IMPERIAL

DETERMINING THE GASTROINTESTINAL STABILITY OF PEPTIDE THERAPEUTICS VIA CHROMATOGRAPHIC METHODS

A Thesis Submitted to Imperial College London for the Degree of Doctor of Philosophy (Ph.D.)

By

Maxim Sebastian Francis Bird

Department of Chemical Engineering

Imperial College London

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence (CC BY NC-SA).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that; you credit the author, do not use it for commercial purposes and share any derivative works under the same licence.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

DECLARATION OF ORIGINALITY

I hereby state that this thesis was written in its entirety by the candidate and has not previously been submitted in part or via any alternative medium for the application of any of other degree in accordance with the University's Regulations for Research Degree Programmes. Except where specifically indicated by reference in the text, all experimental work in this thesis was planned, designed and completed by the candidate. Any views expressed throughout this thesis are those of the candidate.

Maxim S. F. Bird

September 2025

ACKNOWLEDGEMENTS

Throughout the course of my years as a PhD student, I have received support, assistance, advice and information from a number of people and wish to thank them all. However, there are a few who have out of their way to help me on a number of occasions and so deserve special recognition.

First and foremost, the research carried out in this thesis was supported by Eli Lilly and Company, and the Engineering and Physical Sciences Research Council through the Pharma-SEL Prosperity Partnership (grant numbers: EP/T005556/1 and EP/T518207/1). Therefore, I would like to thank both organisations for allowing me to have the amazing opportunity to continue my studies and move into research within the field of pharmaceutical peptides.

I would also like to extend my gratitude to my academic supervisor Prof. Daryl Williams, who provided continuous encouragement throughout the course of this work regardless of how despondent I became about the outlook. I will be forever indebted to him for his willingness to offer time and support whenever I was struggling either experimentally or personally. Additionally, I want to add particular thanks for his assistance during the write-up process, including proofreading drafts of this thesis.

I would also like to thank all the past and present members of the Surface and Particle Engineering Laboratory at Imperial College during my studies. The engaging and insightful conversations we shared, including brainstorming alternative research approaches or new avenues for research when barriers were encountered, greatly enriched the experience of my studies. I was truly grateful to be part of such a supportive team. In particular, three members of the group deserve specific recognition. Firstly, I wish to thank Kam Yenh Hua, who offered advice with regards to self-interaction chromatography and column packing procedures. Second, I offer my gratitude to Jiaxu Li, who provided assistance to me in the laboratory, especially with regards

to data transfer when I was unavailable. Thirdly, I wish to express my enormous gratitude to Dr Oscar Mercado Valenzo, for not only the peer support he offered throughout the course of our PhD studies but also all the advice and suggestions he offered with regards to chromatographic techniques. I wish to also recognise all the work he provided in ensuring the laboratory operated efficiently especially for all the HPLC maintenance that he carried out.

Finally, I wish to thank my family and close friends for not only offering love and encouragement throughout the course of my studies but also bearing with me when I was overwhelmed. In particular, I wish to thank my parents for never being more than a phone call away and offering me a peaceful environment in which to write my thesis, and my eldest brother and sister-in-law, for everything they did to support me during my final year.

ABSTRACT

Peptide therapeutics normally require parenteral administration, which can cause fear, pain, potential tissue damage and an increased risk of infection. Therefore, extensive research has been conducted into developing strategies which overcome the barriers faced in the oral delivery of peptide therapeutics. These barriers include the susceptibility of peptide structures to denature or aggregate in the gastrointestinal tract as well as their proteolytic degradation via the multiple peptidases present. Traditional cell and animal screening tests are slow and expensive, driving the needs for alternative screening approaches to be evaluated.

Self-interaction chromatography (SIC) represents an automated, high-throughput alternative technique of determining peptide propensity to aggregate through the chromatographic determination of the osmotic second virial coefficient. This thesis describes how SIC has been adapted for studying the stability and aggregation potential of peptide therapeutics under human intestinal conditions, and this data was compared to conventional dynamic light scattering and size exclusion chromatography measurements. The results of investigations with semaglutide and liraglutide showed that the second virial coefficient measurements were able to recognise transitions in glucagon-like-peptide-1 (GLP-1) therapeutics physical behaviour, but required prior understanding of the peptide oligomeric state behaviour in solution. When applied to insulin, it was found that the technique required more adaptations, as insulin is a peptide which readily formed stable oligomers.

This thesis also outlines the development and implementation of a covalently immobilised pepsin reactor. A study using peptide hydrophobicity standards found that these immobilised enzymes maintained their selectivity relative to solution reaction-based assays. However, a greater than 50-fold decrease in activity and 5-fold reduction in substrate affinity was observed, which was compensated for by a 500-fold increase in concentration. Additionally, the

immobilised pepsin showed good stability over a 12-week period with an average reduction in activity and substrate-affinity of 5.9% and 4.8%, respectively, between weeks 2 and 8.

LIST OF PUBLICATIONS

Conference Presentations/Posters

Bird, M. S. F., Williams, D. R., Self-Interaction chromatography as a tool for determining oral peptide stability under GI conditions. TIDES USA, Boston, MA, USA, 2024

Bird, M. S. F., Williams, D. R., Self-Interaction chromatography as a tool for determining oral peptide stability under GI conditions. AIChE Annual Meeting, Orlando, FL, USA, 2023

Bird, M. S. F., Williams, D. R., Chromatographic Analysis of Enzymatic Degradation of Peptides under Gastric Conditions. Pharmaceutical Manufacturing Forum, London, UK, 2023

Manuscripts in Preparation

Bird, M. S. F., Li, J., Lombardi, L., Williams, D. R., Oral Peptide Delivery Strategies and their Screening Techniques, in preparation

Bird, M. S. F., Williams, D., Self-Interaction chromatography as a Predictor of GLP-1 Aggregation Propensity, in preparation (Adapted from Chapters 7 & 8)

Bird, M. S. F., Williams, D., Immobilised Pepsin Reactors: An Oral Peptide Screening Tool, in preparation (Adapted from Chapters 9)

TABLE OF CONTENTS

1. Introduction 27

1.1. Peptide Therapeutics 27

1.1.1. Peptides versus large molecule biologics versus small molecules 27

1.1.2. Peptide Therapeutics Today 30

1.2. Oral Peptide Delivery 31

1.2.1. Drivers for oral peptide delivery 31

1.2.2. Barriers to oral peptide delivery 32

1.2.3. Oral peptide delivery strategies 35

1.3. Drug discovery phase of peptide therapeutic development 37

1.4. Thesis Outline 39

2. Peptide aggregation: a review 42

2.1. Peptide Structure 42

2.2. Fibrillation and Aggregation Mechanism 43

2.2.1. Peptide Aggregate forms 43

2.2.2. Oligomers 44

2.2.3. Micellular Structures 44

2.2.4. Amyloid Fibrils 45

2.2.5. Amorphous Aggregates 46

2.3. Factors affecting aggregation 47

2.3.1. Peptide Sequence 47

2.3.2. Peptide Modifications 48
2.3.3. Peptide concentration 50
2.3.4. pH 51
2.3.5. Temperature 52
2.3.6. Excipients 53
2.4. Peptide Stability Measurement Techniques 56
2.4.1. Dynamic Light Scattering (DLS) 57
2.4.2. Size Exclusion Chromatography (SEC) 60
2.4.3. Self-interaction Chromatography (SIC) 61
3. Gastrointestinal proteolysis & conditions: a review 65
3.1. Structure of the Gastrointestinal Tract 65
3.1.1. Gastric Conditions 66
3.1.2. Intestinal Conditions 67
3.1.3. Simulated Gastrointestinal Fluids 68
3.2. Gastrointestinal Enzymes 69
3.2.1. Protease Mechanism of Action 71
3.2.2. Condition Dependence of Enzyme Activity 72
3.2.3. Enzyme Inhibitors 74
3.2.4. Effect of peptide structure 75
3.3. Proteolytic Activity Measurement Techniques 76
3.3.1. Immobilised Enzymes 77
4. Research gap and thesis objectives 84

4.1. Research Gap 84
4.2. Thesis Objectives 86
5. Materials & methods 88
5.1. Materials 88
5.1.1. Insulin 88
5.1.2. Glucagon-Like Pepide-1 (GLP-1) Analogues 89
5.1.3. Peptide Hydrophobicity Library 93
5.1.4. Other Peptides 93
5.1.5. Pepsin 97
5.1.6. Column Packing Resins 101
5.1.7. Other Materials 102
5.2. Methods 103
5.2.1. High-Performance Liquid Chromatography (HPLC) Equipment 103
5.2.2. UV-Vis spectroscopy 105
5.2.3. Resin preparation procedures 106
5.2.4. Column Packing 113
5.2.5. Peptide Aggregation Determination Methods 116
5.2.6. Proteolytic Reactions 123
5.3. Error Analysis 127
5.3.1. UV-Vis spectroscopy 127
5.3.2. Self-interaction chromatography 127
5.3.3. Dynamic light scattering 129

