP pH 2_ offline fractionation (16 fractions)	Gradient: see RP gradient	H_2O:ACN:FA 75%:25%:0.1%	256 k * 4 k	55 - 59 minutes * 16 (based on ion accumulation time) = ~ 912 minutes	11.7 GB * 16 =187.2 GB	~ 30 minutes (processed in parallel)

Table S4.2. Summary of the experimental details of tryptic digested yeast in the 2DMS experiments. The offline fraction gradient applied in C18 RP SPE cartridge is shown in Supplementary Figure S4.2C.

Rolling Energy Equations Applied by the 12 T solariX

Charge state	Slope	Intercept
1+	0.027	17.044
2+	0.039	-3.558
3+/4+/5+	0.029	-1.045

Table S4.3. Online CAD automatic MS/MS rolling energy parameters applied by 12 T Bruker solariX for nLC MS/MS analysis. Using the values in Table S4.3, the collision energy for 1+ charge state peptides with *m/z* 300 – 1,500 are ranging from 25.1 eV to 57.5 eV, the collision energy for 2+ charge state peptides with *m/z* 300 – 1,500 are ranging from 8.1 eV to 54.9 eV, and the collision energy for the 3+/4+/5+ charge state peptides with *m/z* 300 – 1,500 are ranging from 7.6 eV to 42.5 eV.
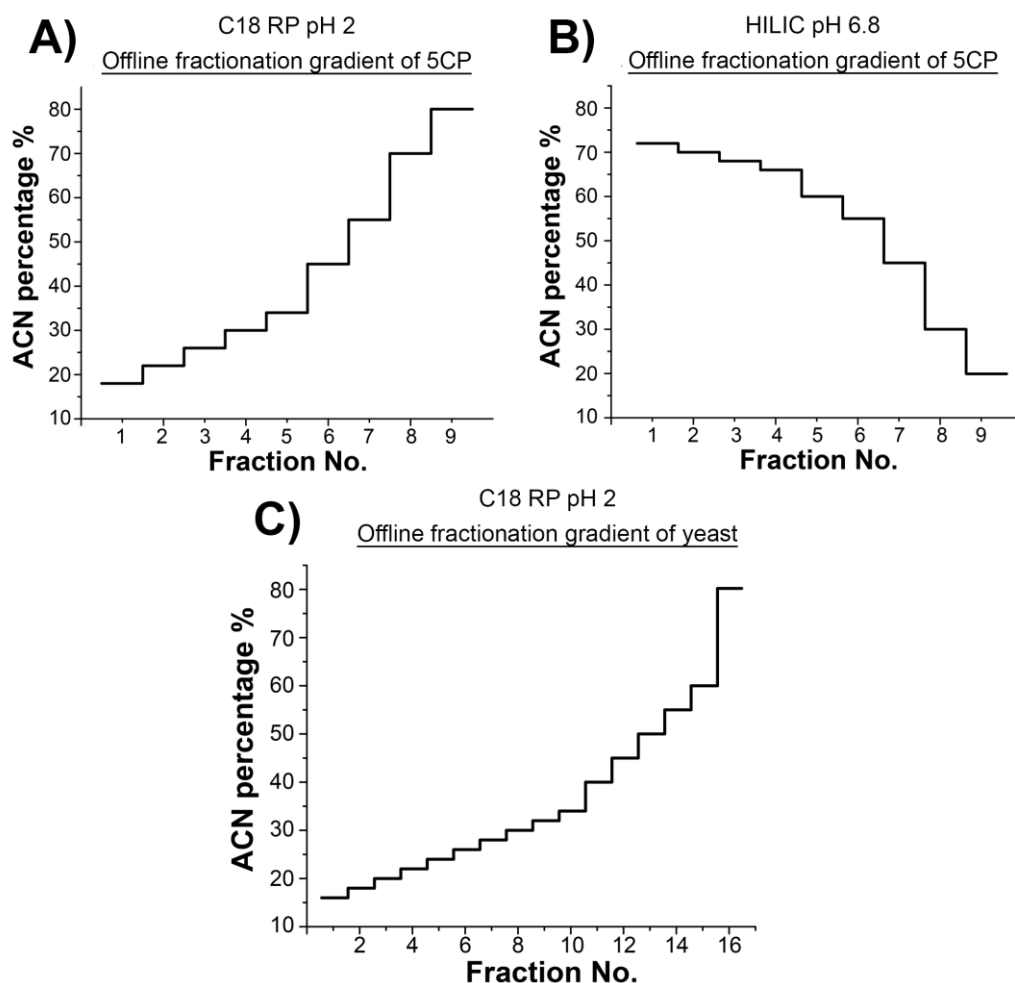
Figure S4.2. Offline fractionation gradients of 5CP using (A) C18 RP SPE cartridge at pH 2 and (B) HILIC SPE cartridge at pH 6.8. The ACN (H_2O) percentages used for 5CP peptides elution in the C18 RP SPE cartridge are 18% (82%), 22% (78%), 26% (74%), 30% (70%), 34% (66%), 45% (55%), 55% (45%), 70% (30%), and 80% (20%); and 9 offline fractions in total were collected. The ACN (H_2O) percentages used for peptides elution of 5CP in the HILIC SPE cartridge are 72% (28%), 70% (30%), 68% (32%), 66% (34%), 60% (40%), 55% (45%), 45% (55%), 30% (70%), and 20% (80%); and 9 offline fractions in total were collected. (C) The offline fractionation gradient graph of yeast using C18 RP SPE cartridge at pH 2. The ACN (H_2O) percentages used for yeast peptides separation are 16% (84%), 18% (82%), 20% (80%), 22% (78%), 24% (76%), 26% (74%), 28% (72%), 30% (70%), 32% (68%), 34% (66%), 40% (60%), 45% (55%),

50% (50%), 55% (45%), 60% (40%), and 80% (20%); and 16 offline fractions in

total were collected.

(A) nLC gradient of BSA and 5CP

Total runtime: 90 minutes

Effective gradient time: 30 minutes

(B) nLC gradient of Yeast

Total runtime: 180 minutes

Effective gradient time: 120 minutes
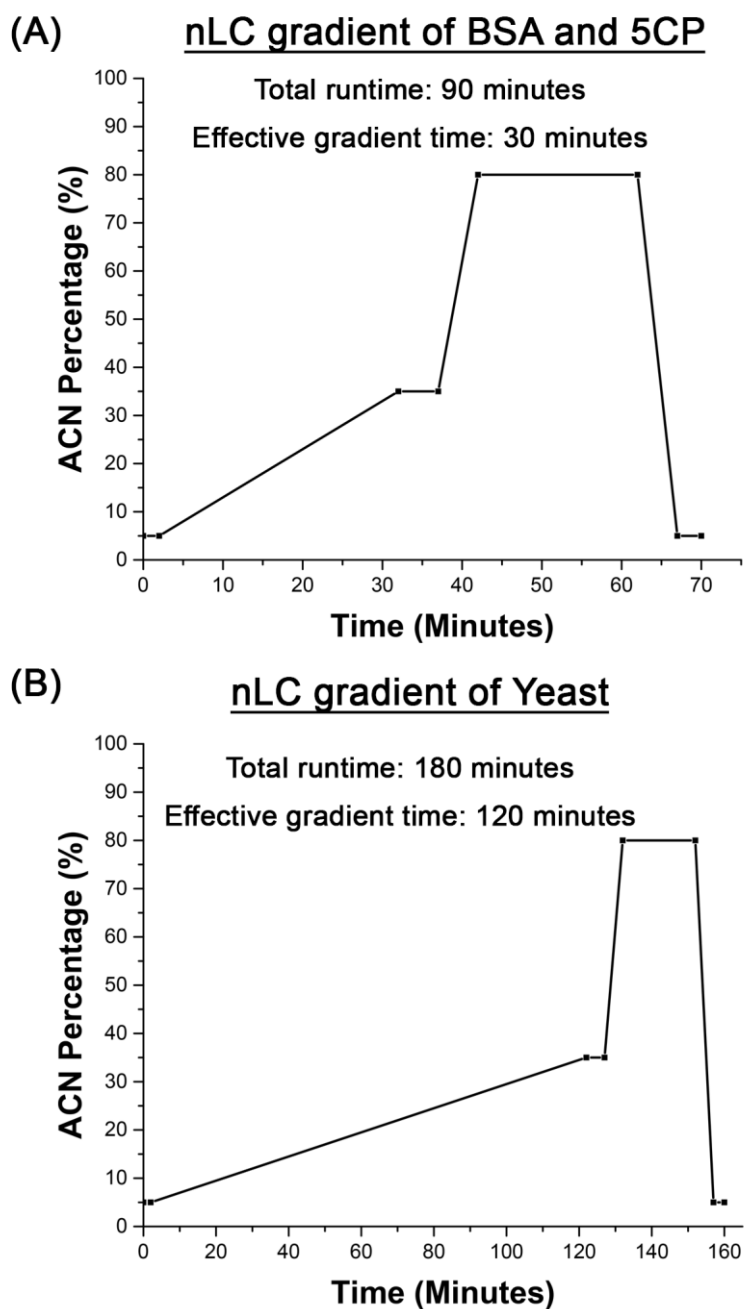
Figure S4.3. The nLC gradient of the tryptic digested (A) BSA and 5CP, as well as (B) yeast. The solvents used for C18 RP nLC columns are H_2O and ACN with 0.1% formic acid. The effective gradients for peptide separation are between 5% and 30% ACN in 30 and 120 minutes for standard proteins and yeast samples respectively, including the time of sample loading and nLC column re-equilibration, the total runtimes are 90 and 120 minutes for the samples respectively.
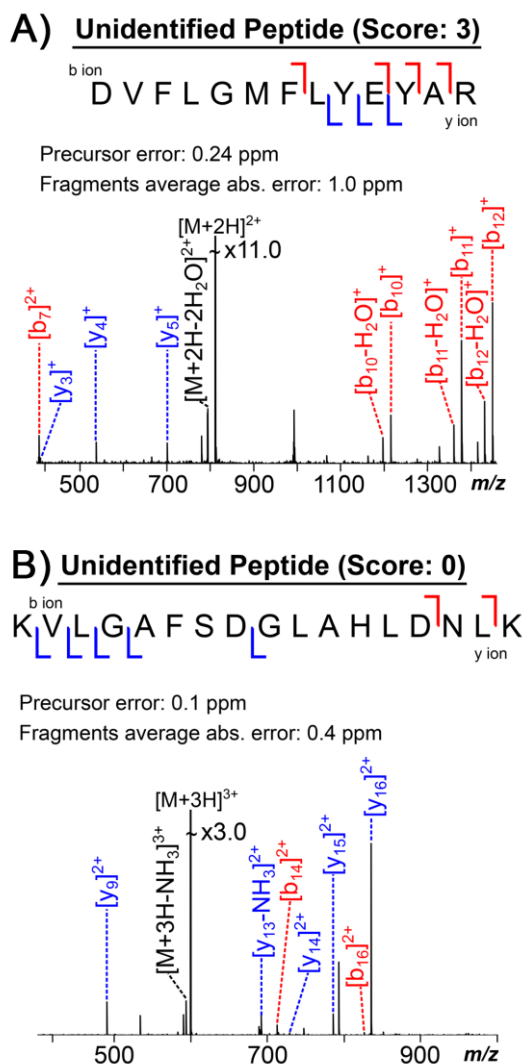
Figure S4.4. 2DMS spectra of (A) DVFLGMFLYEYAR from HSA and (B) KVLGAFSDGLAHLDNLK from HHB, which were classified as unassigned peptides in the MASCOT search. In MASCOT database search, the intensity of fragment ion is one of the key factors which affects the score of peptide assignment, and a minimum of score ~ 13 is required for a confident peptide assignment. In collisionally activated dissociation tandem MS (CAD MS/MS), the intensity of precursor ion decreases while the intensities of fragment ions increase. Thus, the intensities of the fragment ions are similar or even higher than the intensity of precursor ion. In 2DMS spectrum, the intensities of peaks (both precursor and fragment ions) depend on the original intensity as well as the change of magnitude

during fragmentation.[29]   Thus, the intensity of precursor ion is much higher than the intensities of fragment ions in the 2DMS spectrum.   Furthermore, the presence of the strong harmonics in both precursor *m/z* and fragment *m/z* dimension, which are treated as noise peaks in MASCOT searching engine and further reduces the scores of peptides.   Even though the average mass accuracies of precursor and fragment ions are below 1 ppm, the scores for peptides are still quite low.
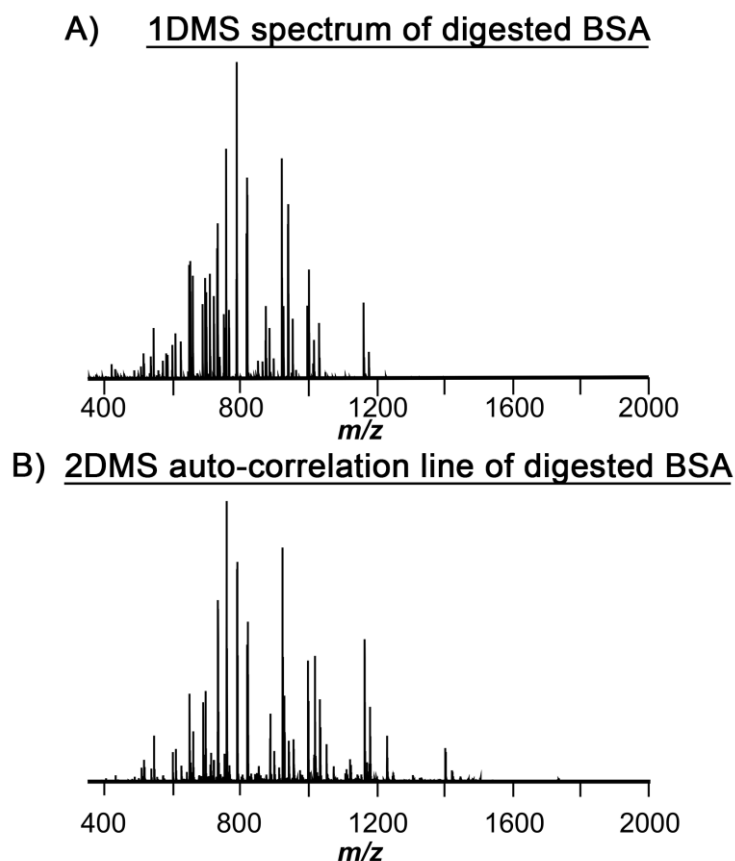
Figure S4.5. Spectra of digested BSA in (A) an 1DMS spectrum and (B) an extraction of auto-correlation line in the 2DMS spectrum. The auto-correlation line is the most closely compacted scan in the 2DMS spectrum, a high resolution scan, therefore, is required. For the 1DMS spectrum, the data were acquired with 4 M data-points, 10 spectra accumulation; and the acquisition time was ~ 1 minute. The resolution of m/z 449.8 in the 1DMS spectrum is ~ 351 K, which is 14 times higher than the resolution of the 2DMS with 256 k data-points on the fragment m/z dimension (see Figure 4.2B). In order to obtain a similar resolution (4 M) in the fragment m/z dimension, a 3-hour 2DMS experiment would be required. The precursor ions (on the auto-correlation line) of digested BSA in the 2DMS spectrum are similar to the 1DMS spectrum of the digested BSA, thus the peaks in 1DMS can be used as the accurate masses of precursor ions for database searches for 2DMS.
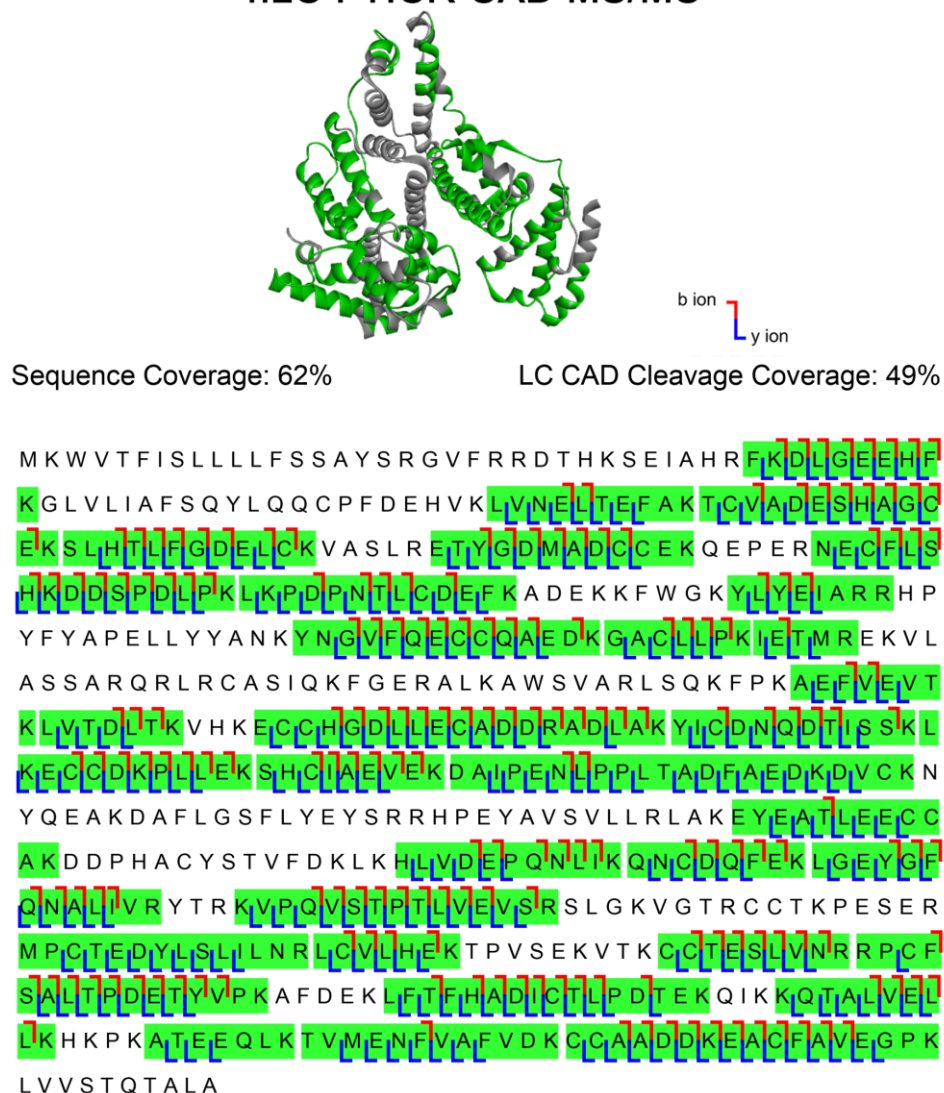
Figure S4.6. Fragmentation map of the digested BSA with 62% sequence coverage and 49% cleavage coverage using nLC FTICR CAD MS/MS. The peptides assigned by nLC MS/MS are highlighted in green.
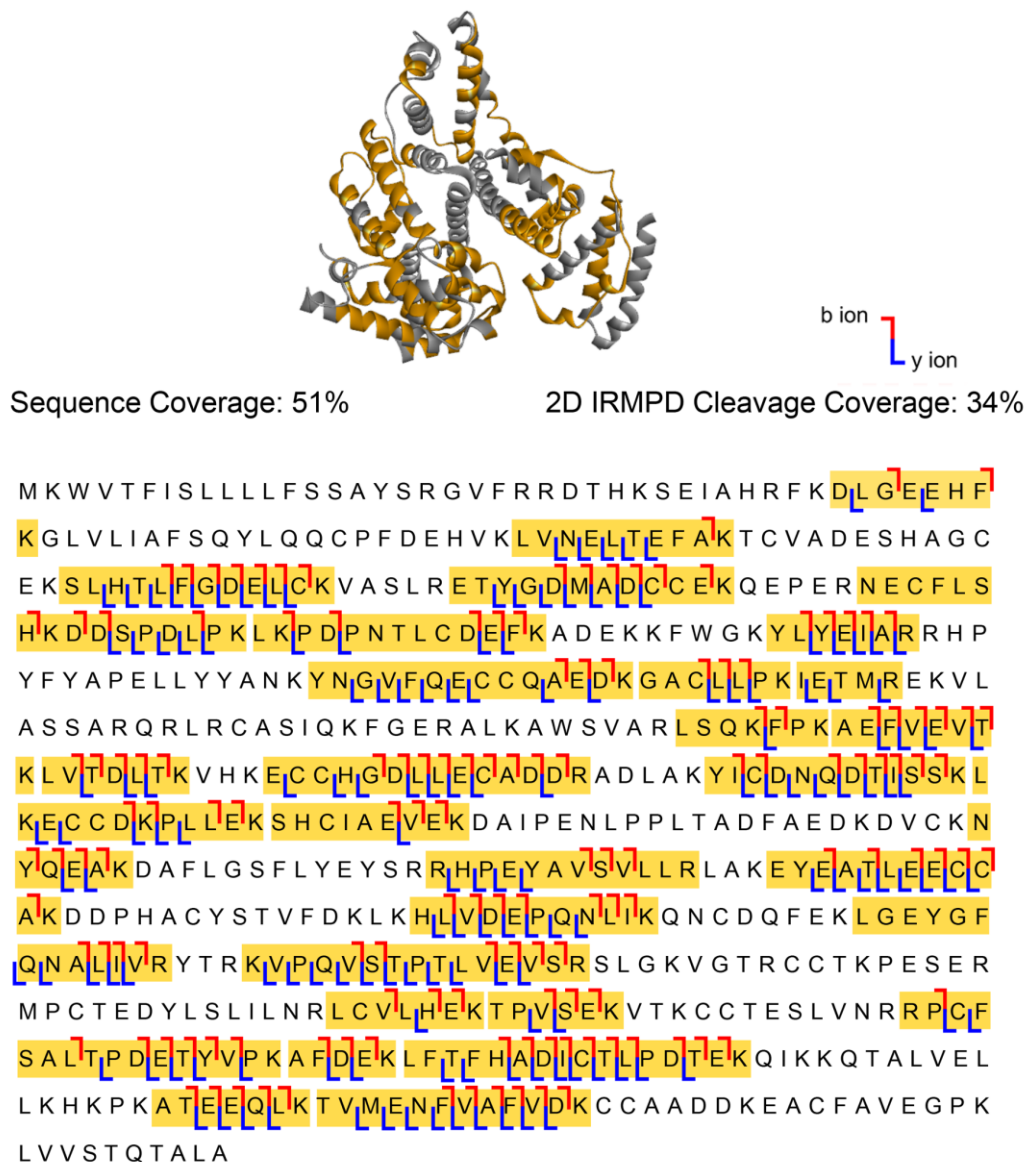
Figure S4.7. Fragmentation map of the digested BSA with 51% sequence coverage and 34% cleavage coverage using IRMPD 2DMS. The peptides assigned by 2DMS are highlighted in yellow.

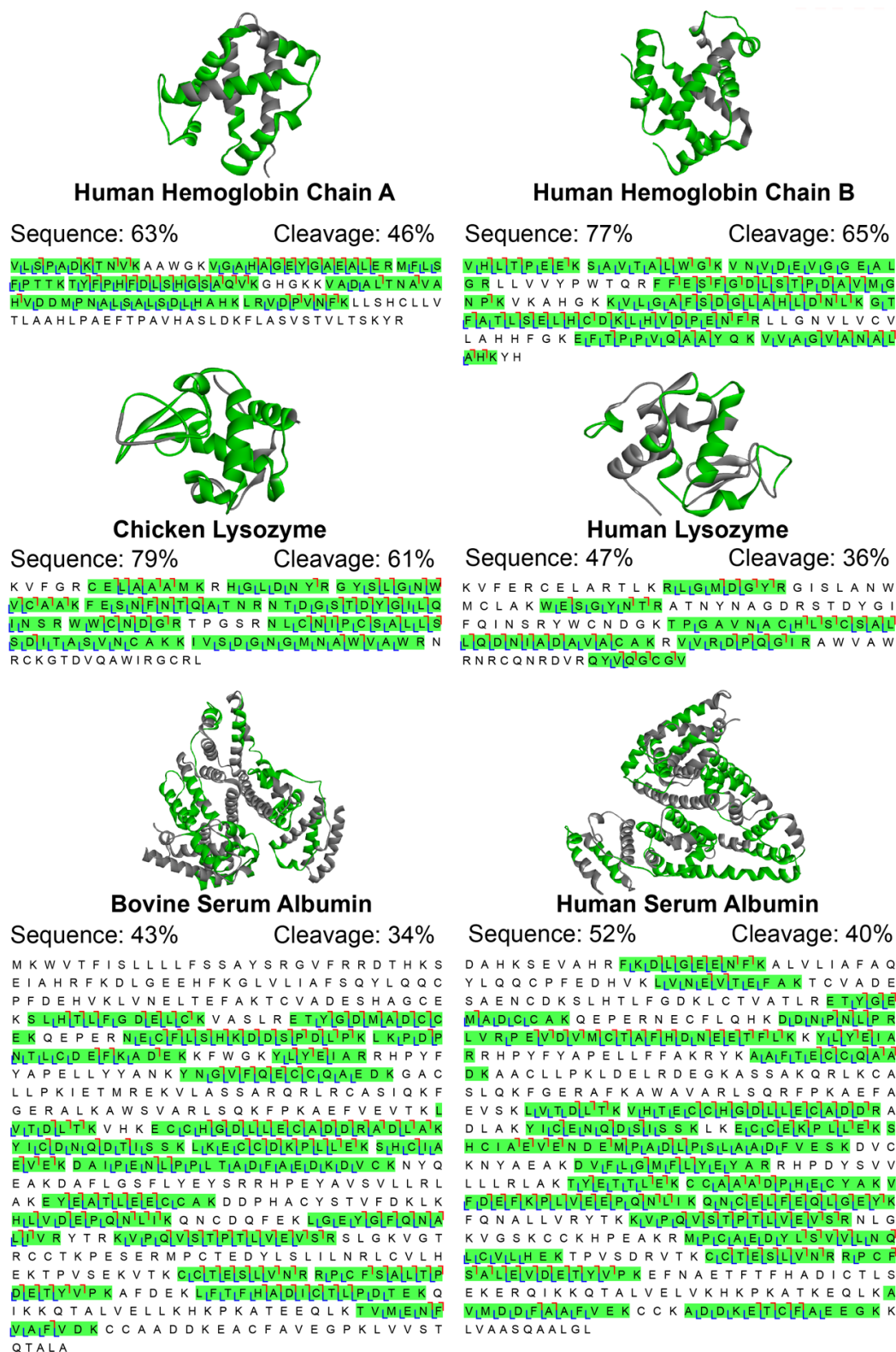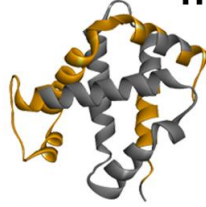# Fragmentation map of digested 5CP using nLC FTICR CAD MS/MS

b ion
y ion

Human Hemoglobin Chain A

Sequence: 63%   Cleavage: 46%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S
F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A
H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V
T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B

Sequence: 77%   Cleavage: 65%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L
G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G
N P K V K A H G K K V L G A F S D G L A H L D N L K G T
F A T L S E L H C D K L H V D P E N F R L L G N V L V C V
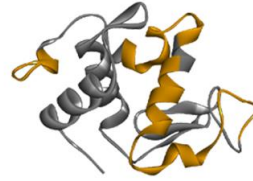L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L
A H K Y H

Chicken Lysozyme

Sequence: 79%   Cleavage: 61%

K V F G R C E L A A A M K R H G L L D N Y R G Y S L G N W
V C A A K F E S N F N T Q A T N R N T D G S T D Y G I L Q
I N S R W W C N D G R T P G S R N L C N I P C S A L L S
S D I T A S V N C A K K I V S D G N G M N A W V A W R N
R C K G T D V Q A W I R G C R L

Human Lysozyme

Sequence: 47%   Cleavage: 36%

K V F E R C E L A R T L K R L L G M D G Y R G I S L A N W
M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I
F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L
L Q D N I A D A V A C A K R V V R D P Q G I R A W V A W
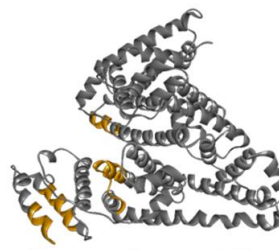R N R C Q N R D V R Q Y V Q G C G V

Bovine Serum Albumin

Sequence: 43%   Cleavage: 34%

M K W V T F I S L L L L F S S A Y S R G V F R R D T H K S
E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C
P F D E H V K L V N E L T E F A K T C V A D E S H A G C E
K S L H T L F G D E L C K V A S L R E T Y G D M A D C C
E K Q E P E R N E C F L S H K D D S P D L P K L K P D P
N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F
Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C
L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F
G E R A L K A W S V A R L S Q K F P K A E F V E V T K L
V T D L T K V H K E C C H G D L L L E C A D D R A D L A K
Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A
E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q
E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L
A K E Y E A T L E E C C A K D D P H A C Y S T V F D K L K
H L V D E P Q N L I K Q N C D Q F E K L G L E Y G F Q N A
L I V R Y T R K V P Q V S T P T L V E V S R S L G K V G T
R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H
E K T P V S E K V T K C C T E S L V N R R P C F S A L T F
D E T Y V P K A F D E K L F T F H A D I C T L P D T E K Q
I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F
V A F V D K C C A A D D K E A C F A V E G P K L V V S T
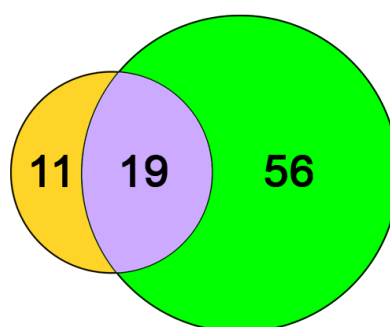Q T A L A

Human Serum Albumin

Sequence: 52%   Cleavage: 40%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q
Y L Q Q C P F E D H V K L V N E V T E F A K T C V A D E
S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E
M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R
L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A
R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A
D K A A C L L P K L D E L R D E G K A S S A K Q R L K C A
S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A
E V S K L V T D L T K V H T E C C H G D L L E C A D D R A
D L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S
H C I A E V E N D E M P A D L P S L A A D F V E S K D V C
K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V
L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V
F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K
F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G
K V G S K C C K H P E A K R M P C A E D Y L S V V L L N Q
L C V L H E K T P V S D R V T K C C T E S L V N R R P C F
S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S
E K E R Q I K K Q T A L V E L V K H K P K A T K E Q L K A
V M D D F A A F V E K C C K A D D K E T C F A E E G K K
L V A A S Q A A L G L

Figure S4.8. Fragmentation maps of the digested 5CP using nLC FTICR CAD MS/MS. The peptides assigned by nLC MS/MS are highlighted in green.

A) Fragmentation map of digested 5CP using IRMPD 2DMS

(Figure S4.9. Continue)

B) Venn Diagram of peptides identifed by 2DMS and nLC MS/MS

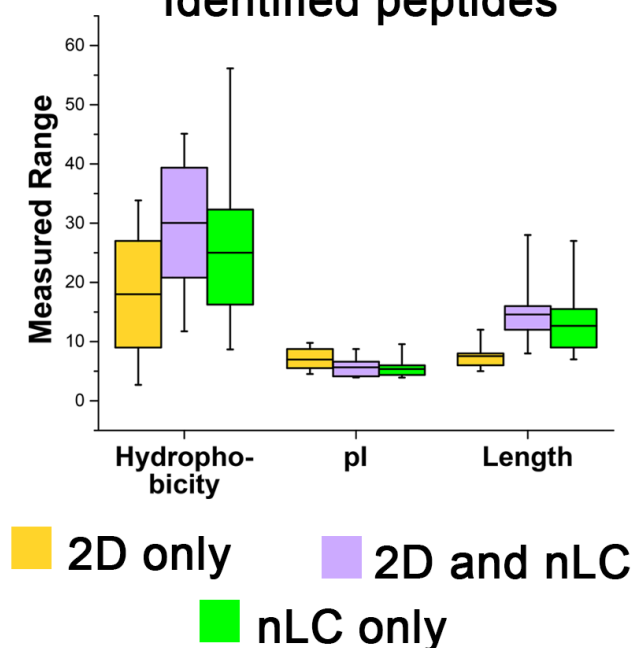## C) Physical Properties of Identified peptides

Figure S4.9. 2DMS results of 5CP without prior purification and separation. (A) Fragmentation maps of the digested 5CP using IRMPD 2DMS. The peptides assigned by 2DMS are highlighted in yellow. (B) Venn diagram showing the difference of peptides identified by 2DMS and nLC MS/MS. (C) Box chart showing the physical properties range of hydrophobicity, pI, and

347

length of the peptides identified by 2DMS and nLC MS/MS.   The number of peptides being identified by 2DMS is less than nLC MS/MS; however, 11 out of 86 peptides (13%) are exclusively detected by 2DMS.   The box chart further shows the peptides identified by 2DMS are more hydrophilic, basic, and shorter than those in nLC MS/MS.
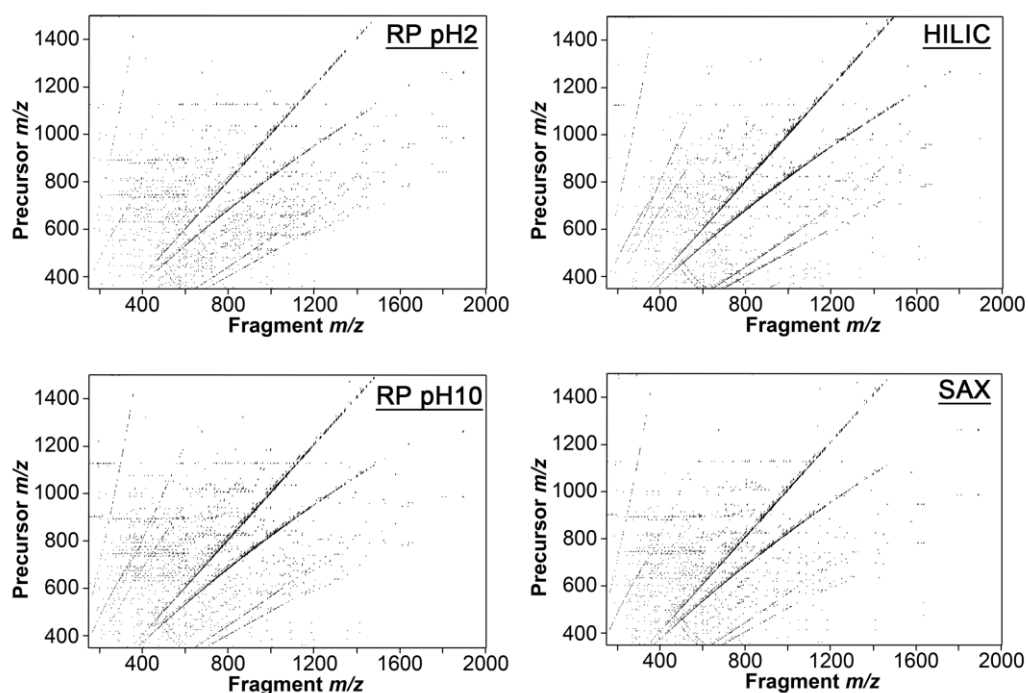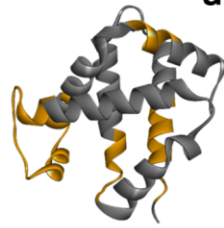
Figure S4.10. 2DMS spectra (256 k * 4 k) of digested 5CP enriched using SPE

cartridges of C18 RP at pH 2, HILIC at pH 6.8, C18 RP at pH 10, and SAX at pH

10.

A) Fragmentation map of digested 5CP (with C18 RP pH 2 enrichment) using IRMPD 2DMS

b ion
y ion

Human Hemoglobin Chain A

Sequence: 36%   Cleavage: 18%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B
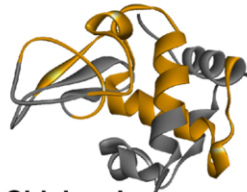
Sequence: 61%   Cleavage: 33%

VHLTPEEKSAVTALWGKVNVDEVGGEAL
GRLLVVYPWTQRFFESFGDLSTPDAVMG
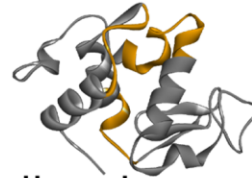NPKVKAHGKKVLGAFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFTPPVQAAYQKVVAGVANAL
AHKYH

Chicken Lysozyme

Sequence: 48%   Cleavage: 28%

KVFGRCELAAAMKRHGLDNYRGYSLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILQ
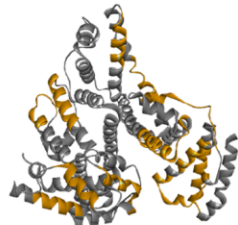INSRWWCNDGRTPGSRNLCNIPCSALLS
SDITASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL

Human Lysozyme

Sequence: 23%   Cleavage: 11%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGYNTRATNYNAGDRSTDYGI
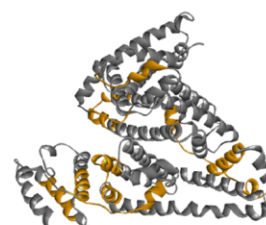FQINSRYWCNDGKTPGAVNACHLSCSAL
LQDNIADAVACAKRVVRDPQGIRAWVAW
RNRCQNRDVRQYVQGCGV

Bovine Serum Albumin

Sequence: 26%   Cleavage: 14%

MKWVTFISLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNELTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMLADCC
EKQEPERNECFLSHKDDSPDLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVELVTKL
VTDLTKVHKECCHGDLLECADDRADLAK
YICDNQDTISSKLKECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEECAKDDPHACYSTVFDKLK
HLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin

Sequence: 25%   Cleavage: 10%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVNEVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
EVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSISSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHKPKATKEQLKA
VMDDFAAFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

(Figure S4.11. Continue)
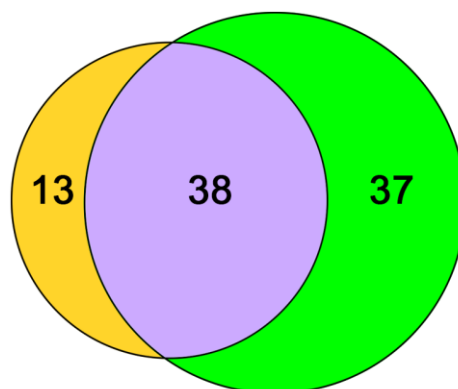
Figure S4.11. 2DMS results of 5CP with C18 RP pH 2 SPE cartridge

purification. (A) Fragmentation maps of digested 5CP with peptide enrichment

using 2DMS with C18 RP SPE cartridge enrichment. The peptides assigned by

2DMS are highlighted in yellow. (B) Venn diagram showing the difference of

peptides identified by 2DMS and nLC MS/MS.   (C) Box chart showing the physical properties range of hydrophobicity, pI, and length of the peptides identified by 2DMS and nLC MS/MS.   Compared to the non-enriched 2DMS experiment (see Supplementary Information, Figure S4.9), the number of unique peptide detected by 2DMS increases by 2 only; while the number of common peptide detected by commonly between 2DMS and nLC MS/MS has been doubled.   Furthermore, the hydrophobicity range of the exclusive peptides detected by 2DMS contains similar average hydrophobicity the peptides being detected by nLC MS/MS.   This result suggests more peptides can be observed by applying C18 RP pH 2 SPE cartridge enrichment as it helps to remove non-peptide peaks in the spectrum and enhance the intensities of peptides. However, the peptides that can be observed in the 2DMS after RP enrichment containing similar physical properties to those peptides being detected in nLC MS/MS as RP-LC has been applied in both methods, in which the hydrophilic peptides are less likely to retain on the SPE RP cartridge.