5.3.4. Size-exclusion chromatography 129

5.3.5. Proteolytic Reaction Rates 130

6. Self-interaction chromatography – Stationary phase analysis 131

6.1. Introduction 131

6.2. Materials and Methods 132

6.3. Results & Discussion 133

6.3.1. Peak Elution Analysis 133

6.3.2. End-capping Comparison 136

6.3.3. Flowrate studies 142

6.3.4. pH variation study 144

6.4. Conclusions 146

7. Self-interaction Chromatography of a model peptide - Semaglutide 148

7.1. Introduction 148

7.2. Materials & Methods 150

7.3. Results 151

7.3.1. Semaglutide Stability 151

7.3.2. Semaglutide Lipid Stability Analysis 161

7.4. Conclusions 170

8. Self-interaction chromatography of Insulin and Liraglutide 172

8.1. Introduction 172

8.2. Materials & Methods 173

8.3. Results 174

8.3.1. Liraglutide Stability 174
8.3.2. Insulin Stability 190
8.4. Conclusions 209
9. Gastric Proteolytic Digestion of Peptides via immobilised pepsin reactors 211
9.1. Introduction 211
9.2. Materials and methods 213
9.3. Resin selection and pepsin immobilisation 214
9.4. Initial qualitative analysis of immobilised pepsin activity 216
9.5. Expanding to a complex peptide structure - Semaglutide 224
9.6. Varied Column Loading 232
9.7. Immobilised versus Free Solution Activity 234
9.8. Column Ageing Study 238
9.9. Expanded Peptide study 240
9.9.1. GLP-1 Analogues 240
9.9.2. Insulin 245
9.9.3. Oxytocin and Vasopressin Analogues 250
9.9.4. Bradykinin Fragments 252
9.9.5. Substance P 254
9.9.6. Gonadotropin-releasing hormone analogues 255
9.9.7. Summary of Observations 257
9.10. Conclusions 258
10. Conclusions and Future Work 261

10.1. Summary and Conclusions 261

10.2. Future Work 265

11. References 270

12. Supplementary Materials 295

12.1. Calibration curves 295

12.1.1. Peptide immobilisation solutions 295

12.1.2. Pepsin immobilisation solutions 295

12.1.3. Peptide self-interaction chromatography concentrations 296

12.2. Example elution curves 296

12.3. Size-exclusion chromatography calibration curve 297

12.4. Example raw dynamic light scattering correlograms 297

12.4.1. Semaglutide 297

12.4.2. Liraglutide 299

12.4.3. Insulin 301

12.5. Exemplary Mass Spectra 302

LIST OF FIGURES

Figure 2.1: Schematic representation of general pathways of aggregation and fibrillation 43

Figure 2.2: Diagram representing the key stages of peptide fibrillation 46

Figure 2.3: Schematic of Dynamic Light Scattering Experimental Setup 58

Figure 2.4: Representative schematic of the process of separation in Size Exclusion Chromatography 60

Figure 2.5: Schematic of idealised self-interaction chromatography packed column setup 62

Figure 3.1: Illustration of the human gastrointestinal tract with labelled relevant sections 66

Figure 3.2: Illustration of the relationship between Proteolytic activity and Temperature 74

Figure 3.3: Illustration of the four primary methods of enzyme immobilisation 78

Figure 5.1: Visual representation of the three-dimensional structure of human insulin 88

Figure 5.2: Representation of the human insulin amino acid sequence including cysteine bridges 89

Figure 5.3: Amino acid sequences of GLP-1 and its analogues liraglutide and semaglutide 90

Figure 5.4: Structure of modified Alanine residue Amino-isobutyric acid (Aib) 92

Figure 5.5: Amino acid sequence of taspoglutide with differences to GLP-1 highlighted in pink and red 94

Figure 5.6: Amino acid sequences of bradykinin (2-9) and [Lys-des-Arg^{9}]-bradykinin 95

Figure 5.7: Amino acid sequence of substance P 95

Figure 5.8: Amino acid sequences of [Lys^{8}]-vasopressin, terlipressin and oxytocin 96

Figure 5.9: Amino acid sequence of Luteinising Hormone releasing hormone (LHRH) and its synthetic analogue Goserelin 97

Figure 5.10: Structures of modified amino acids present in gonadotropin-releasing hormones 97

Figure 5.11: Schematic representation of three-dimensional structure of human and porcine pepsin 98

Figure 5.12: Acid-base nucleophilic mechanism of peptide bond hydrolysis by pepsin 99

Figure 5.13: Schematic of LC-20 HPLC system setup 103

Figure 5.14: Schematic of Nexera HPLC system setup 104

Figure 5.15: Nucleophilic substitution mechanism of the tresyl functionalised surface groups with primary amine groups 107

Figure 5.16: Proposed nucleophilic addition mechanism of glutaraldehyde to unfunctionalised surface alcohol groups 108

Figure 5.17: Mechanism for Schiff's base (imine) formation between aldehyde and primary amine groups 109

Figure 5.18: Mechanism of Schiff's base (imine) reduction to a Secondary amine via a Cyanoborohydride reducing agent 109

Figure 5.19: Senterra II Raman Microscope 20x objective lens focussed image of AF-Tresyl-650M resin bead 113

Figure 5.20: Diagrammatic chromatographic elution curve to represent the parameters required for peak tailing calculations 115

Figure 5.21: Illustrative chromatographic elution curve depicting the parameters required to calculate the column HETP 119

Figure 5.22: Example peptide elution profile through a packed immobilised enzyme column 124

Figure 5.23: Illustration of the pepsin solution reaction workflow 125

Figure 6.1: Acetone elution profile comparison for four AF-Tresyl-650M column stationary phases 134

Figure 6.2: Acetone void volumes versus NaCl Concentration at (a) pH 6.0 and (b) pH 8.0 135

Figure 6.3: Chromatogram comparisons for 10 µL injection of semaglutide in UV-Vis detectors 136

Figure 6.4: Retention volume ratios of semaglutide and acetone 140

Figure 6.5: Van Deemter Plot of the variation in the efficiency of 0.1 mg/mL semaglutide elution with mobile phase velocity 142

Figure 6.6: Retention volume ratios of semaglutide and acetone 144

Figure 7.1: Dynamic Light Scattering radius size distributions for 0.5 mg/mL semaglutide 152

Figure 7.2: Plot of semaglutide hydrodynamic radius under various simulated gastric conditions 153

Figure 7.3: Plot of semaglutide hydrodynamic radius as measured via DLS 155

Figure 7.4: Averaged Raman spectra of Tris-end-capped AF-Tresyl-650M resins with (red) and without (blue) semaglutide immobilised onto the surface 156

Figure 7.5: Example dynamic light scattering measurements of 0.1 mg/mL semaglutide in water 158

Figure 7.6: Elution chromatograms of 0.1 mg/mL semaglutide at pH 8.0 and pH 6.0 159

Figure 7.7: Plot of the Second Osmotic Virial Coefficient of semaglutide 161

Figure 7.8: Exemplary Dynamic Light Scattering size distributions for 0.5 mg/mL semaglutide 163

Figure 7.9: Plots of hydrodynamic radii of 0.5 mg/mL semaglutide species 165

Figure 7.10: Elution profiles of 2% acetone and 0.1 mg/mL semaglutide through a semaglutide-fixated SIC column 167

Figure 7.11: Plots of the second osmotic virial coefficient of semaglutide 169

Figure 8.1: Example dynamic light scattering triplicate measurement of 0.5 mg/mL liraglutide 175

Figure 8.2: Plot of liraglutide hydrodynamic radius as measured via DLS 179

Figure 8.3: Example DLS measurements of 0.5 mg/mL liraglutide after 24 hours incubation 180

Figure 8.4: Example SEC elution curves of 0.5 mg/mL liraglutide samples 181

Figure 8.5: Plots of SEC estimated molecular weights of 0.5 mg/mL liraglutide 184

Figure 8.6: Averaged Raman spectra of Tris-end-capped AF-Tresyl-650M resins with (red) and without (blue) liraglutide immobilised onto the surface 185

Figure 8.7: Example dynamic light scattering measurements of 0.1 mg/mL liraglutide in water 187

Figure 8.8: Example elution profiles of 2% Acetone solution (red) and 0.1 mg/mL liraglutide through a liraglutide fixated stationary phase at pH 6.0 (blue) and pH 8.0 (Green) 188

Figure 8.9: Plot of the Second Osmotic Virial Coefficient of liraglutide 190

Figure 8.10: Example DLS measurements of 0.5 mg/mL insulin 191

Figure 8.11: Plots of fraction of insulin species with a hydrodynamic radius <5 nm 193

Figure 8.12: Plots of radius of insulin species (<5 nm) at 0.5 mg/mL concentration 196

Figure 8.13: Example SEC elution profile of 0.5 mg/mL insulin 197

Figure 8.14: Plots of SEC estimated molecular weights of 0.5 mg/mL insulin 199

Figure 8.15: Averaged Raman spectra of tris-end-capped AF-Tresyl-650M resins with (red) and without (blue) insulin immobilised onto the surface 201

Figure 8.16: Example dynamic light scattering measurements of 0.1 mg/mL insulin in water 202

Figure 8.17: Dead column (a) injection and (b) elution profiles of 0.1 mg/mL insulin (unlabelled) or 2% v/v Acetone 204

Figure 8.18: Example elution profiles of 0.1 mg/mL insulin through an insulin fixated AF-Tresyl-650M packed column 206

Figure 8.19: Plots of the change in the second virial coefficient of insulin with NaCl concentration 208

Figure 9.1: Raman Spectra of Glutaraldehyde-activated and pepsin-functionalised propylamino silica resin 216

Figure 9.2: Mass spectra of hydrophobicity standard peptides P1-P6 (A-F) 218

Figure 9.3: Mass spectra of P1-P3 (A-C) 220

Figure 9.4: Mass spectra of P5-P6 (A-B) 222

Figure 9.5: Mass spectra to illustrate the fragmentation patterns of proteolytic degradation products 223

Figure 9.6: The LC-MS base peak profiles of the pepsin proteolytic degradation profiles of P4 224

Figure 9.7: Mass Spectra of semaglutide 225

Figure 9.8: Example LC-MS base peaks of proteolytic degradation products of 0.1 mg/mL semaglutide 227

Figure 9.9: Reaction profiles of semaglutide and three of its proteolytic degradation fragments 229

Figure 9.10: Reaction profiles of two of the proteolytic degradation fragments of semaglutide 230

Figure 9.11: Example base peak of semaglutide proteolytic degradation products after 30 minutes incubation 231

Figure 9.12: Variation in LC-MS detected relative abundance of semaglutide fragments 234

Figure 9.13: Lineweaver-Burk plots to establish the proteolytic activity of (A) free solution and (B) immobilised pepsin 235

Figure 9.14: Variation in (A) maximum reaction velocity and (B) Michaelis-Menten constant of immobilised pepsin 239

Figure 9.15: LC-MS base peak chromatogram of semaglutide after elution through an immobilised pepsin reactor 241

Figure 9.16: LC-MS base peak chromatogram of liraglutide after elution through an immobilised pepsin reactor 242

Figure 9.17: LC-MS base peak chromatogram of taspoglutide after elution through an immobilised pepsin reactor 243

Figure 9.18: Amino acid sequences with immobilised pepsin proteolytic cleavage sites highlighted of GLP-1 analogues 245

Figure 9.19: LC-MS base peak chromatogram of insulin after elution through an immobilised pepsin reactor 246

Figure 9.20: Human and bovine insulin sequences and their proteolytic degradation sites 248

Figure 9.21: Lys^{8}-vasopressin, terlipressin and oxytocin sequences and their predicted pepsin proteolytic cleavage sites 250

Figure 9.22: LC-MS base peak chromatograms of (A) oxytocin, (B) [Lys^{8}]-vasopressin and (C) terlipressin 251

Figure 9.23: LC-MS base peak chromatogram of (A) bradykinin(2-9) and (B) Lys-des-Arg^{9}-bradykinin after elution through an immobilised pepsin reactor 253

Figure 9.24: Bradykinin (2-9) and Lys-des-Arg^{9}-bradykinin sequences and their predicted pepsin proteolytic cleavage sites 253

Figure 9.25: LC-MS base peak chromatograms of substance P after elution through a column packed with (A) an unfunctionalised and (B) a pepsin immobilised propylamino silica resin 254

Figure 9.26: Substance P amino acid sequence and its proteolytic degradation sites 255

Figure 9.27: LC-MS base peak chromatogram of (A) Luteinizing Hormone releasing hormone (LHRH) and (B) Goserelin after elution through an immobilised pepsin reactor 256

Figure 9.28: LHRH and Goserelin sequences and their proteolytic degradation sites 257

Figure 12.1: Plotted calibration between UV-Vis absorption at 280 nm and concentration of (a) semaglutide, (b) liraglutide and (c) insulin in immobilisation solution 295

Figure 12.2: Plotted calibration between UV-Vis absorption at 280 nm and concentration of pepsin 295

Figure 12.3: Plotted calibration between UV-Vis absorption at 280 nm and concentration of (a) semaglutide, (b) liraglutide and (c) insulin in deionised water 296

Figure 12.4: Example elution profile of 0.1 mg/mL semaglutide in pH 5.0 mobile phase 296

Figure 12.5: Plotted calibration between protein standard mixture molecular weights and retention time in a TSKgel G3000SW$_{XL}$ column 297

Figure 12.6 Correlograms from dynamic light scattering measurements of 0.5 mg/mL semaglutide 297

Figure 12.7: Correlograms from dynamic light scattering measurements of 0.1 mg/mL semaglutide 298

Figure 12.8: Raw correlograms of dynamic light scattering of 0.5 mg/mL semaglutide 299

Figure 12.9: Correlogram of dynamic light scattering measurements of 0.5 mg/mL liraglutide 299

Figure 12.10: Correlograms of dynamic light scattering measurements of 0.5 mg/mL liraglutide 300

Figure 12.11: Correlograms from dynamic light scattering of 0.1 mg/mL liraglutide in water 300

Figure 12.12: Correlograms of dynamic light scattering measurements of 0.5 mg/mL insulin 301

Figure 12.13: Correlograms of dynamic light scattering measurements of 0.1 mg/mL insulin 301

Figure 12.14: Mass spectra of the major peaks of semaglutide after elution through an immobilised pepsin reactor 302

Figure 12.15: LC-MS (A) Base peak chromatogram and (B) primary peak spectrum of liraglutide after elution through an end-capped propylamino silica packed column 303

Figure 12.16: Mass spectra of the major peaks of liraglutide after elution through an immobilised pepsin reactor 304

Figure 12.17: LC-MS (A) Base peak chromatogram and (B) primary peak spectrum of taspoglutide after elution through an end-capped propylamino silica packed column 305

Figure 12.18: Mass spectra of the major peaks of taspoglutide after elution through an immobilised pepsin reactor 305

Figure 12.19: Mass spectra of the major peaks of insulin after elution through an immobilised pepsin reactor 306

Figure 12.20: Exemplary mass spectra of the major peaks of vasopressin after elution through a column packed with ethanolamine end-capped propylamino silica resin (A-C) 307

Figure 12.21: Mass spectra of the major peaks of terlipressin after elution through an immobilised pepsin reactor 307

Figure 12.22: Mass spectra of the major peaks of substance P after elution through (A) a column packed with ethanolamine end-capped propylamino silica resin and (B-D) an immobilised pepsin reactor 308

LIST OF TABLES

Table 1.1: Comparison of the properties of peptide therapeutics relative to small molecule and biologic drugs 29

Table 3.1: Gastrointestinal proteolytic enzymes and details of their secretion site and specificity 71

Table 5.1: RPLC peptide retention standard sequences and molecular weights [321] 93

Table 5.2: Porcine pepsin residues interacting with residues of an ideal heptapeptide substrate 100

Table 5.3: General effect substrate residue location has on likelihood of proteolytic activity of pepsin 101

Table 5.4: Toyopearl AF-Tresyl-650M properties 102

Table 5.5: Propylamino Silica resin properties 102

Table 6.1: Pre- and Post-column UV-Vis Absorption area ratios 139

Table 7.1: Measured and calculated parameters used for calculation of semaglutide osmotic second virial coefficient (B_{22}) 157

Table 7.2: Updated parameters used for the new self-interaction chromatography column 166

Table 8.1: Parameters used for calculation of liraglutide osmotic second virial coefficient (B_{22}) 186

Table 8.2: Calculated parameters applied in the calculation of the osmotic second virial coefficient (B_{22}) of insulin 201

Table 9.1: Varied reaction conditions and resultant pepsin loading concentrations 232

Table 9.2: Observed MS peaks of semaglutide after elution through an immobilised pepsin reactor 241

Table 9.3: Observed MS peaks of liraglutide after elution through an immobilised pepsin reactor 242

Table 9.4: Observed MS peaks of taspoglutide after elution through an immobilised pepsin reactor 243

Table 9.5: Observed MS peaks of insulin after elution through an immobilised pepsin reactor 246

Table 9.6: Summary of observed immobilised pepsin reactions with each peptide compared to sequence analysis 257

Table 9.7: Updated list of the effects substrate residue location has on the likelihood of proteolytic activity of pepsin 258

TABLE OF NOMENCLATURE

SYMBOL	MEANING
Aib	Amino isobutyric acid
API	Active Pharmaceutical Ingredient
B_2	Dimensionless Osmotic Second Virial Coefficient
B_{22}	Osmotic Second Virial Coefficient
B_{HS}	Volume of Exclusion
CD	Circular Dichroism
CLEA	Cross-linked Enzyme Aggregate
DLS	Dynamic Light Scattering
EMA	European Medicines Agency
FDA	US Food and Drug Administration
FTIR	Fourier Transform Infrared Spectroscopy
GI	Gastrointestinal
GIT	Gastrointestinal Tract
GLP-1	Glucagon-Like Peptide-1
HETP	Height equivalent to a Theoretical Plate
HPLC	High Performance Liquid Chromatography
LC-MS	Liquid Chromatography Mass Spectrometry
LHRH	Luteinizing Hormone Releasing Hormone
MALS	Multi Angle Light Scattering
MEA	Monoethanolamine
NMR	Nuclear Magnetic Resonance
R_H	Hydrodynamic Radius

SAXS	Small angle X-ray Scattering
SEC	Size Exclusion Chromatography
SGF	Simulated Gastric Fluid
SGIF	Simulated Gastrointestinal Fluid
SIC	Self-Interaction Chromatography
SIF	Simulated Intestinal Fluid
SLS	Static Light Scattering
SNAC	Sodium N-[8-(2-hydroxybenzoyl) amino] caprylate
PS	Surface Density
Φ	Phase Ratio

CHAPTER 1

1. INTRODUCTION

1.1. Peptide Therapeutics

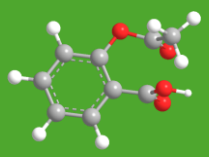
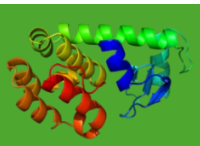
1.1.1. Peptides versus large molecule biologics versus small molecules

Peptides are a naturally occurring class of organic molecules with a structure consisting of chains of amino acids, typically 10-50, linked together by amide bonds [1]. Endogenous peptides act as signalling molecules for various physiological functions within the body, ranging from hormones to neurotransmitters such as glucagon-like peptide-1 (GLP-1) [2] and substance P [3], respectively. Therefore, due to the intrinsic and specific receptor-binding action of their structures, they represent an attractive starting point for the development of novel therapeutics with a safe toxicological profile and high efficacy. The resultant peptide therapeutics may be categorised into three groups, according to their structure. First, native peptides, identical to the endogenous species such as insulin for the treatment of Type 1 Diabetes Mellitus [4]. Second, analogue peptides, modified variants of native peptides such as semaglutide to treat Type 2 Diabetes Mellitus and Obesity [5]. And third, heterologous peptides, novel sequences not found in nature such as Bivalirudin [6]. Native peptides are often not ideal for pharmacological use due to possessing either poor stability or short half-lives within the body, hence leading to the requirement to research and develop novel compounds [7].