A) Fragmentation map of digested 5CP (with HILIC enrichment) using IRMPD 2DMS

(Figure S4.12. Continue)

B) Venn Diagram of peptides identifed by 2DMS and nLC MS/MS
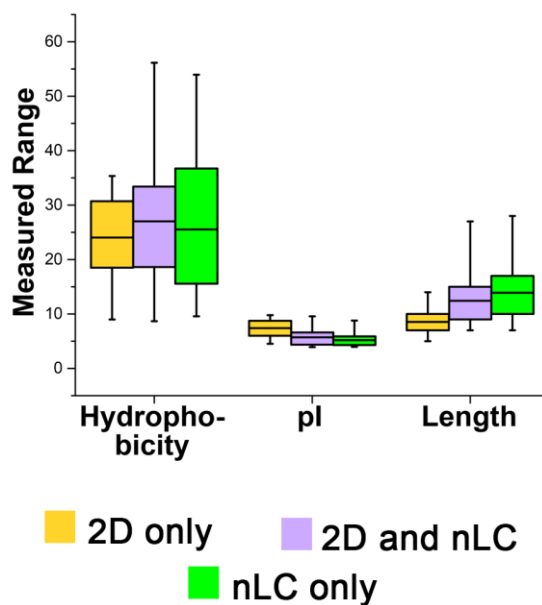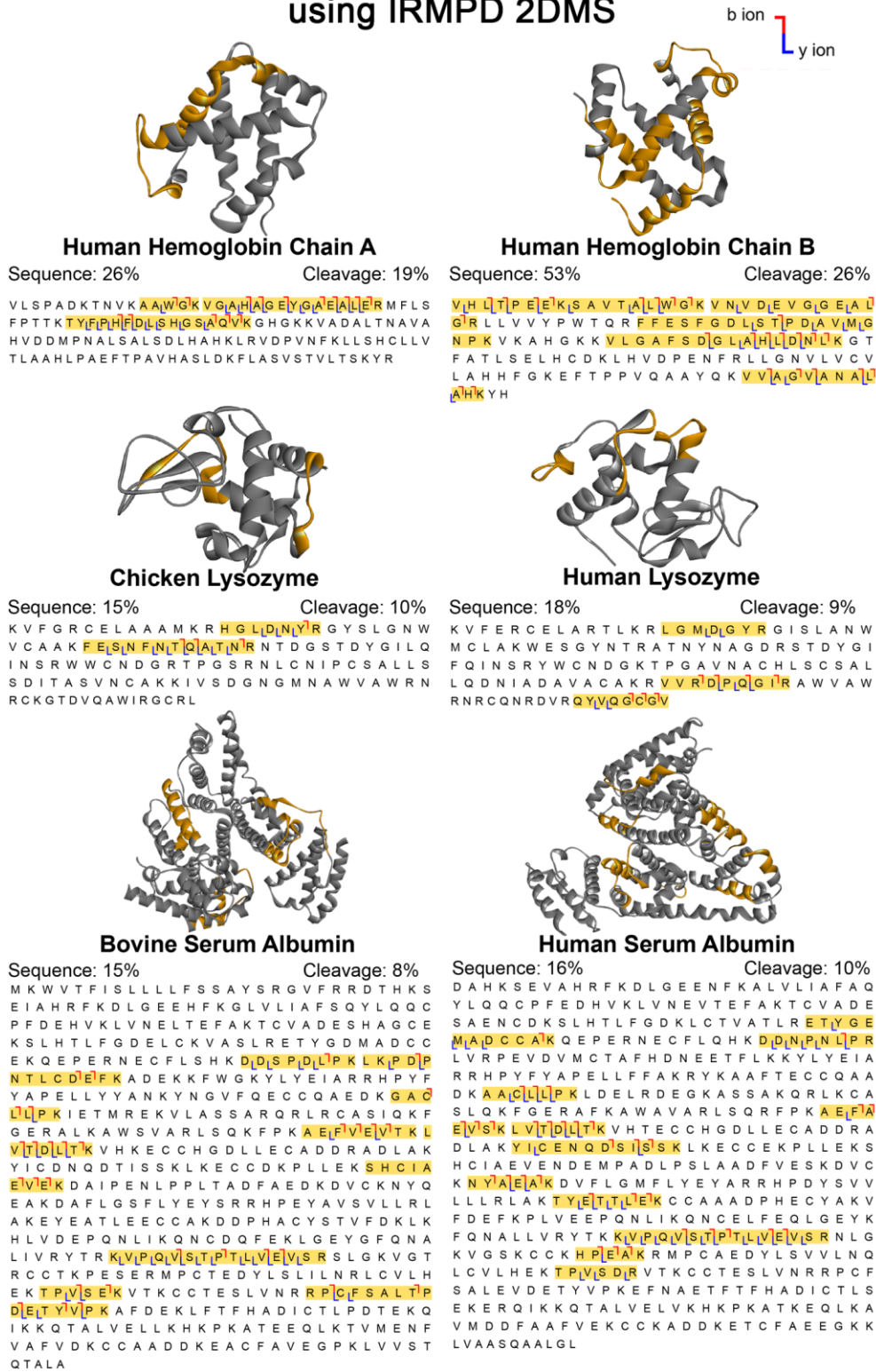

C) Physical Properties of Identified peptides

Figure S4.12. 2DMS results of 5CP with HILIC SPE cartridge purification. (A)

Fragmentation maps of digested 5CP with peptide enrichment using HILIC SPE

354

cartridge at pH 6.8 and then analysed by IRMPD 2DMS. The peptides assigned by 2DMS are highlighted in yellow. (B) Venn diagram showing the difference of peptides b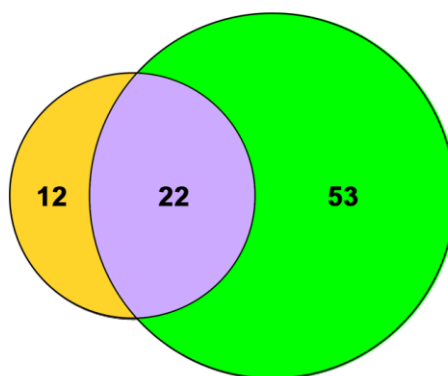eing identified by 2DMS and nLC MS/MS. (C) Box chart showing the physical properties range of hydrophobicity, pI, and length of the peptides identified by 2DMS and nLC MS/MS. The number of peptides identified in the HILIC enriched 2DMS is similar to the number of peptides that has been identified in non-enriched 5CP sample (see Supplementary Information, Figure S4.9). The hydrophobicity of the peptides being exclusively identified in the HILIC enriched 2DMS is much lower than the peptides that identified in nLC MS/MS. Despite the number of peptides being identified in the 2DMS (HILIC enrichment) does not increase sharply, the physical properties of the identified peptides are significantly different from those being detected in nLC MS/MS which suggests combining the HILIC enrichment 2DMS with nLC MS/MS can enhance the sequence and cleavage coverage of the proteomics sample.

A) Fragmentation map of digested 5CP (with C18 RP pH 10 enrichment) using IRMPD 2DMS

(Figure S4.13. Continue)

B) Venn Diagram of peptides identifed by 2DMS and nLC MS/MS
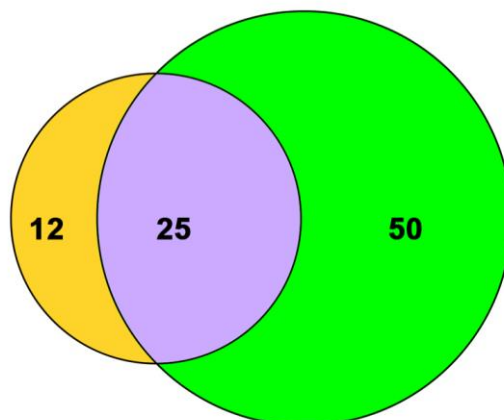
C) Physical Properties of Identified peptides
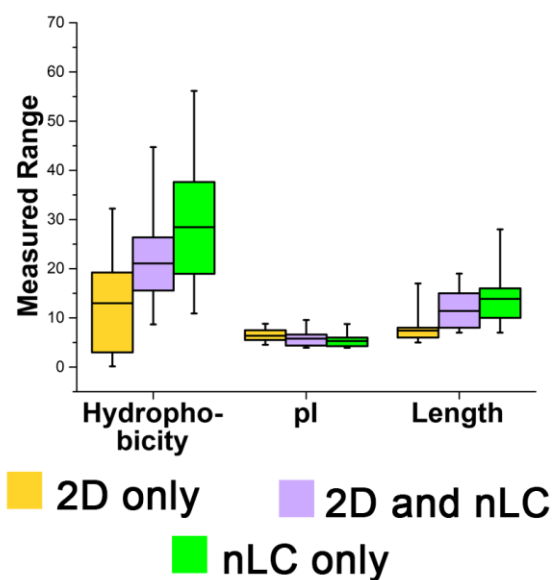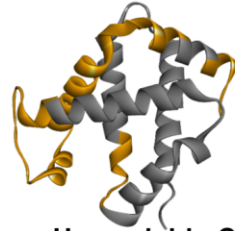
Figure S4.13. 2DMS results of 5CP with C18 RP pH 10 SPE cartridge purification. (A) Fragmentation maps of digested 5CP with peptide enrichment using 2DMS with C18 RP SPE cartridge enrichment at pH 10. The peptides

357

assigned by 2DMS are highlighted in yellow. (B) Venn diagram showing the difference of peptides being identified by 2DMS and nLC MS/MS. (C) Box chart to show the physical properties range of hydrophobicity, pI, and length of the peptides that were identified by 2DMS and nLC MS/MS. Similar observations in the physical properties of the peptides being identified using RP SPE cartridge at pH 2 and pH 10, but the number of peptides identified by RP SPE cartridge after pH 10 is less than the number of identified peptides at pH 2 (see Supplementary Information, Figure S4.11).

A) Fragmentation map of digested 5CP (with SAX enrichment) using IRMPD 2DMS

b ion
y ion
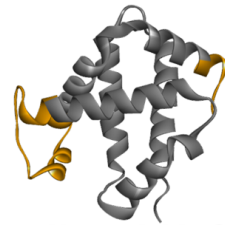
Human Hemoglobin Chain A

Sequence: 23%   Cleavage: 14%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S
F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A
H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V
T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B

Sequence: 23%   Cleavage: 11%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L
G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G L
N P K V K A H G K K V L G A F S D G L A H L D N L K G T
F A T L S E L H C D K L H V D P E N F R L L G N V L V C V
L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L
A H K Y H

Chicken Lysozyme

Sequence: 42%   Cleavage: 24%

K V F G R C E L A A A M K R H G L L D N Y R G Y S L G N W
V C A A K F E S N F N T Q A T N R N T D G S T D Y G I L Q
I N S R W W C N D G R T P G S R N L C N I P C S A L L S
S D I T A S V N C A K K I V S D G N G M N A W V A W R N
R C K G T D V Q A W I R G C R L
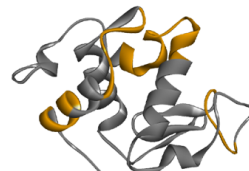
Human Lysozyme

Sequence: 26%   Cleavage: 12%

K V F E R C E L A R T L K R L G M D G Y R G I S L A N W
M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I
F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L
L Q D N I A D A V A C A K R V V R D P Q G I R A W V A W
R N R C Q N R D V R Q Y V Q G C G V
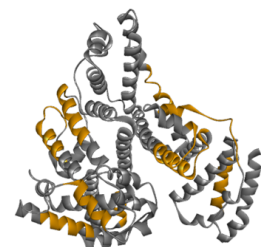
Bovine Serum Albumin

Sequence: 21%   Cleavage: 11%

M K W V T F I S L L L L L F S S A Y S R G V F R R D T H K S
E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C
P F D E H V K L V N E L T E F A K T C V A D E S H A G C E
K S L H T L F G D E L C K V A S L R E T Y G D M A D C C
E K Q E P E R N E C F L S H K D D S P D L P K L K P D P
N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F
Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C
L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F
G E R A L K A W S V A R L S Q K F P K A E F V E V T K L
V T D L L T K V H K E C C H G D L L E C A D D R A D L L A K
Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A
E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q
E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L
A K E Y E A T L E E C C A K D D P H A C Y S T V F D K L K
H L V D E P Q N L I K Q N C D Q F E K L G E Y G F Q N A
L I V R Y T R K V P Q V S T P T L V E V S R S L G K V G T
R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H
E K T P V S E K V T K C C T E S L V N R R P C F S A L T P
D E T Y V P K A F D E K L F T F H A D I C T L P D T E K Q
I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F
V A F V D K C C A A D D K E A C F A V E G P K L V V S T
Q T A L A

Human Serum Albumin

Sequence: 10%   Cleavage: 6%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q
Y L Q Q C P F E D H V K L V N E V T E F A K T C V A D E
S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E
M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R
L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A
R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A
D K A A C L L P K L D E L R D E G K A S S A K Q R L K C A
S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A
E V S K L V T D L L T K V H T E C C H G D L L E C A D D R A
D L L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S
H C I A E V E N D E M P A D L P S L A A D F V E S K D V C
K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V
L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V
F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K
F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G
K V G S K C C K H P E A K R M P C A E D Y L S V V L N Q
L C V L H E K T P V S D R V T K C C T E S L V N R R P C F
S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S
E K E R Q I K K Q T A L V E L V K H K P K A T K E Q L K A
V M D D F A A F V E K C C K A D D K E T C F A E E G K K
L V A A S Q A A L G L

(Figure S4.14. Continue)

B) Venn Diagram of
peptides identifed by
2DMS and nLC MS/MS
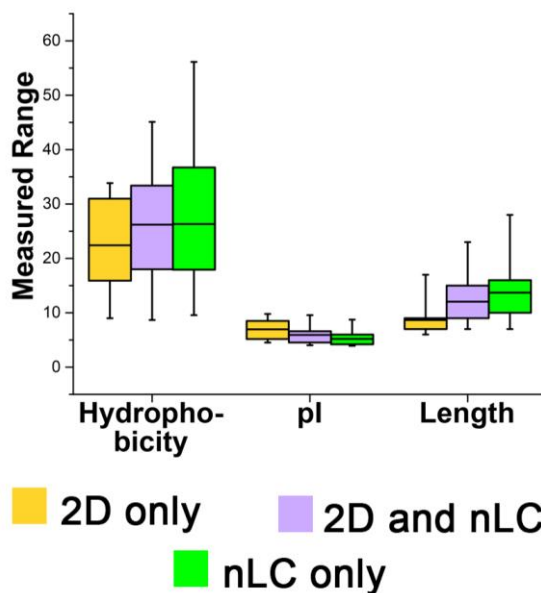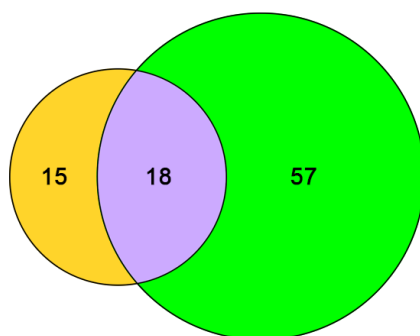

C) Physical Properties of
Identified peptides
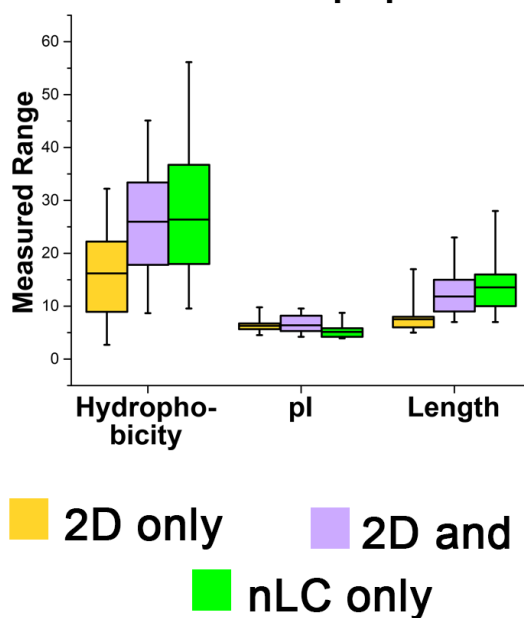
Figure S4.14. 2DMS results of 5CP with SAX SPE cartridge purification. (A)

Fragmentation maps of digested 5CP with peptide enrichment using 2DMS with

SAX SPE cartridge enrichment. The peptides assigned by 2DMS are highlighted

in yellow.   (B) Venn diagram showing the difference of peptides being identified by 2DMS and nLC MS/MS.   (C) Box chart diagram showing the physical properties range of hydrophobicity, pI, and length of the peptides that were identified by 2DMS and nLC MS/MS.   The enrichment with SAX cartridge shows the highest number of exclusive peptides being detected by 2DMS which suggests complementary data can be obtained when sample is analysed by SAX enriched 2DMS as well as nLC MS/MS.

Combined fragmentation map of digested 5CP (with 4 SPEs enrichment) using IRMPD 2DMS

Identified by >1 columns
HILIC only
RP at pH 2 only
SAX only
RP at pH 10 only

b ion
y ion

Human Hemoglobin Chain A

Sequence: 52%   Cleavage: 34%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S
F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A
H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V
T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B

Sequence: 60%   Cleavage: 43%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L
G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G
N P K V K A H G K K V L L G A F S D G L A H L D N L K G T
F A T L S E L H C D K L H V D P E N F R L L G N V L V C V
L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L
A H K Y H
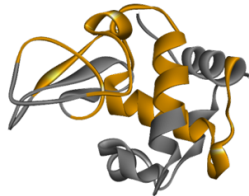
Chicken Lysozyme

Sequence: 48%   Cleavage: 34%

K V F G R C E L A A A M K R H G L D N L Y R G Y S L G N W
V C A A K F E S N F N T Q A T N R N T D G S T D Y G I L Q
I N S R W W C N D G R T P G S R N L C N I P C S A L L S
S D I T A S V N C A K K I V S D G N G M N A W V A W R N
R C K G T D V Q A W I R G C R L

Human Lysozyme

Sequence: 38%   Cleavage: 21%

K V F E R C E L A R T L K R L G M D G Y R G I S L A N W
M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I
F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L
L Q D N I A D A V A C A K R V V R D P Q G I R A W V A W
R N R C Q N R D V R Q Y V Q G C G V
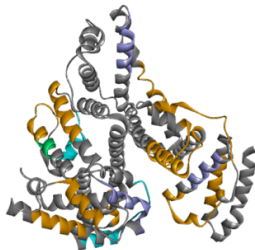
Bovine Serum Albumin

Sequence: 34%   Cleavage: 20%

M K W V T F I S L L L L L F S S A Y S R G V F R R D T H K S
E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C
P F D E H V K L V N E L L T E F A K T C V A D E S H A G C E
K S L H T L F G D E L C K V A S L R E T Y G D M A D C C
E K Q E P E R N E C F L S H K D D S P D L L P K L K P D P
N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F
Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C
L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F
G E R A L K A W S V A R L S Q K F P K A E F V E V T K L
V T D L T K V H K E C C H G D L L E C A D D R A D L A K
Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A
E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q
E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L
A K E Y E A T L L E E C C A K D D P H A C Y S T V F D K L K
H L L V D E P Q N L I K Q N C D Q F E K L G E Y G F Q N A
L I V R Y T R K V P Q V S T P T L L V E V S R S L G K V G T
R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H
E K T P V S E K V T K C C T E S L V N R R P C F S A L T P
D E T Y V P K A F D E K L F T F H A D I C T L P D T E K Q
I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F
V A F V D K C C A A D D K E A C F A V E G P K L V V S T
Q T A L A

Human Serum Albumin

Sequence: 39%   Cleavage: 20%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q
Y L Q Q C P F E D H V K L V N E L V T E F A K T C V A D E
S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E
M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R
L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A
R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A
D K A A C L L L P K L D E L R D E G K A S S A K Q R L K C A
S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A
E V S K L V T D L T K V H T E C C H G D L L E C A D D R A
D L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S
H C I A E V E N D E M P A D L P S L A A D F V E S K D V C
K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V
L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V
F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K
F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G
K V G S K C C K H P E A K R M P C A E D Y L S V V L N Q
L C V L H E K T P V S D R V T K C C T E S L V N R R P C F
S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S
E K E R Q I K K Q T A L V E L V K H K P K A T K E Q L K A
V M D D F A A F V E K C C K A D D K E T C F A E E G K K
L V A A S Q A A L G L

Figure S4.15.   The combined fragmentation maps of 5CP using 4 different SPE

cartridges enrichment and then analysed by 2DMS.
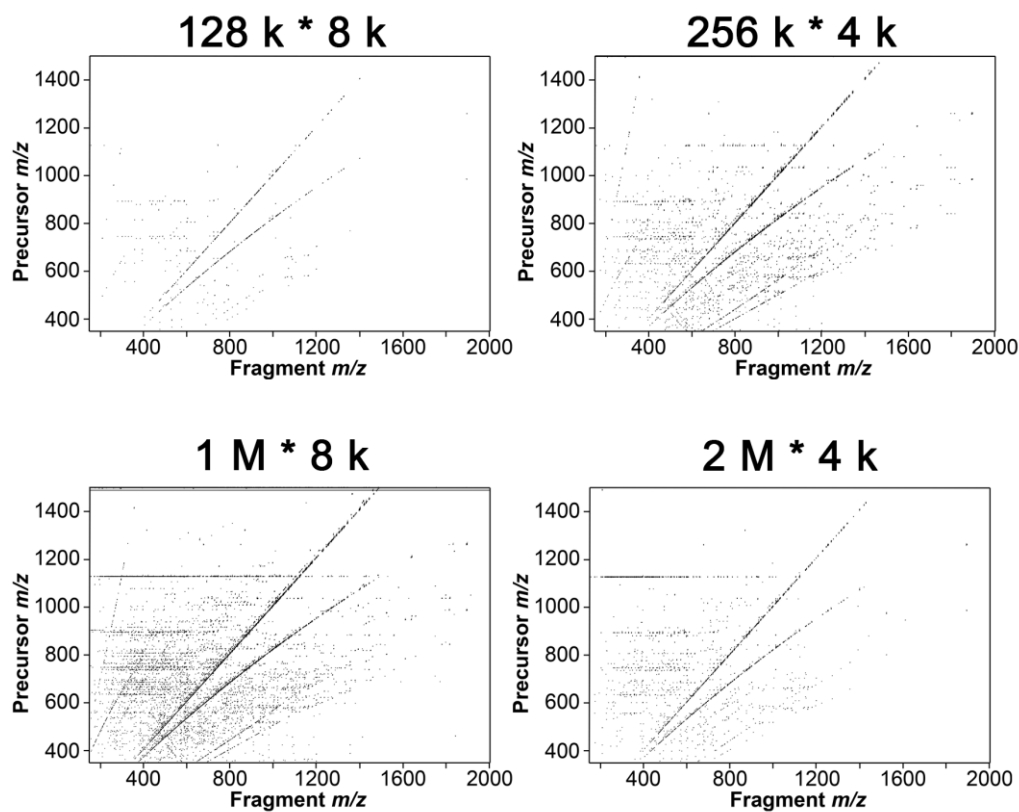
Figure S4.16. 2DMS spectra of the digested 5CP with C18 RP pH 2

enrichment conducting in various resolutions of fragment and precursor *m/z*

dimensions. The 2DMS contour plots are displayed in same intensity.