Within pharmacology, typically therapeutics are categorised into two major classes, namely small molecule drugs and biologic therapeutics. The former of these refer to chemical compounds with a molecular weight of typically <700 Da [8], which may act to regulate biological processes by binding with receptors or forming complexes with target sites [9]. By contrast biological therapeutics are compounds containing active molecules produced by living

organisms and are typically large complex species (with molecular weights >5000 Da [10]) ranging from hormones to monoclonal antibodies and even vaccines. Small molecule therapeutics, due to their less complex structure, typically have a simpler manufacture procedure, which may be more easily modified if required [11], and so benefit from lower costs of production. However, biologics exhibit significantly lower stability and so require more sensitive handling during their manufacturing [12]. Since small molecules have lower molecular weights and more hydrophobic structures relative to biologic therapeutics, they possess higher membrane permeability and as such, in accordance with Lipinski's rule of 5, higher oral bioavailabilities [13]. Therefore, they may be mass produced in tablet or pill form, thereby increasing their commercial viability and reducing the cost to consumers relative to biologics, which require parenteral, and often direct, administration [14]. Despite their routes of administration, biologics typically achieve peak therapeutic concentrations more slowly than small molecules, although this is often accompanied by longer half-lives, which may overall make dosing control more challenging [12]. Due to their small size and thus surface area, small molecule drugs typically target only the active site or receptor and are not able to inhibit large surface interactions such as protein-protein interactions [15]. As such, they have significantly lower specificity relative to biologics and may have unforeseen, adverse side effects, such as distorting structures or off-target binding, which contribute to their often-poor toxicological profile [12]. Conversely, biologics have much higher specificity for their target sites due to their chemical structure and size allowing for more ergonomic interactions and so much stronger affinity. Furthermore, unlike small molecule drugs, their metabolism and excretion pathways are often relatively predictable, as they are much more akin to endogenous species, being broken down enzymatically and excreted via urine or faeces. This may explain in part why biologics were found to require an approximately 2.5-year shorter development time than small molecule drugs [16]. However, biologics still pose other risks such as immunogenic reactions or encephalopathy due to their hinderance of specific physiological processes [12].

Table 1.1: Comparison of the properties of peptide therapeutics relative to small molecule and biologic drugs

	 Small Molecules	 Peptides	 Large Molecule Biologics
Molecular Weight	< 700 Da	~1000 – 5000 Da	>5000 Da
Manufacturing Cost	Cheap (Chemical Synthesis)	Moderate – High (Solid-Phase Synthesis)	Very High (Biological production)
Storage	Stable at Room Temperature	May require refrigeration	Cold chain required
Delivery Route	All	Subcutaneous Intravenous Intramuscular (some Oral)	Subcutaneous Intravenous Intramuscular
Membrane Permeability	High	Limited	Sparingly
Immunogenicity	Low	Moderate	High
Toxicity	Common off-target effects	Low	Very Low (Generally)
Specificity	Moderate	High	Very High

Peptide therapeutics are a subclass of biologics distinct from large molecule biologics on account of their shorter amino acid sequences, resulting in their reduced ability to form stabilising tertiary and quaternary folding structures. Therefore, their biochemical and thus pharmacological properties differ, meaning that strategies that improve the therapeutic action of either small molecules or large molecule biologics, such as monoclonal antibodies, may not necessarily have the same desired effect for peptides. They are known to exhibit some properties more comparable to small molecules than larger molecule biologics, with some associated benefits and disadvantages (see Table 1.1). As they have the same chemical composition as other biologics, peptides have similar modes of action and so exhibit similarly high specificity

and efficacy, but with significantly lower immunogenicity [17]. Furthermore, since they are smaller with less complex structures, they a more cell penetrating [18], though relative to small molecule drugs they still have poor membrane permeability and are generally unable to cross the blood-brain barrier [19]. However, whilst still being more expensive than small molecule drugs, peptides <50 amino acids in size are significantly cheaper than large molecule biologics, such as monoclonal antibodies, to manufacture, not generally requiring recombinant techniques but utilising solid-phase synthesis methods pioneered by Merrifield [20]. Finally, due to peptides lacking a stabilising quaternary structure like larger protein biologic therapeutics, they are highly susceptible to proteolytic degradation and so rapidly cleared from circulation [21]. Therefore, without modifications peptides generally exhibit poor in-vivo stability relative to small molecule and larger biologic therapeutics.

1.1.2. Peptide Therapeutics Today

The research and clinical use of peptides as therapeutic agents began over a century ago with the first isolation and medical application of insulin to treat diabetes mellitus [22]. The clinical use of synthetic peptides became feasible in the 1960s with the development of the solid phase peptide synthesis method by Merrifield et al. [20]. However, due to the short half-lives of native peptide structures in the human body, their application for medical treatment purposes was hindered until the development of analogue species with improved pharmacokinetics. As of the beginning of 2025, from abalaparotide to zilucoplan there have been over 100 peptide therapeutics approved by the US Food and Drug Administration (FDA) [23] and European Medicines Agency (EMA) [24]. However, as of the third quarter of 2024, small molecule drugs are estimated to still constitute approximately 58% of the approximately $1.477 trillion pharmaceuticals market [25], whilst the peptide therapeutic market size is estimated to be significantly less at just under $50 billion [26]. Some estimates are as high as $117 billion when factoring in the rapidly growing market of GLP-1 obesity treatments [27]. It can be safely concluded that peptide therapeutics are considered to be a rapidly developing new drug

modality, with predicted compound annual growth rate of >12% [26, 27]. The small molecules and the pharmaceutical industry as a whole are observing only 7% and 9% growth, respectively [25].

1.2. Oral Peptide Delivery

Peptide drugs, due to their large size and typically hydrophilic structure, typically have poor oral bioavailability. Therefore, for the peptide therapeutics currently clinically available, parenteral delivery is estimated to comprise over 80% of the market in terms of the route of administration [26, 27], with only 12 (or 13 if odevixibat is included as a dipeptide structure) being commercially available for oral delivery. The total oral peptide drug market size is estimated to be about $4 billion, primarily due to the release of an oral semaglutide formulation, Rybelsus, in 2019 [28]. However, the oral delivery sector of the peptide therapeutic market is expected to experience rapid growth (>20% compound annual growth rate) due to the investment the sector is experiencing, primarily on account of the large incentive for oral delivery, which will be discussed subsequently.

1.2.1. Drivers for oral peptide delivery

There are three primary drivers to oral delivery over the conventional parenteral routes of peptide therapeutic administration: increased patient compliance, reduced healthcare costs and potentially improved pharmacodynamic properties. Patients may have an aversion to needles and generally prefer an oral dosing schedule [29] as continuous, repeated injections are reported to result in injection site reactions in 0.5-40% cases for all FDA-approved, injectable, biological agents thereby presenting a barrier to this delivery route [30]. Moreover, a survey of 1250 physicians found that 33.2% of their patients had reported not adhering to their insulin injection dosing schedule for at least one day over the previous month, with an average of 3.3 days [31]. Stress and public embarrassment were reported as two of the most common reasons for poor adherence, both of which may be reduced by the more convenient and simple oral

administration route. With respect to the second primary driver, a systematic review of drug administration costs by Tetteh and Morris found oral administration to be the lowest for biologic medicines [32]. This is primarily due to the lack of clinical professional time or equipment required in its administration since it can be self-administered in a pain-free manner without sterile precautions or medical supplies such as syringes and needles being required. Furthermore, a cost-effectiveness analysis comparing the oral and subcutaneous delivery of semaglutide found that the quality-adjusted life years gained through oral administration compensated for the increase in cost incurred due to the reduced bioavailability and so increased amount of active pharmaceutical ingredient [33]. Finally, oral delivery of specific species, depending on their target, may also offer pharmacological advantages. For example, oral insulin and GLP-1 analogues better mimic the physiological release pathway of their endogenous counterparts, which are both secreted into the hepatic portal vein from the pancreas and intestinal L-cells respectively [34, 35]. This reduces any effects of localised concentrations of the peptides and in the case of insulin, the possibility of hyperinsulinemia and so hypoglycaemia.

1.2.2. Barriers to oral peptide delivery

The structure of peptide therapeutics may vary significantly from small dipeptide species such as lisinopril to large lipidated GLP-1 analogues such as semaglutide or to cyclic species such as oxytocin. Therefore, the barriers for oral delivery are not the same for two peptides, nor is there a one-size-fits all strategy to circumvent these barriers. Furthermore, the dietary habits of every individual patient, and consequently the conditions within their gastrointestinal tract (GIT) vary, especially those with GIT-related diseases such as inflammatory bowel disease, bowel cancer or Crohn's disease, resulting in highly individualised drug absorption profiles [36]. Nevertheless, there are five primary barriers to the oral delivery of peptides: the variance in pH through the GIT, the presence of proteolytic enzymes within the GIT, the gastrointestinal (GI) mucus layer, the epithelial membrane and first-pass metabolism.

The pH varies significantly throughout the GI tract, which is challenging for the maintenance of peptide stability. At extreme pH values peptide secondary structures may be disrupted causing a potential loss in functionality, whilst at pHs close to their isoelectric point, the peptide may become susceptible to aggregation or a significant reduction in solubility. Therefore, a continuously changing pH profile is not ideal for ensuring a consistent structural profile. Beginning in the mouth, saliva ensures the maintenance of a relatively neutral pH within the range pH 6.7-7.3 [37]. However, upon reaching the stomach the pH drops significantly to pH 1.7-4.7, before being neutralised by carbonates and bile juices in the duodenum to reach pH 5.9-6.3 [38]. In the jejunum and the ileum, the conditions are slightly basic, measured in the pH 7.4-7.8 range, whilst the colon shows significant variance ranging from mildly acidic to slightly basic (pH 5-8). Measurements of the physical properties, constituents concentrations and pH of the GI fluid within the various sections of the GIT have been collated and reviewed thoroughly by Mudie et al. [39].

The primary function of the GIT is to digest and absorb nutrients and minerals whilst serving as a barrier to prevent the invasion of pathogenic species. Therefore, within the GIT there are several proteolytic enzyme species such as pepsin in the stomach and trypsin in the small intestine, the function of which is to degrade proteins into amino acids and di- and tripeptides for absorption. The proteolytic enzymes present in the GIT have been thoroughly reviewed by Pereira de Sousa et al. [40]. Since peptide therapeutics aim to mimic endogenous peptide species, they are comprised of the same amino acid structure as nutritional proteins and as such are susceptible to proteolytic degradation, thereby impacting their potential bioavailability. For example, in a study determining the proteolytic stability of seventeen peptide therapeutics in GI tract fluids, only three (cyclosporine, desmopressin, and octreotide) exhibited resistance to enzymatic degradation, all three of which have cyclic structures and are commercially available as oral formulations [41].

The epithelial cells within the GI tract are exposed to the external environment and as such have a continuous mucus layer which acts as a selectively permeable barrier for their protection [42]. The primary component of the mucus gel are the goblet-cell secreted mucin glycoproteins, which have a brush-like structure, with negatively charged glycosylated groups extending as bristles intermittently separated by hydrophobic cysteine-rich domains. These form a size-excluding gel matrix and due to their structure, impede the permeation of hydrophobic and positively charged peptides, thereby increasing their exposure to brush border enzymes [42]. As such, the mucus layer further reduces oral peptide bioavailability, especially for microparticle delivery strategies.

The epithelial membrane varies throughout the GIT, but the intestines is traditionally the target drug absorption site, composed predominantly of four cell types: enterocytes (>80%), mucin-secreting goblet cells, enzyme-secreting Paneth cells and microfold cells (<1%) [43]. These cells, bound together via transmembrane proteins forming tight-junctions, form a continuous monolayer between the lamina propria and the intestinal lumen, which serves as a semi-permeable barrier [44]. Peptide therapeutics are generally, but not always, too large to paracellularly permeate through the tight junctions, which typically have a pore size of between 8 and 10 Å, thereby inhibiting species with a molecular weight of more than 600 Da [45]. Moreover, peptides due to their high molecular weight and hydrophobicity generally have poor passive transmembrane permeability. The epithelial barrier remains the most significant barrier to oral peptide delivery.

Orally administered therapeutics, upon absorption through the GI membrane, enter the hepatic portal system, which delivers the absorbed species to the liver. The liver contains a number of proteolytic enzyme species which may metabolise the orally delivered peptide and so reduce the fraction of drug which reaches systemic circulation [46]. This behaviour is referred to as the first-pass effect and adds an additional barrier to the bioavailability of orally delivered therapeutics. However, of the five barriers discussed previously, the primary focus of the work outlined in this

thesis will relate to the first two, namely the variation in pH and conditions within the GI tract and how that impacts peptide physical stability of peptides and the susceptibility of peptides to proteolytic degradation.

1.2.3. Oral peptide delivery strategies

Oral delivery strategies typically aim to improve the bioavailability of the active pharmaceutical ingredient (API) by one of two methods or both. Firstly, they may either stabilise or shield the species against physical denaturation or degradation within the GI tract environment. Secondly, they may act to increase the permeability of the API across the GI epithelium. The specifics of the various techniques are outside the scope of the work performed in this thesis, but a brief, non-exhaustive summary shall be provided here for context. In this summary, the various oral peptide delivery strategies shall be divided into two subcategories, namely encapsulating and non-encapsulating methods. Within a given formulation, multiple approaches may be used in tandem. Alternative oral delivery technologies such as luminal unfolding microneedle injectors [47], self-orientating millimetre-scale applicators [48] and RaniPills [49] will not be discussed as they are in essence parenteral delivery methods administered orally.

Encapsulation strategies within oral delivery formulations involve loading the API into a nanostructure which primarily aims to shield the API from the GI environment but also may be functionalised to improve permeability. These methods include the use of enteric coatings as protective layers encasing an oral pill, which are commonly used in small molecule oral delivery to target intestinal release. Enteric coatings are composed of copolymer materials, which in the low pH of the gastric environment maintain their structure, but disintegrate as the pH increases in the intestines and their carboxylic acid groups are deprotonated, and so release the API [50]. Alternative encapsulation methods include the use of nanoparticles, which may be divided into polymeric and inorganic nanoparticles, the most common of which are chitosan and silica, respectively [51]. Polymeric nanoparticles are beneficial due to their versatility, as the polymer

and surface functionalisation may be easily modified to maximise their effectiveness. For example chitosan is beneficial as it interacts with epithelial serine residues, modulating tight junction opening [52] whilst its size distribution may be easily controlled [53]. Conversely, inorganic nanoparticles are not degraded within the human body and so accumulate [54], however their porosity is readily tuned, whilst silica has also been found to induce tight junction opening by binding to integrins on the epithelial surface [55]. An alternative encapsulation strategy are solid lipid nanoparticles, which are comprised of solid-lipid core loaded with API encased in a surfactant shell [56]. These have been found to interact strongly with epithelial cells [57] and to initiate a clathrin-dependent endocytosis pathway [58]. Self-(nano)emulsifying drug delivery systems (SEDDS) have been found to target a similar endocytosis mechanism [59]. SEDDS are formulations of lipids, cosurfactants and cosolvents that emulsify upon exposure to aqueous media to entrap a lipophilic API into oil droplets [60]. They have previously been applied successfully in the oral formulation of Cyclosporine A, Neoral [61].

Non-encapsulating oral delivery strategies, in this brief review, shall be divided into three subcategories: species that alter the GI tract environment, solubilising agents and permeation enhancers. The former of these includes species such as enzyme inhibitors and mucolytics. Enzyme inhibitors shall be discussed in more detail in Section 3.2.3. Mucolytics are species which act to break down mucus barriers for example via catalysing the breakdown of disulfide bonds within the mucin glycoprotein structure [62]. For example, formulation with N-acetylcysteine has previously been shown to improve intestinal permeability [63]. Examples of solubilising agents include ionic liquids and hydrophobic ion pairs. Ionic liquids are salts in liquid state under ambient conditions, and are often comprised of organic cations conjugated to anionic, with the most commonly studied example being choline and geranate [64]. They are suggested to stabilise the API and improve its solubility by forming a shell around it via electrostatic interactions [65]. Conversely, hydrophobic ion pairs are hydrophobic counter ions applied to improve the solubilisation of hydrophilic APIs within a lipid-based drug delivery system

[66]. Finally, permeation enhancers are species which act to increase the epithelial permeability of the API. They may be divided in accordance with their mechanism into transcellular and paracellular permeation enhancers. Transcellular permeation enhancers may be further divided into two groups: first, surfactants which disrupt the epithelial bilayer integrity, and second, species which form complexes with the API so as to increase its hydrophobicity and as such its passive transcellular permeability [67]. The most prevalent example of these is sodium N-[8-(2-hydroxybenzoyl)amino]-caprylate (SNAC), which is applied in the oral formulation of semaglutide [68]. Paracellular permeation enhancers disrupt paracellular tight junctions by altering intracellular signalling pathways (1st generation) or interfering with their intercellular interactions (2nd generation) [67]. The most notable paracellular permeation enhancer is sodium caprate, which acts via a 1st generation mechanism and has been used in the oral formulation of octreotide, Mycapssa [69].

1.3. Drug discovery phase of peptide therapeutic development

There are an estimated 5000-10000 drug candidates per approved therapeutic identified during the drug discovery process, with only approximately 250 compounds reaching the pre-clinical phase and 5 entering phase I clinical trials [70]. The cost of developing a novel therapeutic has been estimated to be on average approximately $2.83 billion (USD 2022) by Proudman et al. with pre-clinical research accounting for $692 million (USD 2022) [71]. The full procedure from target identity to therapeutic approval is outside the scope of the work in this thesis, but a brief summary of the drug discovery phase of drug development are required to understand the context of the work carried out.