Fragmentation map of digested 5CP (RP pH 2 enriched)

2DMS (128 k * 8 k)

b ion
y ion
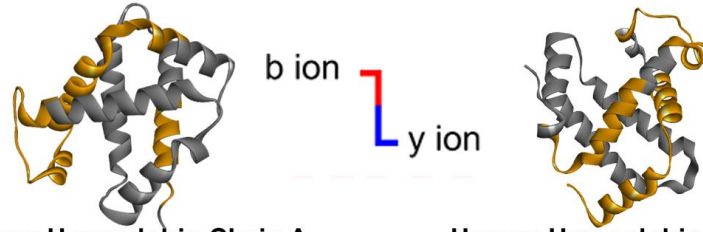
Human Hemoglobin Chain A

Sequence: 40% Cleavage: 20%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B

Sequence: 45% Cleavage: 25%

VHLTPEEKSAVTALLWGKVNVDEVGGEALG
GRLLVVYPWTQRFFESFGDLSTPDAVMG
NPKVKAHGKKVLGAFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFTPPVQAAYQKVVAGVANAL
AHKYH
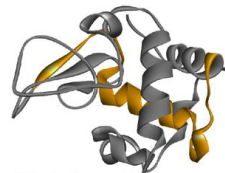
Chicken Lysozyme

Sequence: 24% Cleavage: 11%

KVFGRCELAAAMKRHGLDNYRGYSLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILQ
INSRWWCNDGRTPGSRNLCNIPCSALLS
SDITASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL
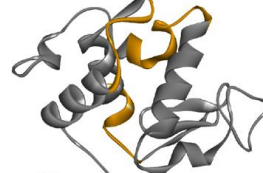
**Human Lysozyme**

Sequence: 21% Cleavage: 8%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGYNTRATNYNAGDRSTDYGI
FQINSRYWCNDGKTPGAVNACHLSCSAL
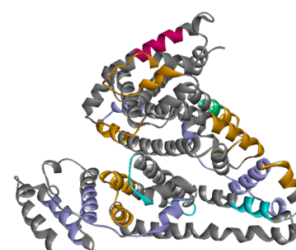LQDNIADAVACAKRVVRDPLQGIRAWLVAW
RNRCQNRDVRQYVQGCGV
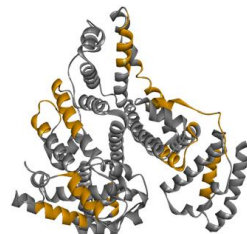
Bovine Serum Albumin

Sequence: 20% Cleavage: 8%

MKWVTFISLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNELTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMADCC
EKQEPERNECFLSHKDDSPDLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVEVTKL
VTDLTKVHKECCHGDLLECADDRADLAK
YICDNQDTISSKLKECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEECCAKDDPHACYSTVFDKLK
HLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin

Sequence: 15% Cleavage: 7%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVINELVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
EVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSISSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHKPKATKEQLKA
VMDDFAALFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

(Figure S4.17. Continue)

2DMS (256 k * 4 k)

b ion

y ion

Human Hemoglobin Chain A

Sequence: 36%   Cleavage:18%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B

Sequence: 61%   Cleavage: 33%

VHLTPEEKSAVTALWGKVNVDEVGGEALI
GRLLVVYPWTQRFFESFGDLSTPDAVMG
NPKVKAHGKKVLGAFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFTPPVQAAYQKVVAGVANAL
AHKYH

Chicken Lysozyme

Sequence: 48%   Cleavage: 28%

KVFGRCELAAAMKRHGLDNYRGYSLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILQ
INSRWWCNDGRTPGSRNLCNIPCSALLS
SDILTASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL

Human Lysozyme

Sequence: 23%   Cleavage: 11%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGLYNTRATNYNAGDRSTDYGI
FQINSRYWCNDGKTPGAVNACHLSCSAL
LQDNIADAVACAKRVVRDPLQGIRAWVAW
RNRCQNRDVRQYVQGCGV

Bovine Serum Albumin

Sequence: 26%   Cleavage: 14%

MKWVTFISLLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNELTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMADCC
EKQEPERNECFLSHKDDSPDLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVELVTKL
VTDLTKVHKECCHGDLLECADDRADLAK
YICDNQDTISSKLKECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEECAKDDPHACYSTVFDKLK
HLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin

Sequence: 25%   Cleavage: 10%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVNEVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
EVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSISSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHKPKATKEQLKA
VMDDFAAFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

Figure S4.17.   The fragmentation maps of the digested 5CP with C18 RP

enrichment at pH 2 using 2DMS with resolutions at 128 k * 8 k (left) and 256 k * 4 k (right).   The peptides assigned by 2DMS are highlighted in yellow.   The average sequence and cleavage coverage per protein in the 128 k * 8 k spectrum are 28% and 13% respectively, while the coverages in the 256 k * 4 k spectrum are 37% and 19% correspondingly.   The data show that in the low resolution 2DMS, the resolution increment in the fragment *m/z* dimension is more important than the resolution increment in the precursor *m/z* dimension.

Fragmentation map of digested 5CP (RP pH 2 enriched)

2DMS (1 M * 8 k)

b ion

y ion

Human Hemoglobin Chain A

Sequence: 40%   Cleavage: 26%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B

Sequence: 69%   Cleavage: 37%

VHLLTPEEKSAVTALLWGKVNLVDEVGGEALI
GRLLVVYPWTQRFFESFGDLLSTPDAVMG
NPKVKAHGKKVLLGAFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFTPPVQAAYQKVVAGVANAL
AHKYH

Chicken Lysozyme

Sequence: 60%   Cleavage: 36%

KVFGRCELAAAMKRHGLLDNYRGYSLLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILLQ
INSRWWCNDGRTPGSRNLCNIPCSALLS
SDILTASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL

Human Lysozyme

Sequence: 30%   Cleavage: 14%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGYNTRATNYNAGDRSTDYGI
FQINSRYWCNDGKTPGAVNACHLSCSAL
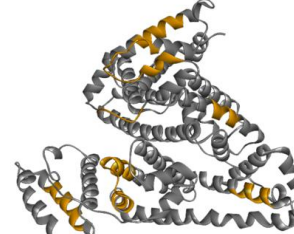LQDNIADAVACAKRVVRDPQGIRAWVAW
RNRCQNRDVRQYVQGCGV

Bovine Serum Albumin

Sequence: 27%   Cleavage: 15%

MKWVTFISLLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNELLTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMADCC
EKQEPERNECFLSHKDDSPDLLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVEVTKL
VTDLTKVHKECCHGDLLECADDRADLAK
YICDNQDTLSSKLKECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEECAKDDPHACYSTVFDKLK
HLLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
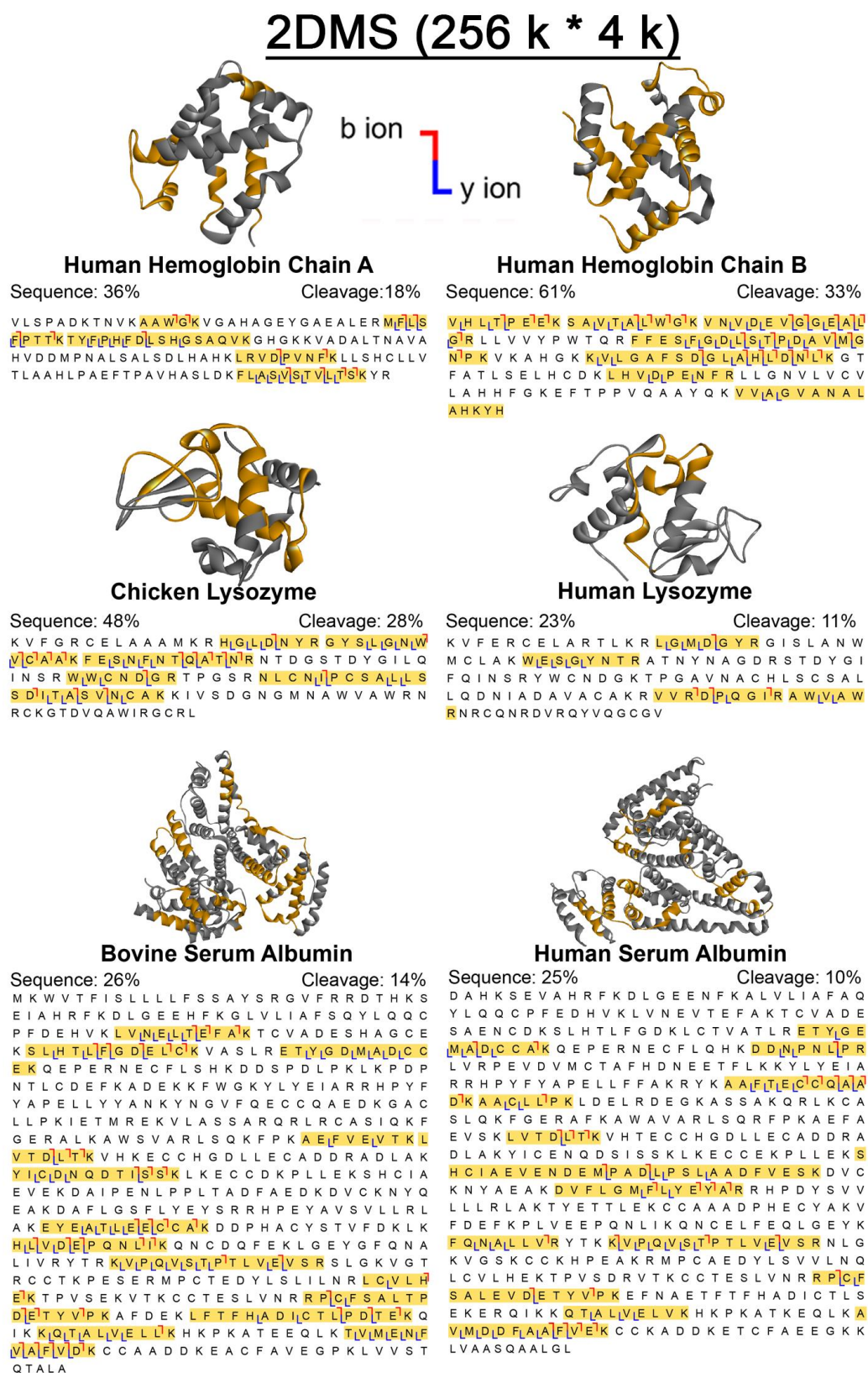IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin

Sequence: 31%   Cleavage: 17%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVNELVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
ELVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHKPKATKEQLKA
VLMDDLFAAFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

(Figure S4.18. Continue)

2DMS (2 M * 4 k)

Human Hemoglobin Chain A

Sequence: 37%   Cleavage: 21%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B

Sequence: 52%   Cleavage: 29%

VHLTPEEKSAVTALLWGKVNLVDEVGGEALL
GRLLVVYPWTQRFFESFGDLLSTPDAVMG
NPKVKAHGKKVLLGALFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFTPPVQAAYQKVVAGVANAL
AHKYH

Chicken Lysozyme

Sequence: 48%   Cleavage: 26%

KVFGRCELAAAMKRHGLDNYRGYSLLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILQ
INSRWWCNDGRTPGSRNLCNIPCSALLS
SDITASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL

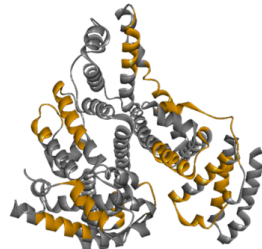**Human Lysozyme**

Sequence: 21%   Cleavage: 9%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGYNTRATNYNAGDRSTDYGI
FQINSRYWCNDGKTPGAVNACHLSCSAL
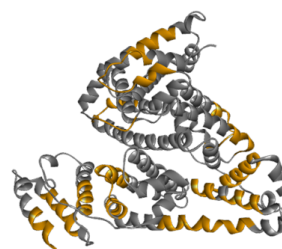LQDNIADAVACAKRVVRDPPQGIRAWVAW
RNRCQNRDVRQYVQGCGV

Bovine Serum Albumin

Sequence: 18%   Cleavage: 10%

MKWVTFISLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNLELLTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMADCC
EKQEPERNECFLSHKDDSPDLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVEVTKL
VTDLTKVHKECCHGDLLECADDRADLAK
YICDNQDTISSKLKECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEELCAKDDPHACYSTVFDKLK
HLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin

Sequence: 15%   Cleavage: 8%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVNELVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
ELVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSISSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHKPKATKEQLKA
VMDDFAAFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

Figure S4.18. The fragmentation maps of the digested 5CP with C18 RP enrichment at pH 2 using 2DMS with resolutions at 1 M * 8 k (left) and 2 M * 4 k

(right).   The peptides assigned by 2DMS are highlighted in yellow.   The average sequence and cleavage coverage per protein in the 1 M * 8 k spectrum are 43% and 24%, respectively, while the coverages in the 2 M * 4 k spectrum are 32% and 17% correspondingly.   The data show that in the high resolution 2DMS, the resolution increment in the precursor *m/z* dimension is more critical than the resolution increment in the fragment *m/z* dimension.
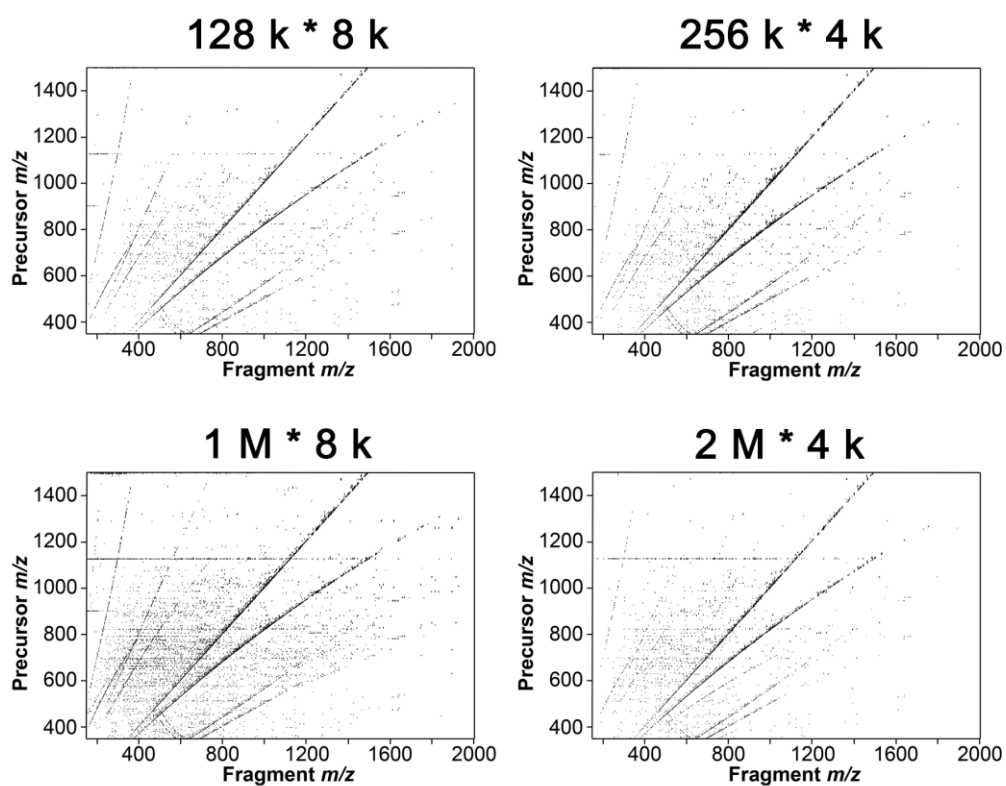
Figure S4.19. 2DMS spectra of the digested 5CP with HILIC enrichment conducting in various resolutions of fragment and precursor *m/z* dimensions. The 2DMS contour plots are displayed with the same intensity.

Fragmentation map of digested 5CP (HILIC enriched)

2DMS (128 k * 8 k)

b ion

y ion

Human Hemoglobin Chain A

Sequence: 26%   Cleavage: 14%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S
F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A
H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V
T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B

Sequence: 37%   Cleavage: 20%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L
G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G
N P K V K A H G K K V L G A F S D G L A H L D N L K G T
F A T L S E L H C D K L H V D P E N F R L L G N V L V C V
L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L
A H K Y H

Chicken Lysozyme

Sequence: 24%   Cleavage: 7%

K V F G R C E L A A A M K R H G L D N Y R G Y S L G N W
V C A A F E S N F N T Q A T N R N T D G S T D Y G I L Q
I N S R W W C N D G R T P G S R N L C N I P C S A L L S
S D I T A S V N C A K K I V S D G N G M N A W V A W R N
R C K G T D V Q A W I R G C R L
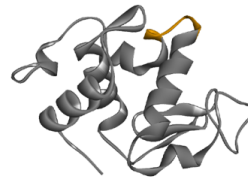
Human Lysozyme

Sequence: 5%   Cleavage: 3%

K V F E R C E L A R T L K R L G M D G Y R G I S L A N W
M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I
F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L
L Q D N I A D A V A C A K R V V R D P Q G I R A W V A W
R N R C Q N R D V R Q Y V Q G C G V

Bovine Serum Albumin

Sequence: 24%   Cleavage: 11%

M K W V T F I S L L L L L F S S A Y S R G V F R R D T H K S
E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C
P F D E H V K L V N E L T E F A K T C V A D E S H A G C E
K S L H T L F G D E L C K V A S L R E T Y G D M A D C C
E K Q E P E R N E C F L S H K D D S P D L P K L K P D P
N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F
Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C
L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F
G E R A L K A W S V A R L S Q K F P K A E F V E V T K L
V T D L T K V H K E C C H G D L L E C A D D R A D L A K
Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A
E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q
E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L
A K E Y E A T L E E C C A K D D P H A C Y S T V F D K L K
H L V D E P Q N L K Q N C D Q F E K L G E Y G F Q N A
L I V R Y T R K V P Q V S T P T L V E V S R S L G K V G T
R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H
E K T P V S E K V T K C C T E S L V N R R P C F S A L T P
D E T Y V P K A F D E K L F T F H A D I C T L P D T E K Q
I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F
V A F V D K C C A A D D K E A C F A V E G P K L V V S T
Q T A L A

Human Serum Albumin

Sequence: 14%   Cleavage: 8%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q
Y L Q Q C P F E D H V K L V N E V T E F A K T C V A D E
S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E
M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R
L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A
R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A
D K A A C L L P K L D E L R D E G K A S S A K Q R L K C A
S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A
E V S K L V T D L T K V H T E C C H G D L L E C A D D R A
D L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S
H C I A E V E N D E M P A D L P S L A A D F V E S K D V C
K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V
L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V
F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K
F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G
K V G S K C C K H P E A K R M P C A E D Y L S V V L N Q
L C V L H E K T P V S D R V T K C C T E S L V N R R P C F
S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S
E K E R Q I K K Q T A L V E L V K H H K P K A T K E Q L K A
V M D D F A A F V E K C C K A D D K E T C F A E E G K K
L V A A S Q A A L G L

(Figure S4.20. Continue)

2DMS (256 k * 4 k)

b ion
y ion

Human Hemoglobin Chain A
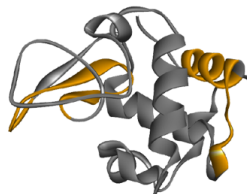
Sequence: 26%   Cleavage: 19%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B

Sequence: 53%   Cleavage: 26%

VHLTPEEKSAVTALWGKVNLVDEVGGEAL
GRLLVVYPWTQRFFESFGDLSTPDAVMG
NPKVKAHGKKVLGAFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFTPPVQAAYQKVVAGVANAL
AHKYH

Chicken Lysozyme

Sequence: 15%   Cleavage: 10%

KVFGRCELAAAMKRHGLDNYRGYSLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILQ
INSRWWCNDGRTPGSRNLCNIPCSALLS
SDITASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL

Human Lysozyme

Sequence: 18%   Cleavage: 9%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGYNTRATNYNAGDRSTDYGI
FQINSRYWCNDGKTPGAVNACHLSCSAL
LQDNIADAVACAKRVVRDPQGIRAWVAW
RNRCQNRDVRQYVQGCGV

Bovine Serum Albumin

Sequence: 15%   Cleavage: 8%

MKWVTFISLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNELTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMADCC
EKQEPERNECFLSHKDDSPDLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVEVTKL
VTDLTKVHKECCHGDLLECADDRADLAK
YICDNQDTISSKLECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEECCAKDDPHACYSTVFDKLK
HLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin

Sequence: 16%   Cleavage: 10%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVNEVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
EVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSISSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHKPKATKEQLKA
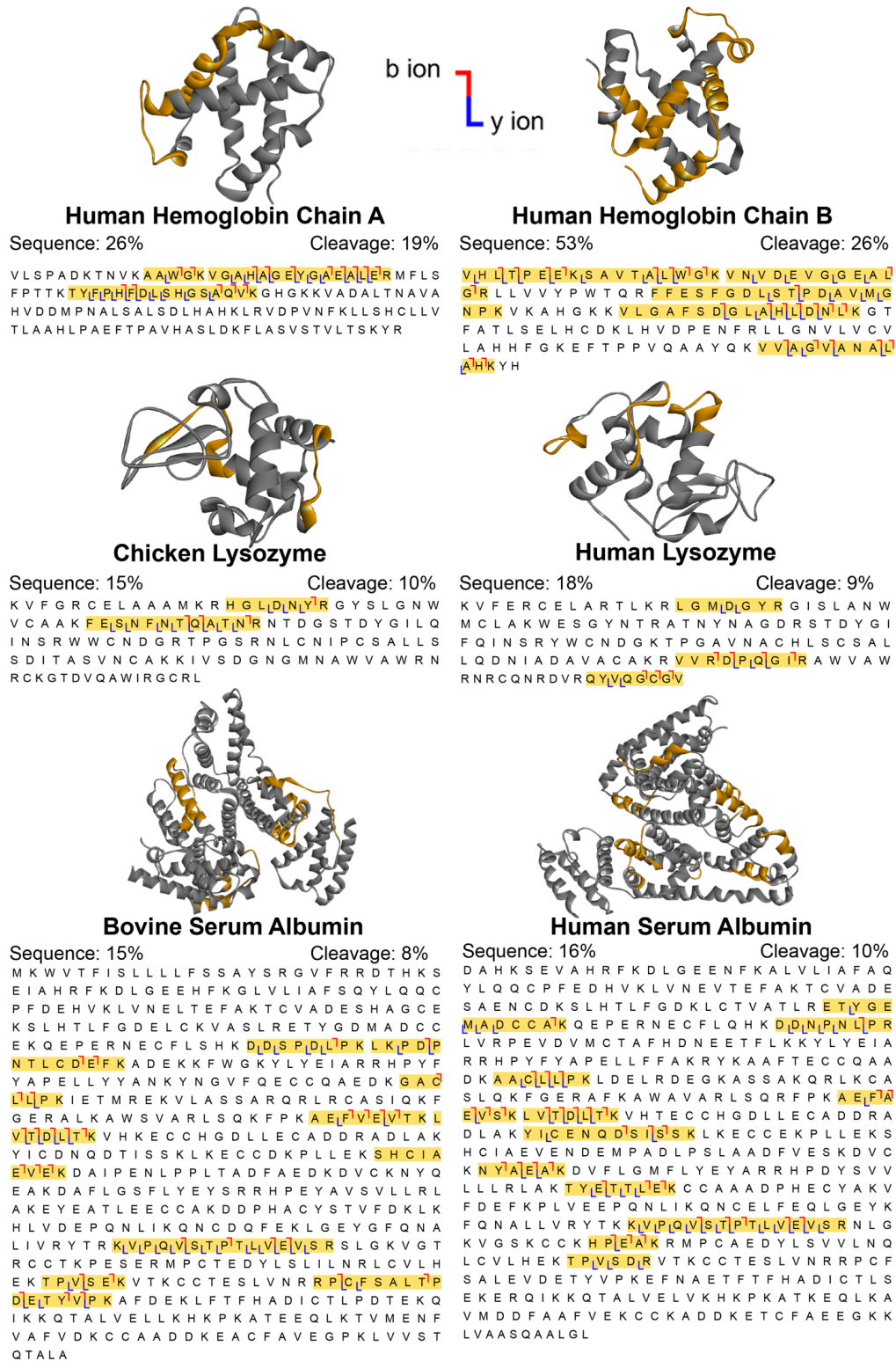VMDDFAAFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

Figure S4.20. The fragmentation maps of the digested 5CP with HILIC

enrichment using 2DMS with resolutions at 128 k * 8 k (left) and 256 k * 4 k (right).   The peptides assigned by 2DMS are highlighted in yellow.   The average sequence and cleavage coverage per protein in the 128 k * 8 k spectrum are 22% and 11% respectively, while the coverages in the 256 k * 4 k spectrum are 24% and 14% correspondingly.   The data show that in the low resolution 2DMS, the resolution increment in the fragment m/z dimension is more important than the resolution increment in the precursor m/z dimension, similar to as shown above (see Supplementary Information, Figure S4.17)

Fragmentation map of digested 5CP (HILIC enriched)

2DMS (1 M * 8 k)
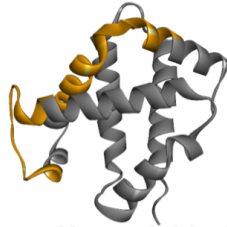
b ion
y ion

Human Hemoglobin Chain A
Sequence: 37%   Cleavage: 26%

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLS
FPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVA
HVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
TLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

Human Hemoglobin Chain B
Sequence: 78%   Cleavage: 46%

VHLTPEEKSAVTALLWGKVNVDEVGGEAL
GRLLVVYPWTQRFFESFGDLSTPDAVMG
NPKVKAHGKKVLGAFSDGLAHLDNLKGT
FATLSELHCDKLHVDPENFRLLGNVLVCV
LAHHFGKEFLTPPVLQAAYLQKVVAGVANALL
AHKYH

Chicken Lysozyme
Sequence: 52%   Cleavage: 32%

KVFGRCELAAAMKRHGLLDNLYRGYSLGNW
VCAAKFESNFNTQATNRNTDGSTDYGILLQ
LNSRWWCNDGRTPGSRNLCNIPCSALLS
SDLITASVNCAKKIVSDGNGMNAWVAWRN
RCKGTDVQAWIRGCRL

Human Lysozyme
Sequence: 24%   Cleavage: 14%

KVFERCELARTLKRLGMDGYRGISLANW
MCLAKWESGYNTRATNYNAGDRSTDYGI
FQINSRYWCNDGKTPGAVNACHLSCSAL
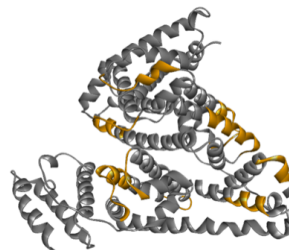LQDNIADAVACAKRVVRDPQGIRAWVAW
RNRCQNRDVRQYVQGCGV
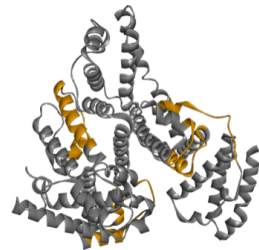
Bovine Serum Albumin
Sequence: 29%   Cleavage: 16%

MKWVTFISLLLLLFSSAYSRGVFRRDTHKS
EIAHRFKDLGEEHFKGLVLIAFSQYLQQC
PFDEHVKLVNELTEFAKTCVADESHAGCE
KSLHTLFGDELCKVASLRETYGDMADCC
EKQEPERNECFLSHKDDSPDLPKLKPDP
NTLCDEFKADEKKFWGKYLYEIARRHPYF
YAPELLYYANKYNGVFQECCQAEDKGAC
LLPKIETMREKVLASSARQRLRCASIQKF
GERALKAWSVARLSQKFPKAEFVEVTKL
VTDLLTKVHKECCHGDLLECADDRADLAK
YICDNQDTISSKLKECCDKPLLEKSHCIA
EVEKDAIPENLPPLTADFAEDKDVCKNYQ
EAKDAFLGSFLYEYSRRHPEYAVSVLLRL
AKEYEATLEECCAKDDPHACYSTVFDKLK
HLVDEPQNLIKQNCDQFEKLGEYGFQNA
LIVRYTRKVPQVSTPTLVEVSRSLGKVGT
RCCTKPESERMPCTEDYLSLILNRLCVLH
EKTPVSEKVTKCCTESLVNRRPCFSALTP
DETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENF
VAFVDKCCAADDKEACFAVEGPKLVVST
QTALA

Human Serum Albumin
Sequence: 27%   Cleavage: 16%

DAHKSEVAHRFKDLGEENFKALVLIAFAQ
YLQQCPFEDHVKLVNEVTEFAKTCVADE
SAENCDKSLHTLFGDKLCTVATLRETYGE
MADCCAKQEPERNECFLQHKDDNPNLPR
LVRPEVDVMCTAFHDNEETFLKKYLYEIA
RRHPYFYAPELLFFAKRYKAAFTECCQAA
DKAACLLPKLDELRDEGKASSAKQRLKCA
SLQKFGERAFKAWAVARLSQRFPKAEFA
EVSKLVTDLTKVHTECCHGDLLECADDRA
DLAKYICENQDSISSKLKECCEKPLLEKS
HCIAEVENDEMPADLPSLAADFVESKDVC
KNYAEAKDVFLGMFLYEYARRHPDYSVV
LLLRLAKTYETTLEKCCAAADPHECYAKV
FDEFKPLVEEPQNLIKQNCELFEQLGEYK
FQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCCKHPEAKRMPCAEDYLSVVLNQ
LCVLHEKTPVSDRVTKCCTESLVNRRPCF
SALEVDETYVPKEFNAETFTFHADICTLS
EKERQIKKQTALVELVKHHKPKATKEQLKA
VMDDFAAFVEKCCKADDKETCFAEEGKK
LVAASQAALGL

(Figure S4.21. Continue)

2DMS (2 M * 4 k)

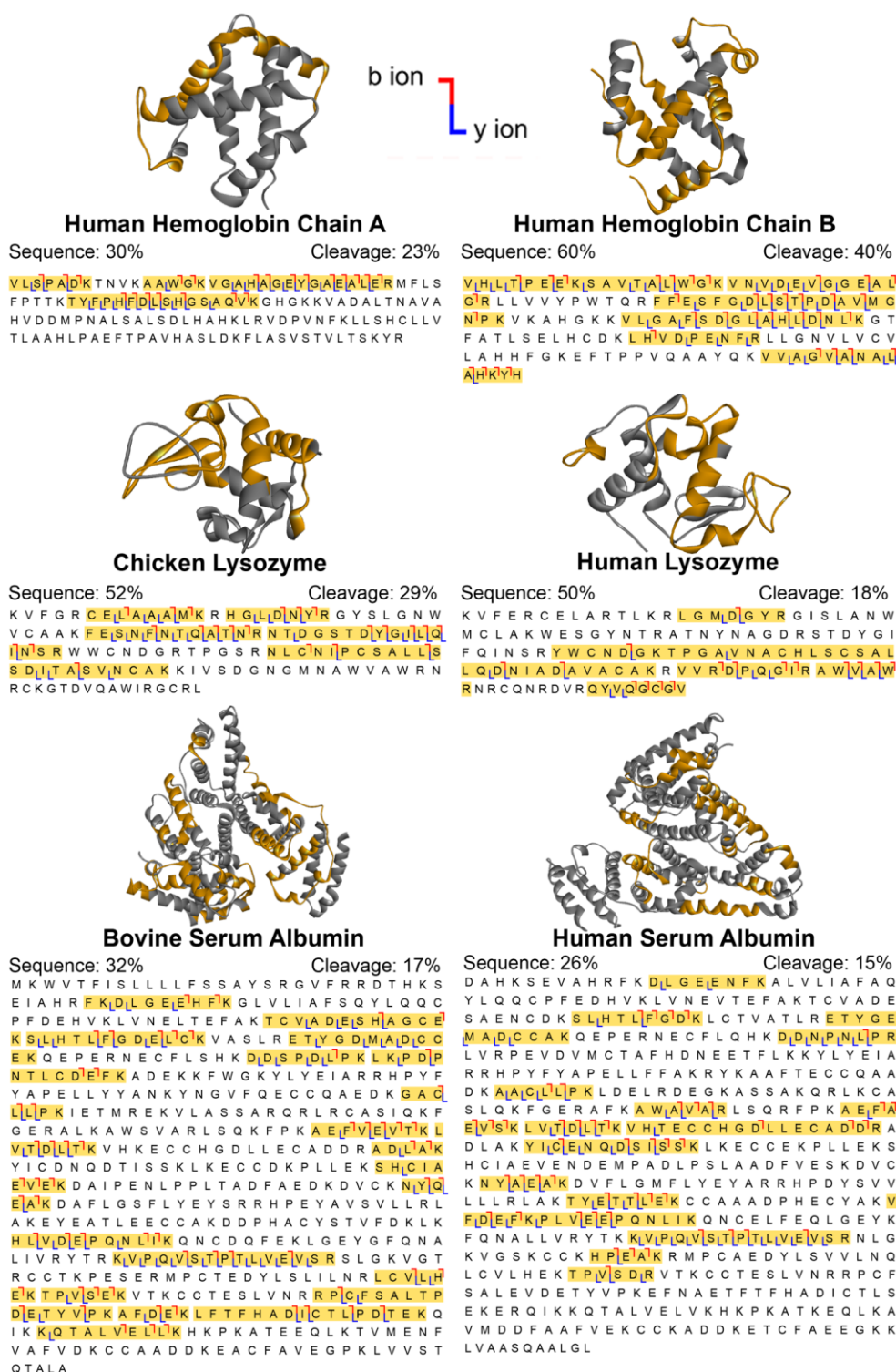**Human Hemoglobin Chain A**
Sequence: 30%   Cleavage: 23%

Human Hemoglobin Chain B
Sequence: 60%   Cleavage: 40%

Chicken Lysozyme
Sequence: 52%   Cleavage: 29%

Human Lysozyme
Sequence: 50%   Cleavage: 18%

Bovine Serum Albumin
Sequence: 32%   Cleavage: 17%

Human Serum Albumin
Sequence: 26%   Cleavage: 15%

Figure S4.21. The fragmentation maps of the digested 5CP with HILIC enrichment using 2DMS with resolutions at 1 M * 8 k (left) and 2 M * 4 k (right).

The peptides assigned by 2DMS are highlighted in yellow. The average sequence and cleavage coverage per protein in the 1 M * 8 k spectrum are 41% and 25% respectively, while the coverages in the 2 M * 4 k spectrum are 41% and 24% correspondingly. The data show in the high resolution 2DMS, the resolution increment in the precursor m/z dimension is more important than the fragment m/z dimension.

A) Fragmentation map of digested 5CP (RP and HILIC enrich) using IRMPD 2DMS (1 M * 8 k)

HILIC at pH 6.8 only | RP C18 at pH 2 only | HILIC at pH 6.8 and RP C18 at pH 2 | b ion | y ion
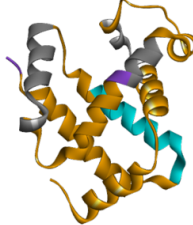
Human Hemoglobin Chain A
Sequence: 52%   Cleavage: 35%

Human Hemoglobin Chain B
Sequence: 80%   Cleavage: 51%

Chicken Lysozyme
Sequence: 66%   Cleavage: 46%

Human Lysozyme
Sequence: 36%   Cleavage: 19%

Bovine Serum Albumin
Sequence: 34%   Cleavage: 21%

Human Serum Albumin
Sequence: 44%   Cleavage: 26%

(Figure S4.22. Continue)

B) Venn Diagram of peptides identifed by 2DMS and nLC MS/MS

SPE	RP (pH 2)	HILIC (pH 6.8)
No. of Peptide	57	58
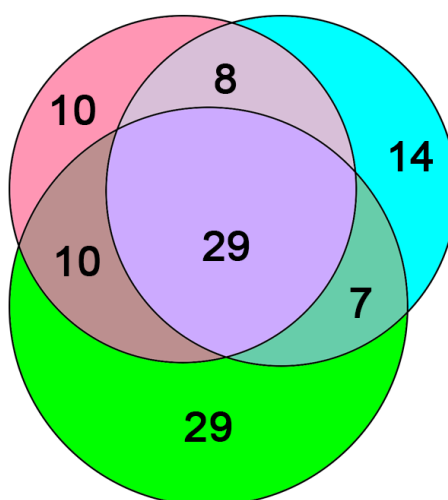
■ 2DMS (RP pH 2)
■ 2DMS (HILIC pH 6.8)
■ nLC MS/MS

C) Physical Properties of Identified peptides
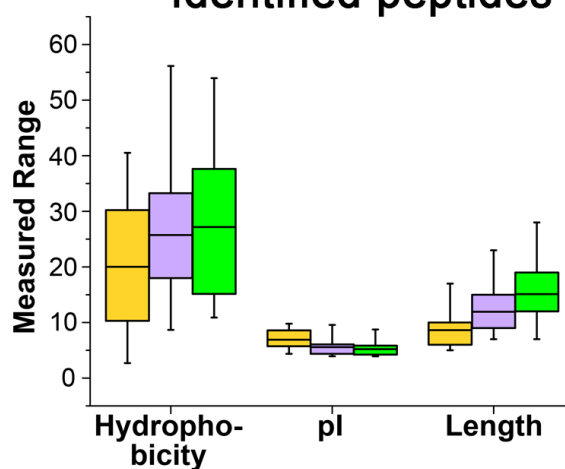
■ 2D only   ■ 2D and nLC
■ nLC only

Figure S4.22. Proteomic results of 5CP using high resolution 2DMS with RP and HILIC enrichment. (A) Fragmentation maps of the digested 5CP analysed using 2DMS with resolution at 1 M * 8 k (fragment *m/z* * precursor *m/z*) with RP and HILIC SPE cartridges enrichment. The average sequence and cleavage coverages per protein are 52% and 33% respectively, which are better than the results obtained from the 2DMS spectra of 256 k * 4 k (Figure 4E). (B) Venn diagram of peptides identified from 2DMS, including RP and HILIC SPE cartridges enrichment, and nLC MS/MS. (C) Box chart showing the physical properties of the identified peptides from 2DMS and nLC MS/MS. The ranges of physical properties being covered by 2DMS are similar to the observations in the results of 2DMS with resolution at 256 k * 4 k (Figure 4C), suggesting that increasing the resolving power of the fragment and precursor *m/z* dimension doesn't greatly enhance the coverage of different types of peptide but it can help to improve the assignment of peptides that originally contain a relative low S/N ratio.

Figure S4.23. 2DMS spectra (256 k * 4 k) of the digested 5CP samples with C18 RP enrichment at pH 2. The sample was then separated into 4 different *m/z* range using the quadrupole before the acquisition of 2DMS. The 2DMS contour plots are displayed with the same intensity.

**2DMS spectra (256 k * 4 k) of the digested 5CP (HILIC enriched)
with mass range isolation applied in the quadrupole**
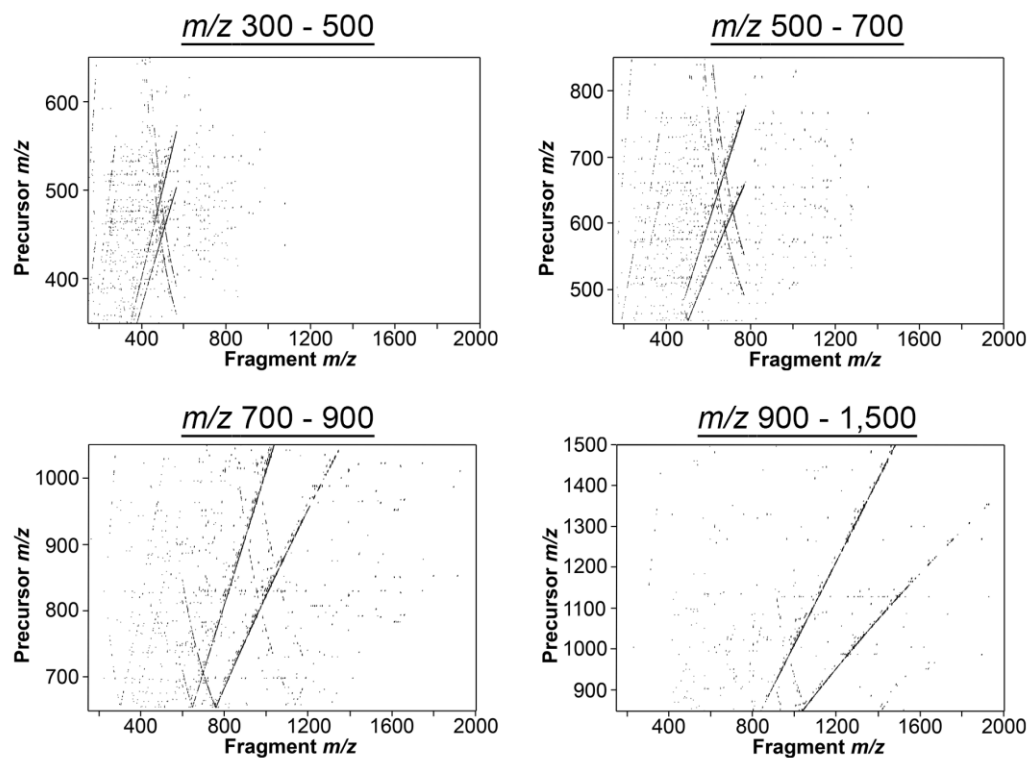
Figure S4.24. 2DMS spectra (256 k * 4 k) of the digested 5CP samples with HILIC enrichment at pH 6.8.   The sample was then separated into 4 different *m/z* range using the quadrupole before the acquisition of 2DMS.   The 2DMS contour plots are displayed with the same intensity.
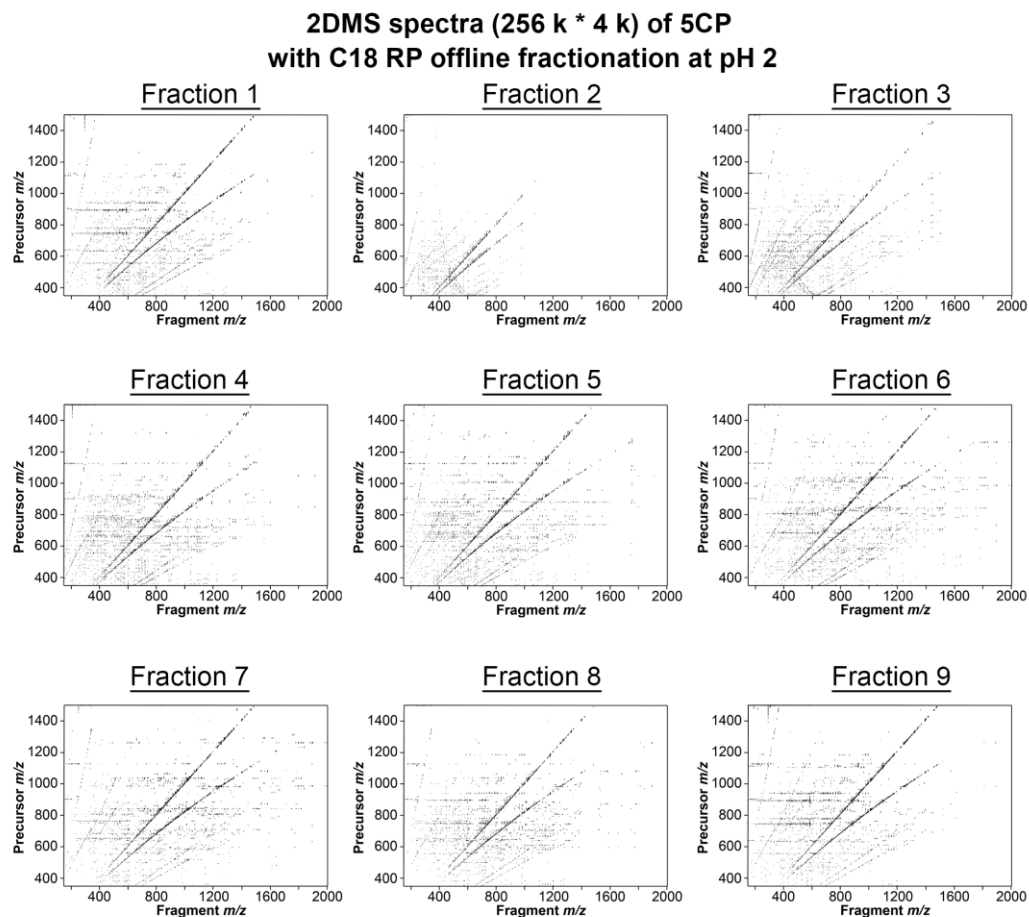
Figure S4.25. 2DMS spectra (256 k * 4 k) of the digested 5CP samples with offline fractionation using C18 RP SPE cartridge at pH 2. The 2DMS contour plots are displayed with the same intensity.