Upon identifying and validating a target receptor, the drug discovery phase of the development of a therapeutic peptide begins. The drug discovery phase consists of three primary stages: hit identification, hit-to-lead selection and lead optimisation. The methods studied within this thesis aim to improve on those within the hit-to-lead selection stage, and so that stage will be analysed

most thoroughly. The other two stages shall be summarised briefly for context. Hit identification consists of screening massive libraries of peptide sequences within a specific combination space, perhaps based on natural peptides or rational design [72]. This is generally performed using display library technology [72], perhaps accompanied by deep-learning approaches [73], to rapidly identify sequences which have affinity for the target receptor. After hit-to-lead selection, lead optimisation primarily involves four objectives. One involves the modification of the lead structure and formulation to improve its absorption, distribution, metabolism and excretion kinetic pathways, such as for example the lipidation of GLP-1 analogues to enhance their albumin binding and reduce proteolytic degradation by dipeptidyl peptidase IV [74]. A second is initial scouting of the scalability of the manufacturing [75] and purification process of the peptide through consideration of its yield and impurity profile [76]. Thirdly, initial off-target safety pharmacology screening may also be initiated during this phase of drug development to minimise later candidate attrition [77]. However, the most important are the more detailed pharmacodynamic and pharmacokinetic studies, including initial, non-good laboratory practice, exploratory in-vivo trials to assess toxicology [78].

During the hit-to-lead stage, properties of the “hit” peptides are measured and their structures refined to “lead” candidates with improved potency, selectivity, stability and developability. The primary studies carried out during this phase of peptide therapeutic development include structure activity studies, which are performed by measuring target receptor binding of the are studied via several potential methods including surface plasmon resonance, radioligand binding, Förster resonance energy transfer or functional cell assays. Understanding of which residues are key for target binding is studied and optimised via d-amino acid scans, alanine scans or N- and C-terminal truncation [79]. In addition, solubility, conformational stability and aggregation propensity of the candidate peptides are studied during this phase of candidate development [80] with several of the experimental techniques outlined in Section 2.4. Since peptide solubility may limit formulation feasibility and dosage strength, formulation selection

also begins during this phase of development. The stability of the peptide structure to chemical degradation such as oxidation, deamidation, hydrolysis or thermal stress are also studied via mass spectrometry methods to identify weakness within the structure and potential degradation pathways [81]. Similarly, experiments incubating the candidate peptide in serum with specific proteases, representative of the targeted delivery route, to measure its proteolytic stability are also conducted to identify protease susceptible residue sites. These techniques are outlined in more detail in Section 3.3. Additionally, permeability assays, such as parallel artificial membrane permeability assays or cell line experiments (e.g. Caco-2 or MDCK cells) are carried out to ensure the target absorption and distribution pathways are feasible [82].

Given the high costs associated with developing a novel drug, even in the preclinical phases, and the "shotgun" approach applied during drug discovery, improving the candidate filtration process will likely result in significant reduced overall costs. Accurate experimental techniques with higher throughput during the hit-to-lead phase of drug discovery would allow for more rapid data collection of different structures and formulations, and so potentially a greater number of candidates screened. Therefore, this may allow for a reduction in the number of expensive in-vivo test required later during the pre-clinical phase of drug development and potentially reduce the number of failed candidate peptide therapeutics. Therefore, the work in this thesis aims to develop self-interaction chromatography and immobilised enzyme reactors as tools for studying peptide aggregation propensity and proteolytic stability, respectively, under GI conditions with a view to their potential application in oral peptide formulation development.

1.4. Thesis Outline

Chapter 2 discusses the structure of peptide therapeutics, their various modifications from native peptides and how these impact their propensity to aggregate. It further outlines the different aggregate structures and discusses the mechanisms through which they form. It finally reviews the different methods through which peptide aggregation may be measured, with a focus

on dynamic light scattering, size exclusion chromatography and self-interaction chromatography.

Chapter 3 reviews the structure of and solution conditions within the gastrointestinal tract before focussing on the GIT proteolytic enzymes and factors affecting their activity. Additionally, the techniques for measuring the proteolytic degradation of peptides are briefly discussed prior to reviewing the various types of immobilised enzyme reactors and methods of measuring their effectiveness. The subsequent Chapter concludes the literature review section of the thesis with a summary of the gap in the previous research and an outline of the thesis objectives.

Chapter 5 discusses the materials used throughout the experimental work, discussing the background to the sample peptides, with a brief explanation for their selection. In particular it discusses the selectivity and proteolytic mechanism of the enzymes used. It also describes the various equipment and experimental procedures followed throughout the course of experimental work described throughout the rest of the thesis.

Chapter 6 outlines the development of the self-interaction chromatography method with respect to its application for studying peptide analytes. It discusses the selection of the resin, end-capping procedure and optimisation of the chromatographic method, including parameters such as the flowrate.

Chapter 7 investigates the application of Self-Interaction Chromatography (SIC) for studying the aggregational propensity of a model peptide, semaglutide, under various simulated intestinal conditions. The results are compared with size distribution data gathered using dynamic light scattering data to establish whether the technique captures the correct trends. Chapter 8 expands the SIC study to two additional peptides, liraglutide and insulin. The results are then compared to size distribution data collected via dynamic light scattering and size exclusion chromatography.

Chapter 9 focuses on the development of an immobilised pepsin reactor and standard operating procedure. It establishes the immobilisation of pepsin before establishing the changes in selectivity and Michaelis-Menten parameters through comparison with solution data, before expanding the study to a library of peptides.

The thesis concludes with Chapter 10, which provides a summary of the results and recommendations for any potential future work.

CHAPTER 2

2. PEPTIDE AGGREGATION: A REVIEW

2.1. Peptide Structure

Peptides are oligomers or polymers of amino acids (known as residues) bound together by amide (peptide) linkages, which are partial double bonds between the carboxy group of one amino acid to the amino group of the subsequent amino acid [1]. The amino acids, which form the basic structural unit of the peptide chain, consist of an α-carbon covalently bound to a hydrogen, an amino group, a carboxyl group and a side chain commonly referred to as the R-group. There are 21 proteinogenic amino acids in eukaryotes with differing R-groups, however, often generally only the 20 canonical ones are listed as selenocysteine is rare and is not generally encoded within the human genome [83]. The difference in physicochemical properties of the various amino acids is due to the different chemical nature of the R groups, which vary in charge, hydrophobicity and size [84]. Within peptides and proteins, the amino acid sequence is commonly notated beginning from the N-terminal (free α-amino group) amino acid to the C-terminal (free α-carboxyl group) amino acid. Traditionally peptides with fewer than 10 residues are referred to as oligopeptides, whilst the term polypeptide is used to define those with 10-100 residues and proteins, those with more than 100 residues [1]. However, more recently, peptides with greater than 50 residues are often referred to as proteins and those with 10-50 residues as peptides. Protein and peptide structure may be classified into three to four tiers [85]. The amino acid sequence constitutes the primary structure of the peptide. The local structure of segments of the backbone of amino acids may form into periodic spatial arrangements referred to as the secondary structure. Under suitable conditions via intermolecular interactions the most commonly formed structures are α-helices and β-pleated sheets. These secondary structures will interact with one another through

intramolecular forces between the various amino acid side chains, causing the structures to fold further into a compact three-dimensional form, providing the overall shape of the individual peptide chain known as the tertiary structure. Proteins containing more than one polypeptide chain, differ from peptides in that they may also form quaternary structures, which refer to the spatial arrangement of multiple tertiary structures of individual polypeptide chains [84]. These complexes may be homogeneous or heterogeneous consisting of the same or different polypeptide chains, respectively.

2.2. Fibrillation and Aggregation Mechanism

2.2.1. Peptide Aggregate forms

The term aggregate, with respect to peptides, can refer to a whole range of multimolecular peptide species, ranging from the smallest dimers and oligomers to macromolecular fibrils and amorphous structures [86]. The final aggregate formed is dependent on a number of factors including the environmental conditions and the peptide sequence and its structural motifs. For peptide therapeutics these aggregate structures generally fall into four categories: oligomers, micellular structures, amyloid fibrils and amorphous aggregates. The general pathway of aggregation is outlined in Figure 2.1 with the formation of micellular structures absent as their formation is highly reliant on peptide structure.

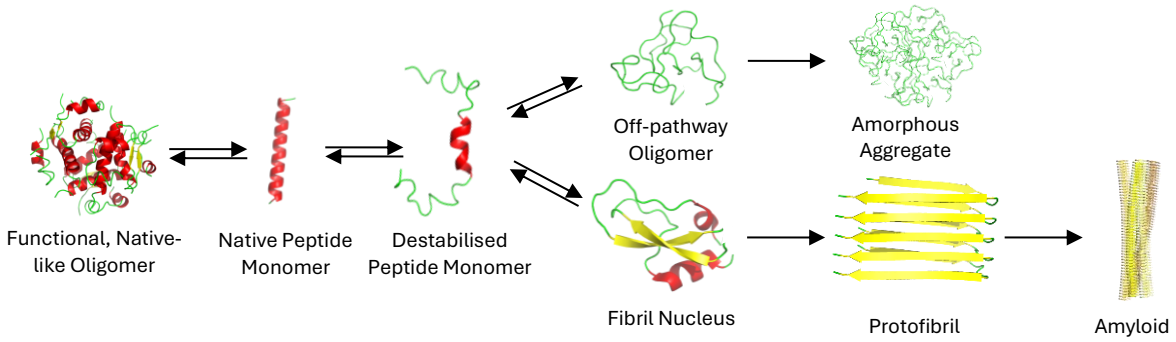


Figure 2.1: Schematic representation of general pathways of aggregation and fibrillation

2.2.2. Oligomers

Peptide oligomers are a poorly defined class of aggregates, but are generally classified as those that do not readily sediment in solution and are comprised of a limited number of monomers such that it has a relatively small size [86]. Oligomers may be categorised as either on-pathway or off-pathway. The former of these are essential intermediates and direct precursors to larger fibrillar species during peptide aggregation [87]. By contrast, off-pathway oligomers are distinct, metastable species with sufficient structural differences that there is an energy barrier to their being converted or incorporated into larger aggregates [88]. Certain peptides, under stable solution conditions but at sufficiently high concentrations, begin to associate to form stable oligomers. This occurs either via intermolecular interactions or via interactions with a specific excipient, with the individual peptides either maintaining their native structure or adopting a relatively similar conformation within the oligomer [89]. This process is generally reversible and provides stability to the peptide against degradation. An example of such a structure would be the hexamer of insulin, the structure of which is detailed in Section 5.1.1, or amyloid β (1-40) in the presence of Zn^{2+} ions [90].