2DMS spectra (256 k * 4 k) of 5CP
with HILIC offline fractionation at pH 6.8
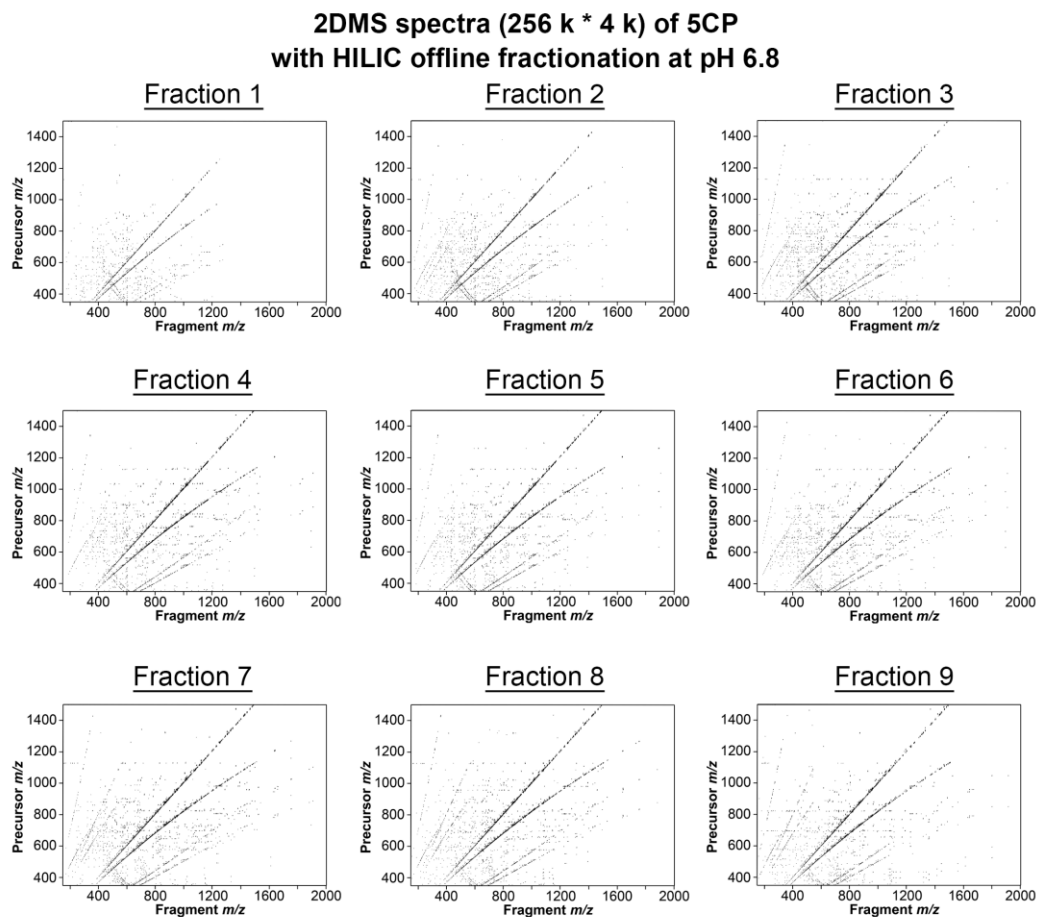
Figure S4.26. 2DMS spectra (256 k * 4 k) of the digested 5CP samples with offline fractionation using HILIC SPE cartridge at pH 6.8. The 2DMS contour plots are displayed with the same intensity.

A)

Fragmentation map of
digested 5CP (RP pH 2 enriched)
using IRMPD 2DMS (quadrupole isolated)
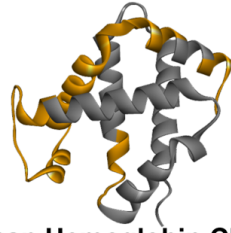
b ion
y ion

Human Hemoglobin Chain A
Sequence: 43% Cleavage: 29%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S
F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A
H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V
T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B
Sequence: 67% Cleavage: 40%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L
G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G
N P K V K A H G K K V L G A F S D G L A H L D N L K G T
F A T L S E L H C D K L H V D P E N F R L L G N V L V C V
L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L
A H K Y H

Chicken Lysozyme
Sequence: 67% Cleavage: 42%

K V F G R C E L A A A M K R H G L D N Y R G Y S L G N W
V C A A K F E S N F N T Q A T N R N T D G S T D Y G I L Q
I N S R W W C N D G R T P G S R N L C N I P C S A L L S
S D I T A S V N C A K K I V S D G N G M N A W V A W R N
R C K G T D V Q A W I R G C R L
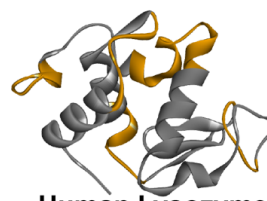
Human Lysozyme
Sequence: 35% Cleavage: 22%

K V F E R C E L A R T L K R L G M D G Y R G I S L A N W
M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I
F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L
L Q D N I A D A V A C A K R V V R D P Q G I R A W V A W
R N R C Q N R D V R Q Y V Q G C G V
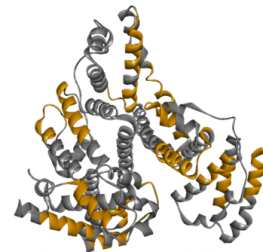
Bovine Serum Albumin
Sequence: 31% Cleavage: 18%

M K W V T F I S L L L L F S S A Y S R G V F R R D T H K S
E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C
P F D E H V K L V N E L T E F A K T C V A D E S H A G C E
K S L H T L F G D E L C K V A S L R E T Y G D M A D C C
E K Q E P E R N E C F L S H K D D S P D L P K L K P D P
N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F
Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C
L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F
G E R A L K A W S V A R L S Q K F P K A E F V E V T K L
V T D L T K V H K E C C H G D L L E C A D D R A D L A K
Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A
E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q
E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L
A K E Y E A T L E E C C A K D D P H A C Y S T V F D K L K
H L V D E P Q N L I K Q N C D Q F E K L G E Y G F Q N A
L I V R Y T R K V P Q V S T P T L V E V S R S L G K V G T
R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H
E K T P V S E K V T K C C T E S L V N R R P C F S A L T P
D E T Y V P K A F D E K L F T F H A D I C T L P D T E K Q
I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F
V A F V D K C C A A D D K E A C F A V E G P K L V V S T
Q T A L A

Human Serum Albumin
Sequence: 33% Cleavage: 19%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q
Y L Q Q C P F E D H V K L V N E V T E F A K T C V A D E
S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E
M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R
L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A
R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A
D K A A C L L P K L D E L R D E G K A S S A K Q R L K C A
S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A
E V S K L V T D L T K V H T E C C H G D L L E C A D D R A
D L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S
H C I A E V E N D E M P A D L P S L A A D F V E S K D V C
K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V
L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V
F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K
F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G
K V G S K C C K H P E A K R M P C A E D Y L S V V L N Q
L C V L H E K T P V S D R V T K C C T E S L V N R R P C F
S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S
E K E R Q I K K Q T A L V E L V K H K P K A T K E Q L K A
V M D D F A A F V E K C C K A D D K E T C F A E E G K K
L V A A S Q A A L G L

(Figure S4.27. Continue)

B) Peptides identified in each *m/z* range fraction
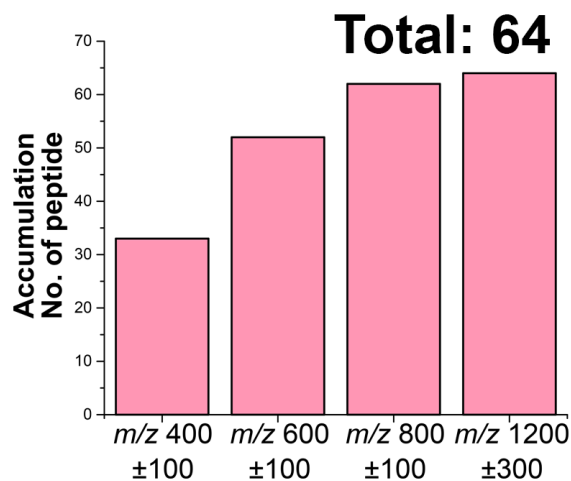
Total: 64

C) Venn Diagram of peptides identifed by 2DMS and nLC MS/MS

D) Physical Properties of Identified peptides

■ 2D only   ■ 2D and nLC   ■ nLC only
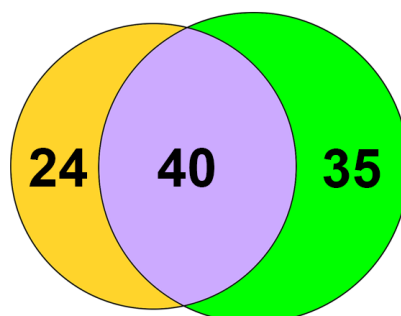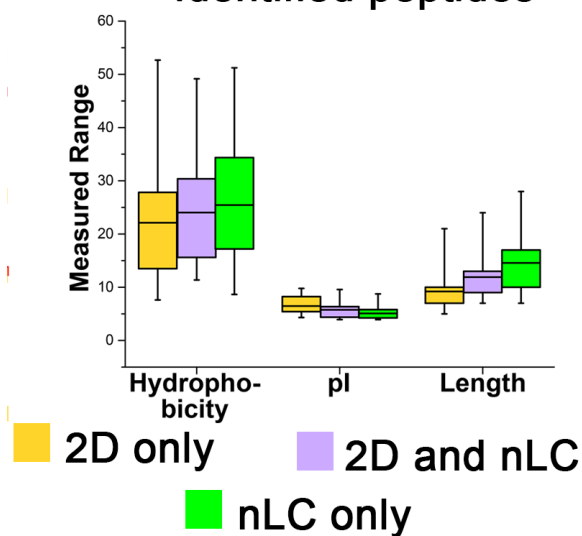
Figure S4.27. (A) Fragmentation maps of the digested 5CP with RP enrichment which was then isolated using the quadrupole and 4 *m/z* fractions were acquired using 2DMS. The peptides assigned by 2DMS are highlighted in yellow. The average sequence and cleavage coverage per protein are 46% and 28%, which are higher than the result obtained previously using the same sample but without *m/z* fractionation (see Supplementary Information, Figure S4.11). (B) The graph to show the accumulation number of peptides being identified in each *m/z* fraction. (C) Venn diagram to show the overlapping of peptides between 2DMS using *m/z* fractionation and nLC MS/MS. (D) The box chart diagram to show the physical properties of the identified in the 2DMS and nLC MS/MS. Compared to the previous results without *m/z* fractionation (see Supplementary Information, Figure S4.11C), the ranges of hydrophobicity and length that can be covered by 2DMS are slightly enhanced, suggesting reduced number of ions in the FTICR cell during 2DMS can improve the identification of the hydrophobic, long peptides.

A) Fragmentation map of digested 5CP (HILIC enriched) using IRMPD 2DMS (quadrupole isolated)

b ion
y ion

Human Hemoglobin Chain A

Sequence: 46%   Cleavage: 31%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B

Sequence: 68%   Cleavage: 45%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G N P K V K A H G K K V L G A F S D G L A H L D N L K G T F A T L S E L H C D K L H V D P E N F R L L G N V L V C V L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L A H K Y H

Chicken Lysozyme

Sequence: 66%   Cleavage: 42%

K V F G R C E L A A A M K R H G L D N Y R G Y S L G N W V C A A K F E S N F N L T Q A T N R N T D G S T D Y G I L Q I N S R W W C N D G R T P G S R N L C N I I P C S A L L S S D L T A S V N C A K K I V S D G N G M N A W V A W R N R C K G T D V Q A W I R G C R L

Human Lysozyme

Sequence: 29%   Cleavage: 19%

K V F E R C E L A R T L K R L G M D G Y R G I S L A N W M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L L Q D N I A D A V A C A K R V V R D P Q G I I R A W V A W R N R C Q N R D V R Q Y V Q G C G V

Bovine Serum Albumin

Sequence: 39%   Cleavage: 24%

M K W V T F I S L L L L L F S S A Y S R G V F R R D T H K S E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C P F D E H V K L V N E L T E F A K T C V A D E S H A G C E K S L H T L F G D E L C K V A S L R E T Y G D M A D C C E K Q E P E R N E C F L S H K D D S P D L P K L K P D P N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F G E R A L K A W S V A R L S Q K F P K A E F V E V T K L V T D L T K V H K E C C H G D L L E C A D D R A D L A K Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L A K E Y E A T L E E C C A K D D P H A C Y S T V F D K L K H L V D E P Q N L I K Q N C D Q F E K L G E Y G F Q N A L I V R Y T R K V P Q V S T P T L V E V S R S L G K V G T R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H E K T P V S E K V T K C C T E S L V N R R P C F S A L T P D E T Y V P K A F D E K L F T F H A D I C T L S D T E K Q I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F V A F V D K C C A A D D K E A C F A V E G P K L V V S T Q T A L A

Human Serum Albumin

Sequence: 40%   Cleavage: 22%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q Y L Q Q C P F E D H V K L V N E V T E F A K T C V A D E S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A D K A A C L L P K L D E L R D E G K A S S A K Q R L K C A S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A E V S K L V T D L T K V H T E C C H G D L L E C A D D R A D L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S H C I A E V E N D E M P A D L P S L A A D F V E S K D V C K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G K V G S K C C K H P E A K R M P C A E D Y L S V V L N Q L C V L H E K T P V S D R V T K C C T E S L V N R R P C F S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S E K E R Q I K K Q T A L V E L V K H K P K A T K E Q L K A V M D D F A A F V E K C C K A D D K E T C F A E E G K K L V A A S Q A A L G L

(Figure S4.28. Continue)

B) Peptides identified in each *m/z* range fraction

Total: 85

C) Venn Diagram of peptides identifed by 2DMS and nLC MS/MS
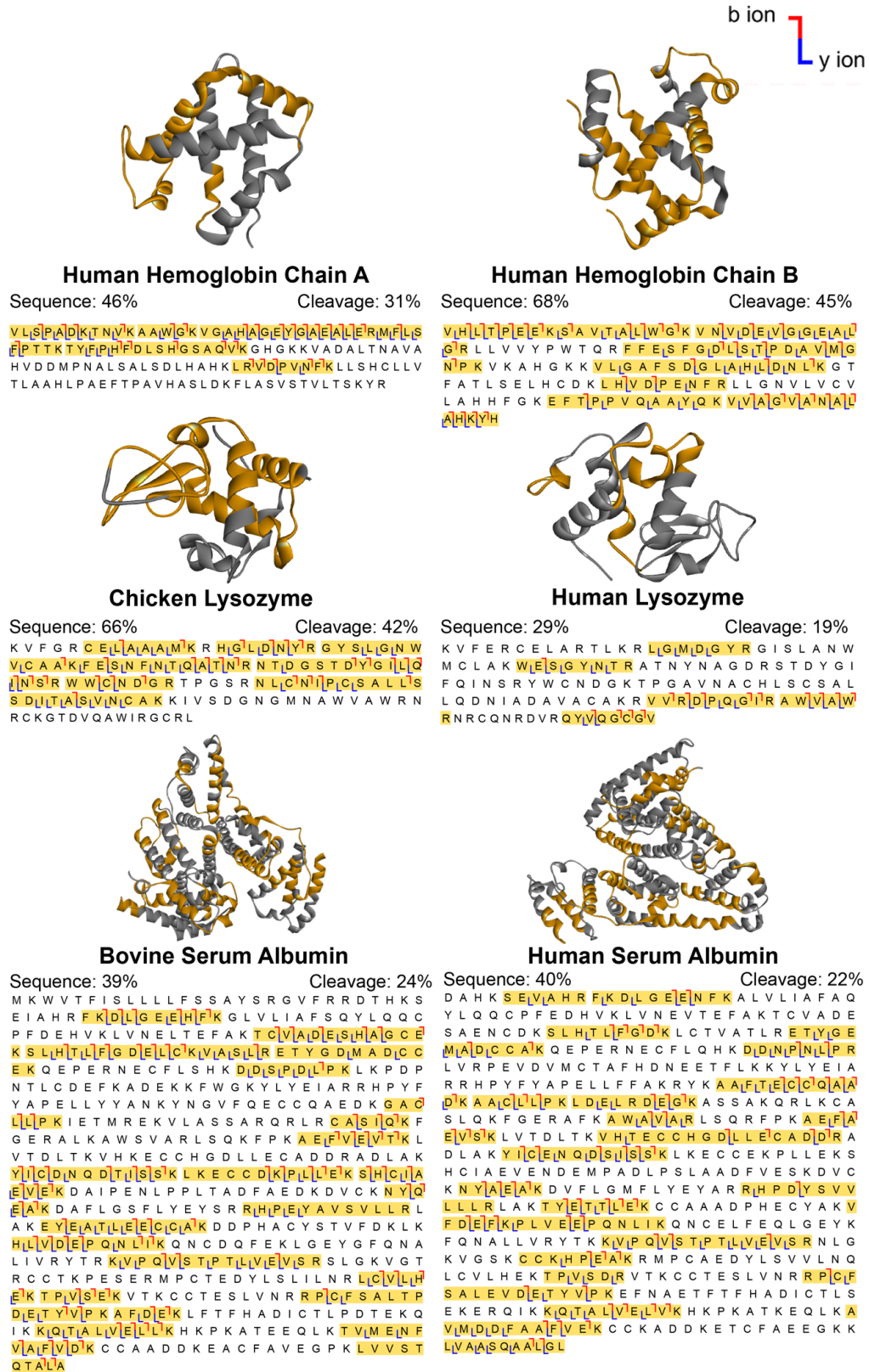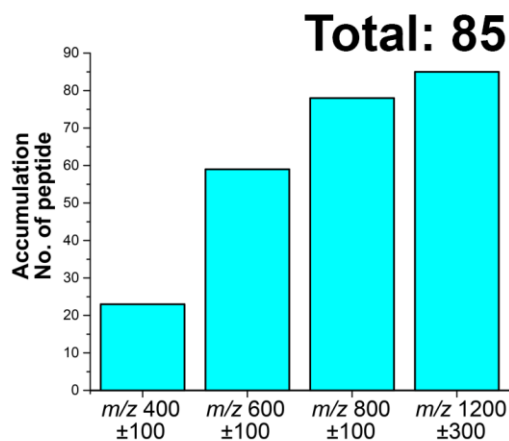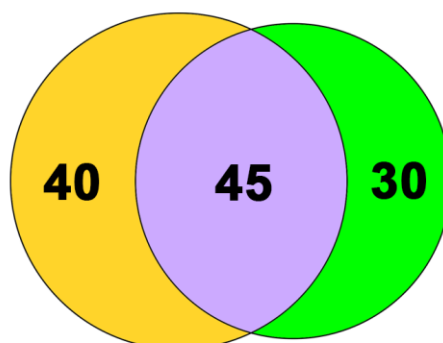
D) Physical Properties of Identified peptides

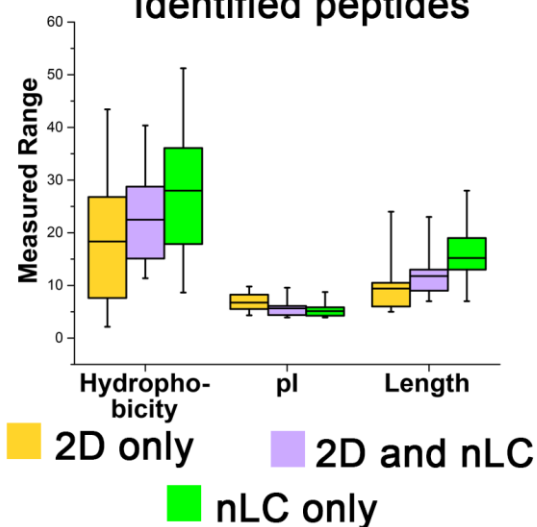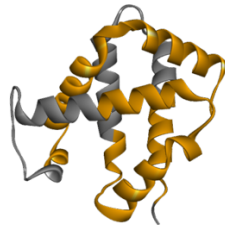2D only   2D and nLC   nLC only

389

Figure S4.28.   (A) Fragmentation maps of the tryptic digested 5CP with HILIC enrichment which was mass range isolated using quadrupole and each mass range was acquired with 2DMS.   The peptides assigned by 2DMS are highlighted in yellow.   The average sequence and cleavage coverage per protein are 48% and 31% respectively, which are higher than the previous result using the same sample but without any fractionation (see Supplementary Information, Figure S4.12).   (B) Cumulative number of identified peptides in each mass range fraction.   (C) Venn diagram showing the overlapping of peptides between 2DMS with mass range fractionation and nLC MS/MS.   (D)   Box chart showing the physical properties of the identified in the 2DMS and nLC MS/MS.   Compared to the previous results without mass range fractionation (see Supplementary Information, Figure S4.12C), the ranges of hydrophobicity and length that can be covered by 2DMS are enhanced, suggesting reduced the number of ions in the FTICR cell during 2DMS can improve the identification of the hydrophobic, long peptides.   The results of the digested 5CP enriched by RP and HILIC SPE cartridges which were then further isolated using quadrupole show that *m/z* range fractionation can effectively enhance the number of peptides being identified in 2DMS, the hydrophobicity range as well as the length of the peptides are also extended.