2.2.3. Micellular Structures

Peptide amphiphiles consist of hydrophilic and hydrophobic chains of amino acids or more typically, a hydrophilic peptide bound to one or more lipid chains, the latter being a lipopeptide [91]. These peptides, under aqueous conditions and above a critical aggregation concentration, agglomerate to form spherical micelles due to the hydrophobic interactions of the hydrophobic regions and electrostatic repulsions of the hydrophilic regions [92]. At sufficiently high concentrations these species then cluster and merge via hydrophobic interactions to form wormlike micelles, which themselves coalesce into twisted ribbons and larger nanofibers [93]. An example of a therapeutic peptide exhibiting spherical micelle formation is daptomycin [94]. Typically, this type of self-assembly behaviour is intrinsic to the structure of the peptide therapeutic and as such would likely be a known property, prior to investigation [95]. However,

with lipidation being a common technique used in improving the half-life of peptide therapeutics, degradation processes may result in this behaviour being exhibited.

2.2.4. Amyloid Fibrils

Amyloid fibrils are aggregates composed of peptide monomers intermolecularly stacked to form cross-β sheet conformations, which arrange into protofibril structures, and in turn to fibrils, characterised by their linear structure [96]. They are typically the most thermodynamically stable conformation a given peptide agglomeration can exhibit; hence they readily form above a critical solution concentration [86]. Their formation typically begins with a peptide, which has a sufficient proportion of its sequence having a random unfolded secondary structure in its native state or for which the native structure is disrupted via solvent or excipient interactions or via chemical reaction resulting in a partially-folded state with an exposed peptide backbone [97-99]. The exposed peptide backbones then reversibly self-assemble into small oligomer structures until reaching a critical size, known as the nucleus of aggregation, for which the rate of aggregation becomes faster than dissociation (primary nucleation) [86] (see Figure 2.2). These oligomers then readily aggregate via β-sheet like interactions to form larger oligomers, protofibrils and fibrils [86, 87, 89]. The surface of these fibrils may then interact with unstable monomers to catalyse the formation of new nuclei (secondary nucleation) or the fibrils themselves may fragment to form nuclei [86, 100]. Therefore, the formation rate of peptide fibrils is characterised by a lag phase, followed by a growth phase and a plateau, with either primary or secondary nucleation being responsible for the growth phase [101, 102]. More detailed analysis of the mechanism and kinetics of fibril formation has been extensively reviewed [86, 89, 100, 102, 103] but is beyond the scope of the work carried out in this thesis.

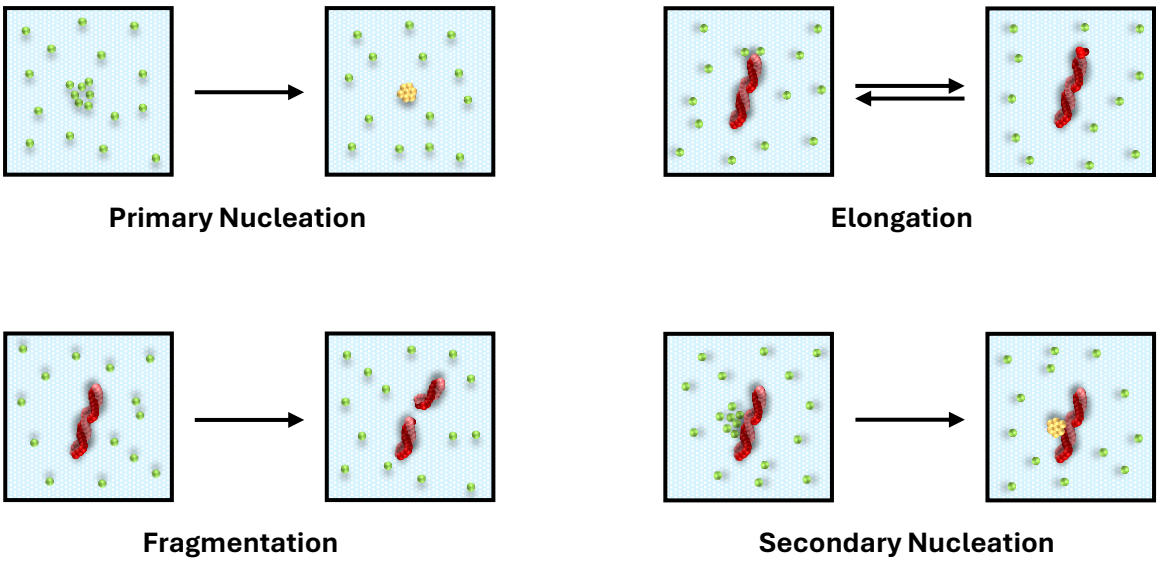


Figure 2.2: Diagram representing the key stages of peptide fibrillation

2.2.5. Amorphous Aggregates

The term amorphous aggregate generally corresponds to large assemblies of peptides with an irregular structure [86]. However, various non-fibrillar aggregates have been shown to have significant degree of structure to their conformations [96, 104, 105], showing similar β-structures to amyloid fibrils [105], exemplified by their ability in some cases to act as precursors to fibrils [96]. Therefore, in terms of peptide aggregation, it is perhaps more accurate to describe non-linear, amorphous aggregates as non-fibrillar. They differ from oligomeric structures in that amorphous aggregates are typically orders of magnitude larger than the monomeric peptide with potential to form deposits [86]. Their formation generally proceeds similarly to that of amyloid fibrils, beginning with a peptide monomer with irregular, random-coil-like structures either due to instability or that being the peptide's native state [89]. These monomers agglomerate via hydrophobic interactions or backbone hydrogen-bonding to form disordered oligomers which may reorganise and compact to form β-sheet structures, characteristic of fibril precursor structures [97-99] or which further associate into the larger amorphous structures discussed [89].

2.3. Factors affecting aggregation

2.3.1. Peptide Sequence

Peptides due to their smaller size relative to proteins and lack of quaternary structure in their monomer state, have a reduced number of electrostatic interactions, hydrogen-bonds and van der Waals forces maintaining their conformation and secondary structures. Therefore, they additionally have relatively less shielding of their hydrophobic groups from the aqueous environment, increasing the probability of and enabling stronger interactions between monomers when in solution. Therefore, individual alterations to a peptide sequence may not just perturb its folding conformation, but also its aggregation propensity by changing the strength of hydrophobic, van der Waals and electrostatic intermolecular interactions [106-108]. It has been shown that specific segments of the amino acid chain may drive the initiation of the aggregation process [109, 110]. These sequence segments are known as aggregation-prone regions (APRs). They are generally 5 to 15 residues long and are comprised of hydrophobic rich, low net charge amino acids that have a high native tendency to form β-sheet structures and may even induce aggregation upon insertion into a peptide chain [111-114]. However, the presence of an APR itself does not necessarily imply peptide aggregation, as the sequence may be shielded by structural folding [115]. Nevertheless, APRs do initiate aggregation upon sufficient peptide unfolding and interaction with the solvent medium [116]. The amyloid aggregation propensity of individual amino acids has been tabulated by Pawar et al. with the aromatic residues (other than Histidine) and cysteine showing the highest proclivity and the charged residues the lowest [117]. Despite this, aggregation has been shown not to be solely dependent on hydrophobicity or even the β-sheet propensity of the sequence [118]. Typically, therapeutic peptides are designed without contiguous stretches of hydrophobic residues or APRs. However, another method of aggregation prevention is through the placement of gatekeeper residues which have charged side chains (e.g. Arg, Lys, Glu, Asp), which act to counteract aggregation through charge repulsion or through

steric hinderance [119-121]. Additionally, Pro and Gly may also act to counteract aggregation via their incompatibility with the β-structure [114, 120].

2.3.2. Peptide Modifications

As outlined in the previous subsection, the physical stability of peptides is highly dependent on their structure. However, in the design of therapeutic peptides, several strategies have been developed to increase the bioavailability, potency or half-life of the active pharmaceutical ingredient (API) via modification of its structure. These range from major structural additions such as lipidation or glycosylation or conformational changes via stapling, cyclisation or disulfide bonds, to minor changes such as amino acid modification. Glycosylation [122], conjugation with polyethylene glycol (PEG) otherwise known as PEGylation [123], phosphorylation [124, 125], acetylation [126], cyclisation [127] and (non-disulfide bond) peptide stapling [128] are not structural facets of any of the peptides investigated in this study and so are outside the scope of this review, but have been reviewed previously. Therefore, this subsection will primarily focus on the effects of lipidation, amino acid modification and disulfide bonds on peptide conformation and physical stability.