Fragmentation map of the digested 5CP (RP fractionation) using IRMPD 2DMS

Figure S4.29. Fragmentation maps of the digested 5CP with offline C18 RP SPE cartridge fractionation using IRMPD 2DMS. The peptides assigned by 2DMS are highlighted in yellow. The sample was fractionated into 9 fractions and 9 2DMS spectra were obtained (see Supplementary Information, Figure S4.25). The average sequence and cleavage coverages are 60% and 42% respectively, which provides the highest coverages among all the methods employed.

Fragmentation map of
the digested 5CP (HILIC fractionation)
using IRMPD 2DMS

b ion
y ion

Human Hemoglobin Chain A

Sequence: 45% Cleavage: 32%

V L S P A D K T N V K A A W G K V G A H A G E Y G A E A L E R M F L S
F P T T K T Y F P H F D L S H G S A Q V K G H G K K V A D A L T N A V A
H V D D M P N A L S A L S D L H A H K L R V D P V N F K L L S H C L L V
T L A A H L P A E F T P A V H A S L D K F L A S V S T V L T S K Y R

Human Hemoglobin Chain B

Sequence: 68% Cleavage: 50%

V H L T P E E K S A V T A L W G K V N V D E V G G E A L
G R L L V V Y P W T Q R F F E S F G D L S T P D A V M G
N P K V K A H G K K V L G A F S D G L A H L D N L K G T
F A T L S E L H C D K L H V D P E N F R L L G N V L V C V
L A H H F G K E F T P P V Q A A Y Q K V V A G V A N A L
A H K Y H

Chicken Lysozyme

Sequence: 66% Cleavage: 47%

K V F G R C E L A A A M K R H G L D N Y R G Y S L G N W
V C A A K F E S N F N T Q A T N R N T D G S T D Y G I L Q
I N S R W W C N D G R T P G S R N L C N I P C S A L L S
S D I T A S V N C A K K I V S D G N G M N A W V A W R N
R C K G T D V Q A W I R G C R L

Human Lysozyme

Sequence: 54% Cleavage: 33%

K V F E R C E L A R T L K R L G M D G Y R G I S L A N W
M C L A K W E S G Y N T R A T N Y N A G D R S T D Y G I
F Q I N S R Y W C N D G K T P G A V N A C H L S C S A L
L Q D N I A D A V A C A K R V V R D P Q G I R A W V A W
R N R C Q N R D V R Q Y V Q G C G V

Bovine Serum Albumin

Sequence: 51% Cleavage: 28%

M K W V T F I S L L L L F S S A Y S R G V F R R D T H K S
E I A H R F K D L G E E H F K G L V L I A F S Q Y L Q Q C
P F D E H V K L V N E L T E F A K T C V A D E S H A G C E
K S L H T L F G D E L C K V A S L R E T Y G D M A D C C
E K Q E P E R N E C F L S H K D D S P D L P K L K P D P
N T L C D E F K A D E K K F W G K Y L Y E I A R R H P Y F
Y A P E L L Y Y A N K Y N G V F Q E C C Q A E D K G A C
L L P K I E T M R E K V L A S S A R Q R L R C A S I Q K F
G E R A L K A W S V A R L S Q K F P K A E F V E V T K L
V T D L T K V H K E C C H G D L L E C A D D R A D L A K
Y I C D N Q D T I S S K L K E C C D K P L L E K S H C I A
E V E K D A I P E N L P P L T A D F A E D K D V C K N Y Q
E A K D A F L G S F L Y E Y S R R H P E Y A V S V L L R L
A K E Y E A T L E E C C A K D D P H A C Y S T V F D K L K
H L V D E P Q N L I K Q N C D Q F E K L G E Y G F Q N A
L I V R Y T R K V P Q V S T P T L V E V S R S L G K V G T
R C C T K P E S E R M P C T E D Y L S L I L N R L C V L H
E K T P V S E K V T K C C T E S L V N R R P C F S A L T P
D E T Y V P K A F D E K L F T F H A D I C T L P D T E K Q
I K K Q T A L V E L L K H K P K A T E E Q L K T V M E N F
V A F V D K C C A A D D K E A C F A V E G P K L V V S T
Q T A L A

Human Serum Albumin

Sequence: 44% Cleavage: 25%

D A H K S E V A H R F K D L G E E N F K A L V L I A F A Q
Y L Q Q C P F E D H V K L V N E V T E F A K T C V A D E
S A E N C D K S L H T L F G D K L C T V A T L R E T Y G E
M A D C C A K Q E P E R N E C F L Q H K D D N P N L P R
L V R P E V D V M C T A F H D N E E T F L K K Y L Y E I A
R R H P Y F Y A P E L L F F A K R Y K A A F T E C C Q A A
D K A A C L L P K L D E L R D E G K A S S A K Q R L K C A
S L Q K F G E R A F K A W A V A R L S Q R F P K A E F A
E V S K L V T D L T K V H T E C C H G D L L E C A D D R A
D L A K Y I C E N Q D S I S S K L K E C C E K P L L E K S
H C I A E V E N D E M P A D L P S L A A D F V E S K D V C
K N Y A E A K D V F L G M F L Y E Y A R R H P D Y S V V
L L L R L A K T Y E T T L E K C C A A A D P H E C Y A K V
F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E Y K
F Q N A L L V R Y T K K V P Q V S T P T L V E V S R N L G
K V G S K C C K H P E A K R M P C A E D Y L S V V L N Q
L C V L H E K T P V S D R V T K C C T E S L V N R R P C F
S A L E V D E T Y V P K E F N A E T F T F H A D I C T L S
E K E R Q I K K Q T A L V E L V K H K P K A T K E Q L K A
V M D D F A A F V E K C C K A D D K E T C F A E E G K K
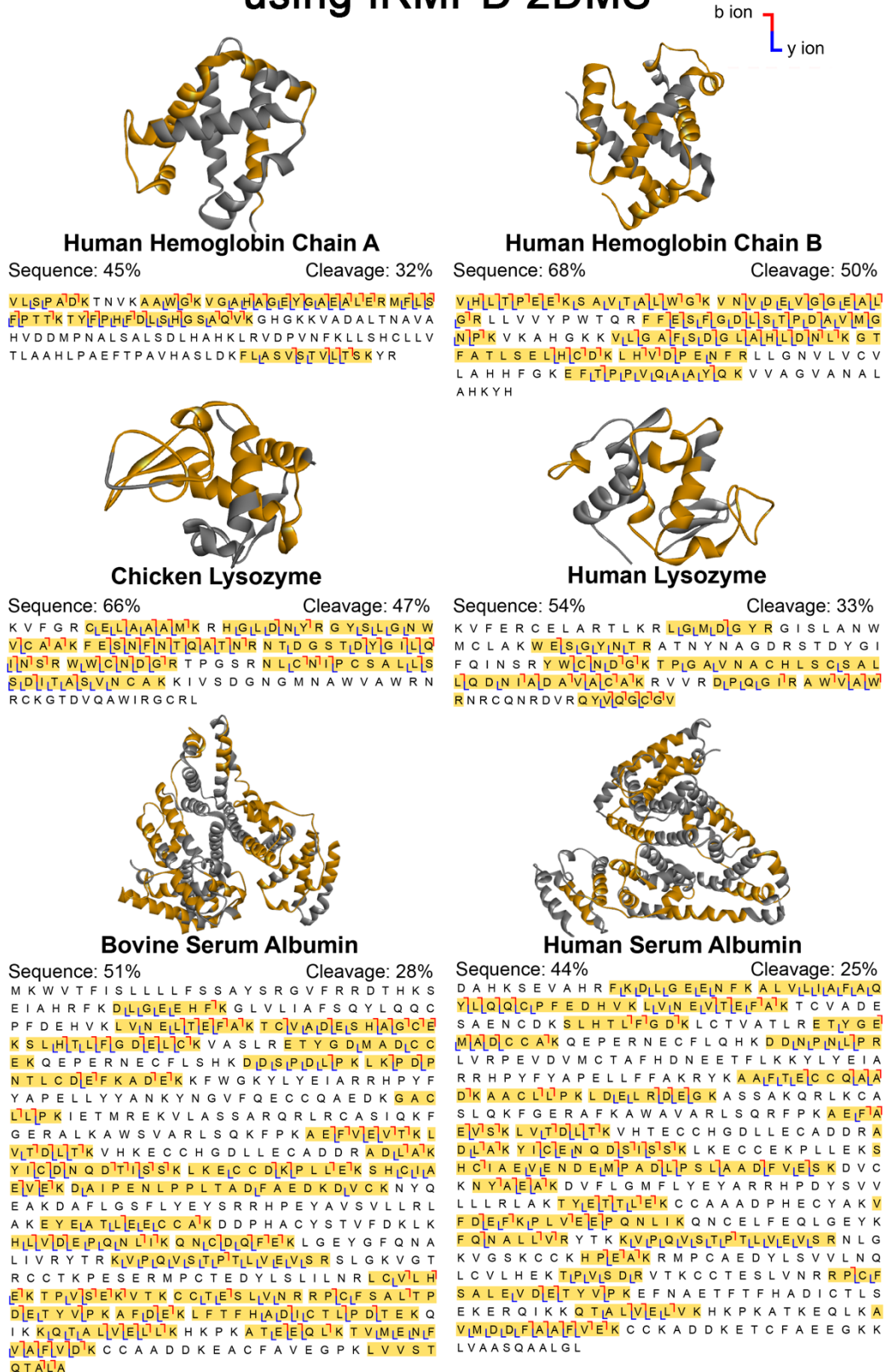L V A A S Q A A L G L

Figure S4.30. Fragmentation maps of the digested 5CP with offline HILIC SPE cartridge fractionation using IRMPD 2DMS. The peptides assigned by 2DMS are highlighted in yellow. The sample was fractionated into 9 fractions and 9 2DMS spectra were obtained (see Supplementary Information, Figure S4.26). The average sequence and cleavage coverages are 55% and 36% respectively.
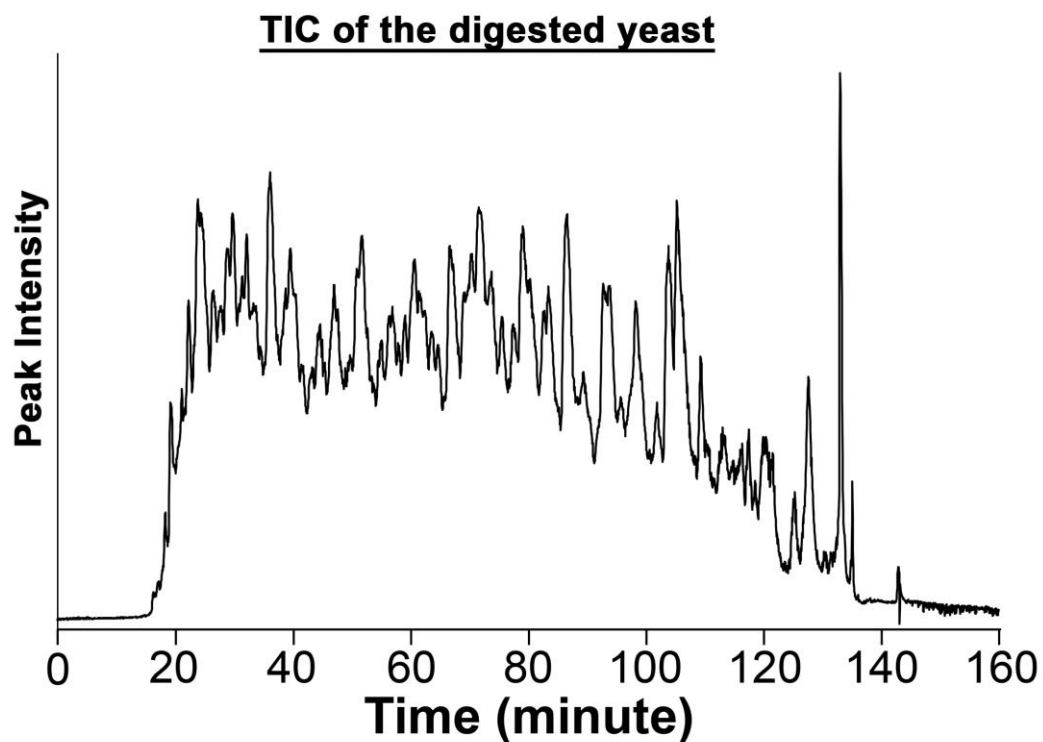
Figure S4.31. The total ion chromatogram (TIC) of the digested yeast analysed using a C18 RP nLC MS/MS with a 120-minute effective gradient, 160-minute total acquisition time.
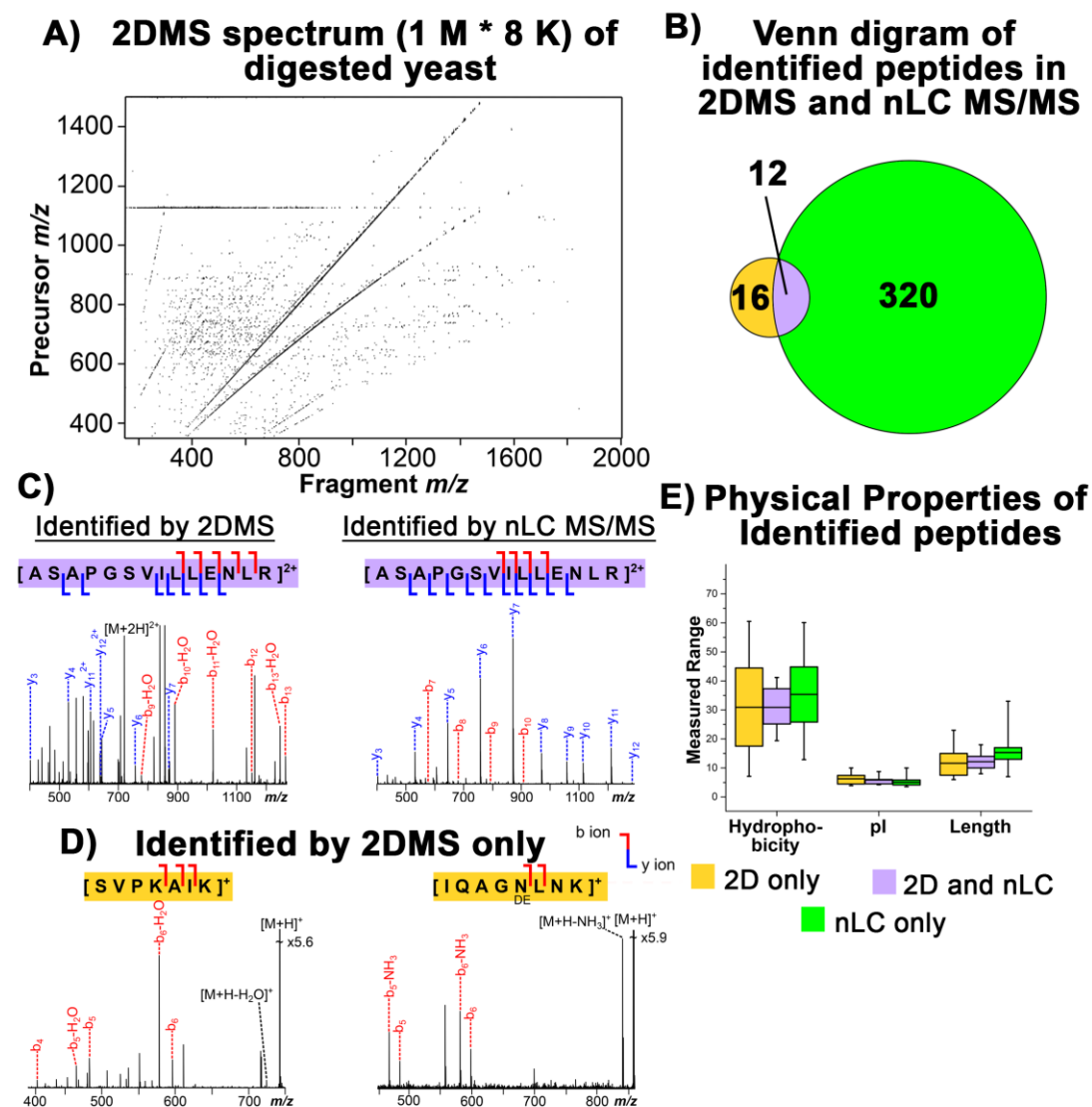
Figure S4.32. Proteomic results of yeast tryptic digest using 2DMS and nLC MS/MS. (A) 2DMS spectrum of yeast with resolution of 1 M * 8 k (fragment *m/z* * precursor *m/z*). (B) Venn diagram showing the overlapping of peptides between 2DMS and nLC MS/MS. 28 peptides were identified using 2DMS, and 16 out of these (57%) are not assigned using nLC MS/MS. (C) Spectra showing the peptide assigned using 2DMS (left) and nLC MS/MS (right). (D) Spectra showing the peptides exclusively assigned using 2DMS. (E) Box chart showing the physical properties of the identified peptides in the 2DMS and nLC MS/MS. The physical properties of the peptides being identified using 2DMS are similar to the observations in 2DMS of 5CP sample, in which more hydrophilic, basic, and

short peptides are assigned by 2DMS.
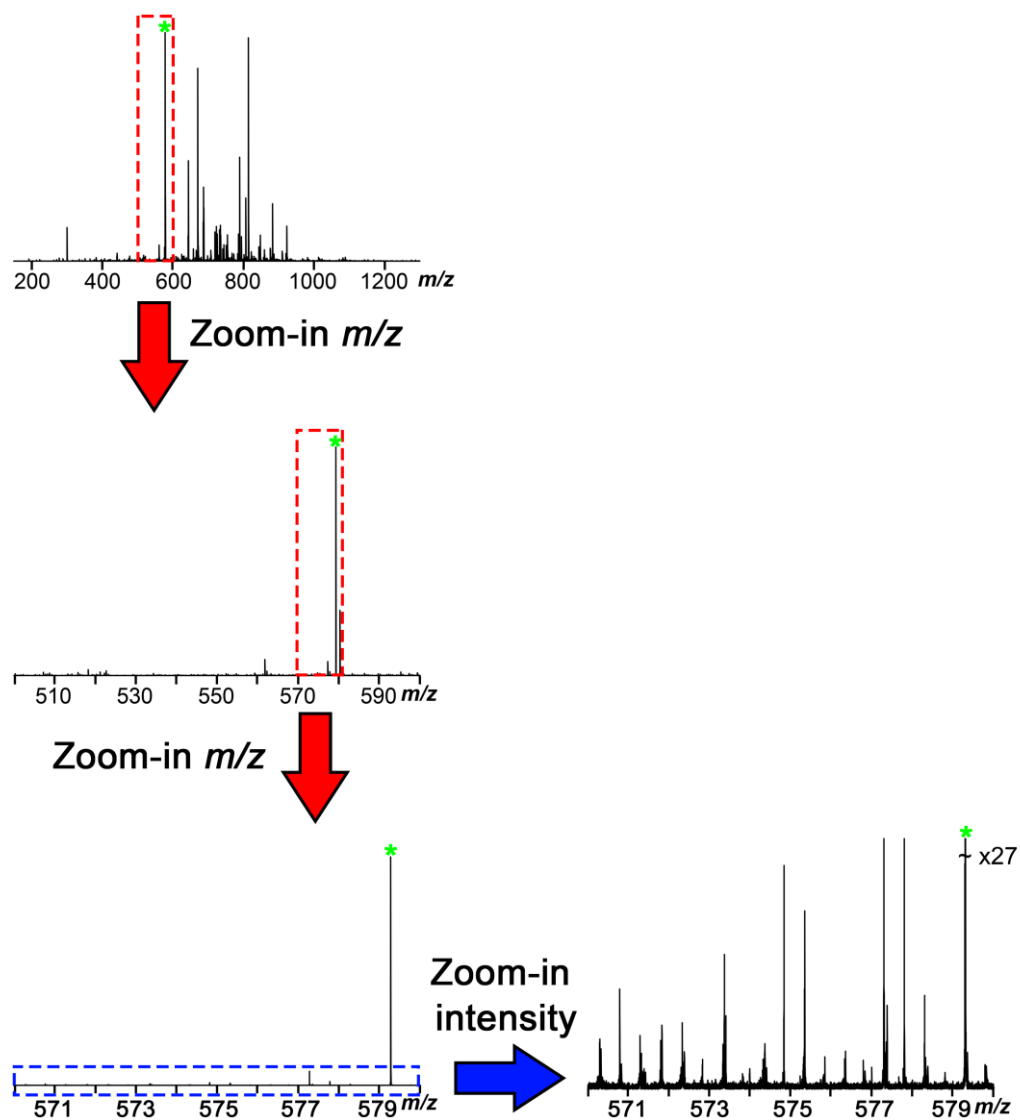
Figure S4.33. 1DMS spectrum of the tryptic digested yeast with 100 accumulation spectra.

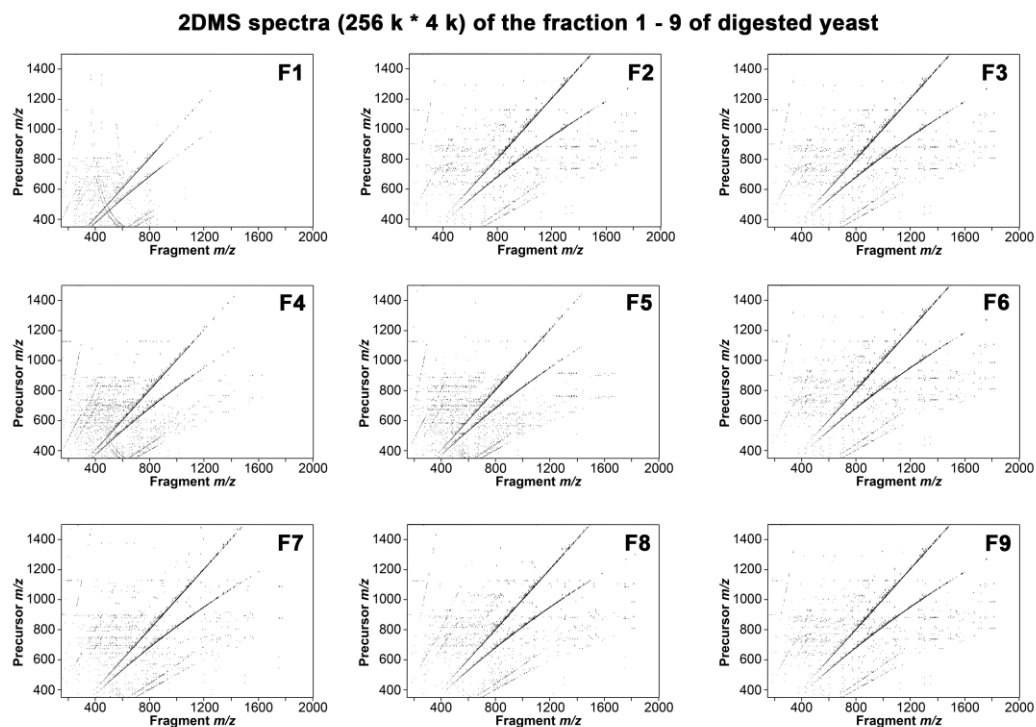**2DMS spectra (256 k * 4 k) of the fraction 1 - 9 of digested yeast**

Figure S4.34. 2DMS spectra of digested yeast with 256 k * 4 k resolution of the fraction 1-9 using C18 RP SPE cartridge at pH 2. The 2DMS contours plots are displayed with the same intensity.
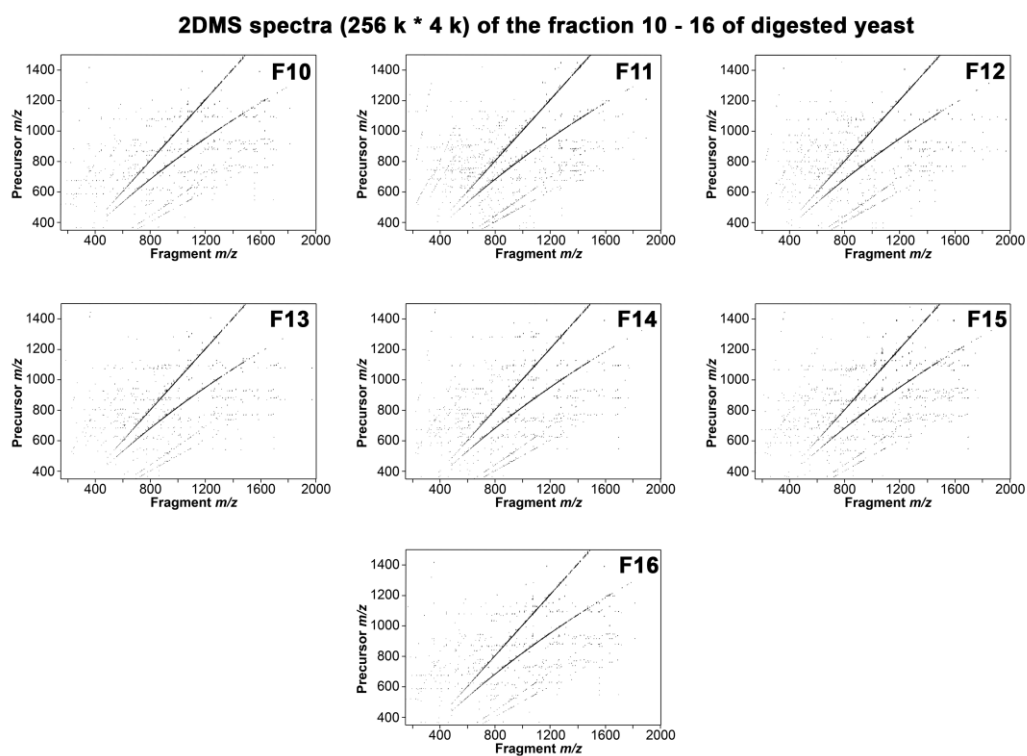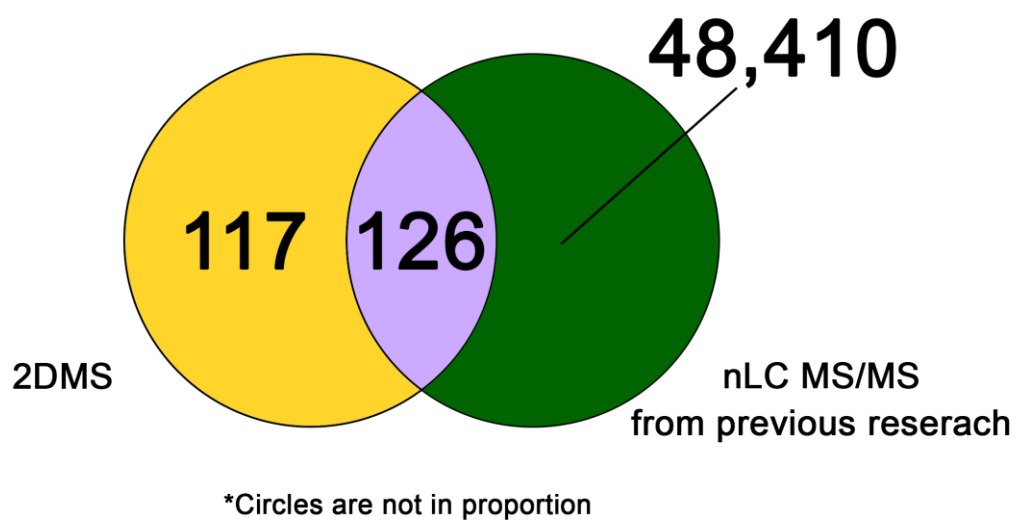
Figure S4.35. 2DMS spectra of digested yeast with 256 k * 4 k resolution of the fraction 10-16 using C18 RP SPE cartridge at pH 2. The 2DMS contours plots are displayed with the same intensity.

Figure S4.36.   Venn diagram showing the overlapping of the yeast peptides identified in the offline fractionated 2DMS experiment and the best set of yeast data obtained from the previous literature.[18]

Chapter 5 Conclusion and Future Work

The work presented within this thesis has demonstrated the application of mass spectrometry (MS) together with various advanced fragmentation techniques for the study of amyloid proteins as well as complex proteomic studies. This chapter aims at summarising and outlining the achievements for each of the previous chapters, future plans are also included which may help to improve the experiments or enhance the data quality in the near future.

Chapter 2: Exploring the Aggregation and Deamidation Mechanisms of Human Islet Amyloid Polypeptide Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry
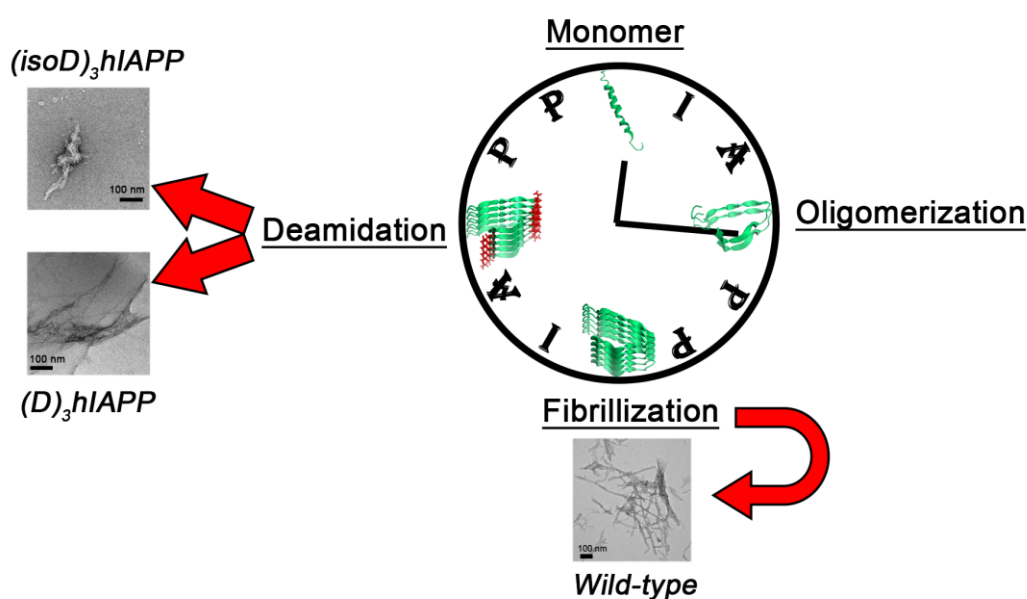
Chapter 2 focused on the study of aggregation mechanisms as well as the deamidation rate and sites of human islet amyloid polypeptide (hIAPP) using ultra-high resolution Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) together with fluorescence spectrometry and transmission electron microscopy (TEM).

From the MS results, hIAPP aggregates rapidly to form early oligomers,

ranging from dimer to pentamer, which were detected in the MS spectrum. The electron capture dissociation tandem MS (ECD MS/MS) results showed the region around Gly-33 and Ser-34 is the most critical region for the formation of dimerized hIAPPs as it is the minimal interaction region between two hIAPP units, this observation was further supported by the ECD MS/MS results from the trimer of hIAPP. Furthermore, the aggregation of individual hIAPP units into oligomers was shown to occur much faster than deamidation.

The collisionally activated dissociation (CAD) MS/MS results of deamidated hIAPP show that Asn-21, Asn-22, and Asn-35 residues are significantly deamidated within one month. The extent of deamidation was higher in the fibrillary pellets than in the solutions, which indicates that the deamidated hIAPP tends to form fibrils faster than wild-type hIAPP. This hypothesis is further supported by accelerated aggregation rate observed for ((D)$_3$hIAPP) and ((isoD)$_3$hIAPP) and a higher percentage of mutant hIAPPs identified in the fibrils of hIAPP solutions spiked with mutants ((D)$_3$hIAPP) or ((isoD)$_3$hIAPP) in wild-type hIAPP solutions.

The TEM images showed the difference in morphology of (isoD)$_3$hIAPP fibrils compared to wild-type hIAPP and ((D)$_3$hIAPP) fibrils, which helps explain the acceleration of aggregation caused by the deamidated hIAPP. Moreover, the amount of amyloid fibril formed is directly proportional to the spiking percentage of mutant hIAPPs; and the effect of seeding ((isoD)$_3$hIAPP) is twice that of seeding the same amount of ((D)$_3$hIAPP) into non-deamidated hIAPP solutions.

The results herein demonstrate FTICR MS with electron capture dissociation is a suitable analytical tool to understand the aggregation mechanism of amyloid proteins, indicating these techniques can be further applied to other amyloid proteins to investigate the aggregation mechanisms. The deamidation sites as

well as the effects of deamidation are shown using MS; however, the individual effect of each of the proposed deamidation sites is still unknown. Synthesising deamidated peptides with various combinations of deamidation site at Asn-21, Asn-22, and Asn-35 will help to narrow down the deamidation effect contributed by each site.

In future, research can focus on studying the effects of each individual deamidation site (Asn-21, Asn-22, and Asn-35) and combinations thereof towards the rate and morphology of hIAPP aggregation via different synthetic deamidated hIAPP configurations. This approach would allow the most critical deamidation site towards hIAPP aggregation to be determined. Results obtained could be particularly useful in further therapeutic development for identifying the most critical protein transformations.

Chapter 3: The Inhibition Pathways of Human Islet Amyloid Polypeptide

Chapter 5 – Conclusion and Future Work
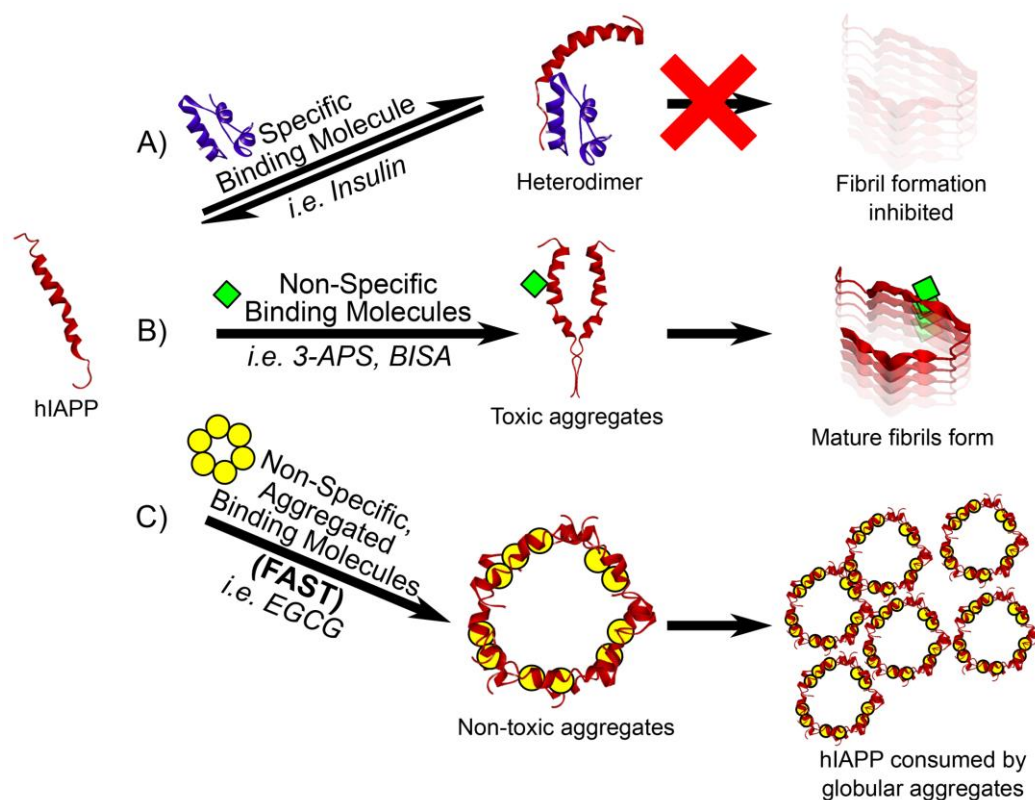
Chapter 3 focused on the study of inhibition pathways of hIAPP using FTICR MS together with fluorescence spectrometry and TEM, in which the potential pathways shown herein may help for accelerating therapeutic development in the coming future.

The experimental results here demonstrated two potential inhibition pathways are occurring to prevent the formation of toxic hIAPP aggregates. Inhibitors like insulin adopt a direct pathway through interaction with hIAPP between Ser-29 and Asn-35 residues which were proposed to be the most critical aggregation region for the formation of hIAPP early oligomers. These inhibitors are not only effective in preventing the formation of toxic hIAPP aggregates, but are also able to dissociate the hetero-complexes back into the monomer units of hIAPP and inhibitor, which is a possible key feature for potential drug development. These inhibitors, however, are only effective to limited amyloid protein sequences and structures, as a result, they are not viable to prevent the aggregation of mutant hIAPPs or other amyloid proteins.

On the other hand, other inhibitors, such as EGCG, can self-aggregate and trigger the formation of non-toxic, amorphous hIAPP aggregates faster than forming the toxic hIAPP aggregates. These inhibitors interact non-specifically to the amyloid proteins, thus they are effective to prevent the formation of toxic aggregates regardless of the sequences and structures of amyloid proteins. Despite the inhibition effect, the formation of amorphous aggregates seems to indicate the hIAPP units are continuously consumed once mixed with these inhibitors, in a long term, very low or even no free hIAPP unit is available for the normal physiological function inside the human body.

The results herein demonstrate MS together with fluorescence spectrometry and TEM help to determine the inhibition pathways of amyloid protein which can

be used to improve therapeutic development. In the future, the scale of this experiment should be expanded so that at least three different potential inhibitors are used in each measurement of a particular inhibition pathway to determine the variations of inhibition effects. Furthermore, the mixing ratio between hIAPP and potential inhibitors was set to 1:1 in all the experiments, however, this value has not been particularly optimised for a potential inhibitor. The lowest effective inhibition ratio should also be determined for each potential inhibitor as the concentration effect of the inhibitor may affect the performance in preventing amyloid fibril formation.

In future, more compounds could be explored for their inhibitory potential both to evaluate possible new inhibitors for potency, but also potentially discover even more new inhibition mechanisms of amyloid protein aggregation. Furthermore, targeted inhibitors with similar features to insulin (specific interaction with hIAPP) and EGCG (self-aggregation) could be synthesised and studied for their effectiveness on preventing the aggregation of hIAPP.

Chapter 4: The Next Dimension in Proteomics
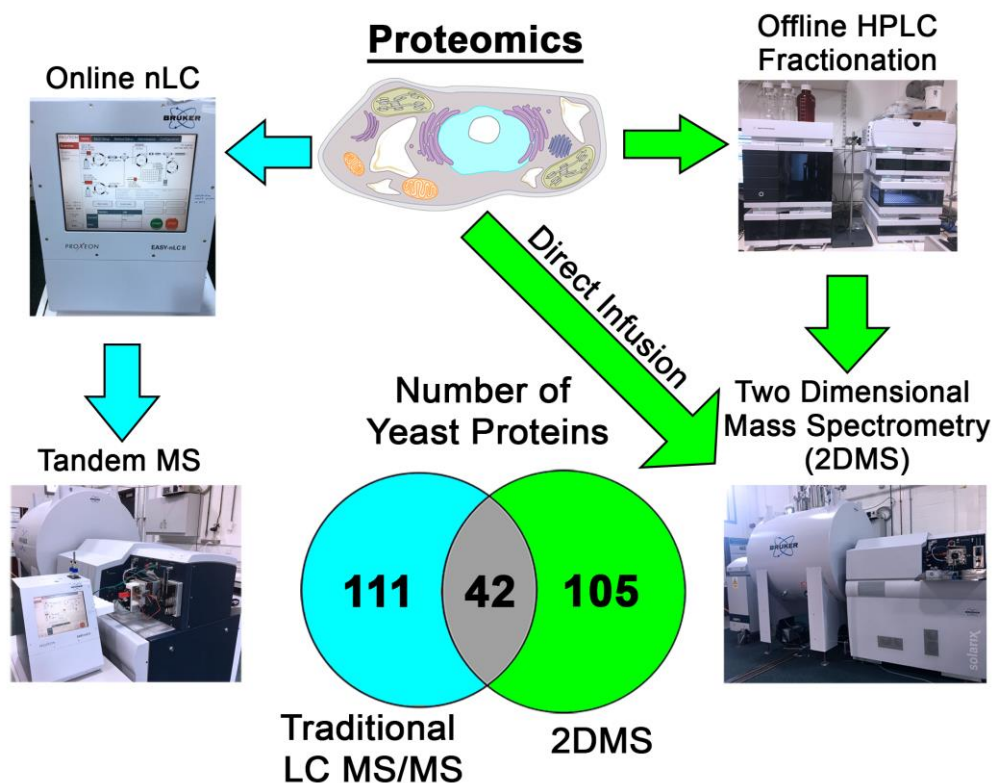
Chapter 4 focuses on studying proteomics using an alternative method – two dimensional mass spectrometry (2DMS) which allows the study of complex samples via direct infusion into the MS without prior online LC separation.

The results presented demonstrate that 2DMS is a viable alternative technique for proteomic studies, showing online LC MS/MS is no longer an essential tool in proteomics. Therefore, many proteomic analyses, such as biomolecular non-covalent interactions, which are usually very limited by available LCMS-compatible solvent systems and the available column chemistry of liquid chromatography, are able to be performed using 2DMS.

Furthermore, complementary protein and peptide assignments from the results of 2DMS and nLC MS/MS suggests a deeper proteome coverage can be achieved by combing the data obtained from 2DMS and nLC MS/MS, which also indicates the proteomes that are exclusively covered by 2DMS are currently under-

studied using purely standard nLC MS/MS techniques, thus 2DMS is not only an alternative but may soon be considered as a required complementary technique for deeper proteome analysis.

2DMS is also shown to have a higher efficiency in observing hydrophilic, basic, or short peptides, suggesting 2DMS is also suitable for application in PTM analysis as most of the important PTMs in disease-related issue, such as phosphopeptides and glycopeptides, are extremely hydrophilic which is challenging to analysis using standard nLC MS/MS technique.

2DMS is an advanced technique that has only recently been applied to proteomic studies, in which there are still lots of unexplored, potential areas for application of 2DMS. Native proteomics is a key priority experiment in the coming future since it is difficult to perform using traditional LC MS/MS which has been mentioned above. Analysis of native proteins is still challenging to perform using 2DMS as the current cell lysate and digestion methods are not well-optimised for native proteomics, thus increased optimisation for the workflow of native proteomics is required. Furthermore, top-down proteomics using 2DMS can also be performed. In top-down proteomics, proteins can be effectively separated by size exclusion chromatography, ion-exchange chromatography, and/or affinity chromatography, in which all the buffers used in these chromatography separations are incompatible with MS in an online LC mode. Thus, 2DMS is an obvious high-performance alternative method for analysing the fractioned proteins directly from these columns without applying the standard RP online chromatography columns.