2.3.2.1. Lipidation

Lipidation is a common strategy in peptide therapeutic development with several benefits including enhancing the peptide's stability to thermal or pH variations, improving peptide cell membrane interaction strength and increasing the half-life of the drug within the body [93, 129, 130]. The functionalisation process is typically performed via an amide bond at the N-terminus or a Lys amine side chain [129, 131], often via a linker species such as L-γGlu, which has been found to be important in ensuring a prolonged half-life when lipidating Peptide YY_{3-36} [132] . However, lipidation may also be performed irreversibly via O-esterification or reversibly via S-esterification at Ser [133] or via a thiol-ester group to Cys [134, 135]. The functionalisation methods have been reviewed in more detail by Kowalczyk et al. [131] and Zhang et al. [129]. This

strategy may drastically alter the physical stability of the peptide and mechanism through which it aggregates. In the case of uncharged mildly hydrophobic peptides, lipidation increases their hydrophobicity without any secondary structure folding mechanisms and thereby increases the peptide's propensity to aggregate in aqueous solutions[129, 136]. In the case of charged peptides, lipidation leads to the formation of an amphiphilic structure [93, 137], which when above its critical micelle concentration, will associate into micelle structures, comprised of a hydrophobic core of the lipid chains bound by hydrophobic interactions and a hydrophilic peptide outer shell interacting with the solvent environment. They may themselves form a fibrillar network possibly leading to hydrogelation after sufficient time in the environment [93, 136, 138]. The lipidation site itself is important as it has the capacity to impact the aggregation propensity of the peptide with studies on a GLP-1 analogue showing that both the length of the lipid chain attached and its placement within the sequence effect the peptide's aggregation propensity [139]. Likewise, Peptide YY, when lipidated at different residues, showed differing propensities to form micelles and hydrogels under various solution conditions [130].

2.3.2.2. Non-proteinogenic amino acid substitution

Substitution of non-proteinogenic amino acids into the primary structure of the template peptide is a well-practised molecular design strategy used to improve the stability of therapeutic peptides. An example of this practice would be the substitution of Ala-8 from native GLP-1 for amino-isobutyric acid (Aib). Aib is a non-natural amino acid analogous to alanine with an additional methylation at the α-Carbon, which functions in semaglutide to protect against proteolytic degradation by dipeptidyl peptidase IV [140]. However, this strategy also impacts the peptide conformation and so the aggregation pathway. For example, Aib is known to improve the stability of α-helix conformations and disrupt the formation of β-sheet thereby inhibiting several fibrillation pathways [141]. Additionally, it was observed to improve the physical stability of glucagon when substituted with Ser-16 [142]. A similar strategy to disrupt the secondary structure of the template peptide is effected through the substitution of *D*-amino acids into the

sequence, which similarly act to inhibit the formation of secondary structures [143]. However, their effectiveness at inhibiting peptide self-assembly is dependent on the site of substitution [144]. Furthermore, it is not necessarily directly correlated to the number of *D*-amino acid substitutions but the disruption of the secondary structure formation, with alternating chiralities of peptide EAK16 showing much poorer propensity to self-assemble than either *D*- or *L*-EAK16 [145].

2.3.2.3. Disulfide Bonds

Disulfide bridges are covalent bonds formed between two Cysteine side-chain β-thiol groups, either within the same peptide chain to form a loop or with a neighbouring peptide chain [146]. These bonds are a common post-translational modification during peptide and protein folding that stabilise the conformational and monomer structures. Disulfide linkages in aggregation-prone regions of the sequence have been found to inhibit the formation of conformations leading to aggregation [147-149], but may also instigate protein oligomerisation in other cases [150-152]. The effect disulfide linkages may have on the rate and path of aggregation of the individual chains may be significant. For example, insulin has been found to exhibit a pH-dependent lag phase of fibrillation, whilst the rate of fibrillation of chain A appeared to be independent of pH [153]. Additionally, chain B has been found to display a concentration-dependent aggregation pathway under acidic conditions.

2.3.3. Peptide concentration

Peptide solution concentration typically has a significant influence on the rate of aggregation which is not unexpected as it is in accordance with most rate laws. However, the relationship is highly dependent on the mechanism of aggregation. For example, the rate of formation of amorphous aggregates of Titin I27, was observed to be a function of the aggregate surface area [154]. Conversely, as briefly described in Section 2.2.4, fibrillation, in the majority of instances, follows nucleation-polymerisation kinetics, meaning that the primary nucleation event is the rate

determining step [155]. For these kinetics, the rate of fibrillation increases and lag-time decreases with increasing peptide concentration, as exemplified by insulin [156]. In the event that secondary nucleation is the rate limiting step of fibrillation, then the relationship between concentration and rate of aggregation is still clearly observed despite requiring a more complex kinetic model [157]. Nevertheless, a negative relationship between increasing peptide concentration and rate of aggregation may be observed, as in the example case of GLP-1 at pH 7.5, although not at pH ≥7.7 [158]. This relationship is suggested to be caused by the increase in formation of off-pathway oligomers, stable under the more acidic conditions and which do not convert to on-pathway oligomers or for which the kinetics of conversion are extremely slow. Thus, their breakdown or conversion becomes a rate-limiting step of fibrillation at the sufficiently high concentrations for which they are observed.

2.3.4.pH

Peptide conformations are most stable against denaturation near or at their isoelectric points due to the reduced electrostatic interactions which can otherwise disrupt their structures [84]. However, due to the lower strength of electrostatic repulsion forces under these solution conditions, attractive intermolecular van der Waals forces, hydrogen bonds and hydrophobic interactions have their greatest effect and generally lead to an increased rate of agglomeration and aggregation of the native species. At pHs further from the isoelectric point, higher net charges result in opposing electrostatic forces that must be overcome to self-associate, leading to a slower rate of aggregation [159-161]. The pH of the peptide environment not only has an effect on the rate of aggregation but perhaps also the conformation the aggregate takes. For example, with more monodisperse and amorphous aggregates of β-lactoglobulin being observed at pHs close to or at the peptide's isoelectric point, whilst amyloid or spherulite fibrils were observed at more extreme pH values where the native monomer folding conformation is disrupted [162]. The kinetics of the formation of these aggregates and fibrils can differ substantially, resulting in a less clear relationship between pH and aggregation rate, particularly

at pHs where both species form in competition [155]. An example of a peptide with such kinetics is GLP-1, which is highlighted more in Section 5.1.2.

2.3.5. Temperature

Aggregation may occur via a number of various pathways. The four major pathways transpire via an intermediate/partially unfolded/denatured state, via self-association of the native state peptides, via chemical reactions between monomer peptides leading to chemically bound aggregates and via chemical degradation of the peptide to an intermediate with a higher propensity to aggregate [163]. Depending on the aggregation pathway most prevalent under the particular solution conditions, the rate-limiting step may display multiple different kinetic orders [163]. The primary role temperature has in peptide aggregation is its direct effect on diffusion and overcoming the activation barriers involved in either the formation of covalent bonds between monomers, chemically degrading the peptide or unfolding the peptide structure. An increase in temperature results in increased stability of the less structured monomer conformations, more prone to aggregation, since electrostatic interactions, van der Waals forces and hydrogen bonds are all exothermic in nature and so are destabilised increasing the rate of peptide unfolding, whilst hydrophobic interactions are endothermic stabilising the aggregate structure [84]. The effects of temperature are not so linear however, as the aggregation pathway with the lowest energy barrier and so conformation of the aggregation nucleus, may change with temperature, causing drastic changes in the rate of aggregation, especially, when both aggregate forms are viable and in competition [164]. Due to aggregation being a multi-step process, requiring reaction, aggregation or peptide denaturation and unfolding, the activation energy observed is the sum of those for rate-limiting steps and the enthalpy change for other aggregation steps that progress at a sufficient rate [163]. Therefore, if there is sufficient variation in the enthalpy change or activation energy of a particular step with temperature, then the process will not display Arrhenius behaviour over a wide temperature range [163]. However, Arrhenius behaviour of aggregation is exhibited for several peptides with a monomeric unfolded structure, for example

β-Amyloid peptide [165-167], or at specific conditions for which the native structure has degenerated, for example insulin at pH 1 [168] or pH 1.6 [169].

2.3.6.Excipients

2.3.6.1. Salts

There are three possible mechanisms through which salts may affect a peptide's propensity to aggregate, either via one predominantly or via a combination of them simultaneously depending on the salt in consideration, the solution conditions and the peptide. The first of these is Debye-Huckel screening, during which the salt ions coalesce non-specifically around residues within the peptide chain of opposing charge, thereby increasing the local ionic strength of the solvent and shielding, and thus reducing, peptide inter- and intramolecular electrostatic interactions [170, 171]. The second mechanism through which ions may affect peptide aggregation is via ion pairing, during which it is theorised that salt ions bind with specific polar or oppositely charged groups on the peptide chain and so alter the isoelectric point of the peptide [172]. The third mechanism is not through direct interaction with peptide itself, but how the ions affect water-water interactions and so the efficiency with which water solvates apolar residues and forms a hydration shell around it [173]. The ability of ions to impact peptide aggregation via this mechanism follows the Hofmeister series [174], an example of which is the aggregation behaviour of α-synuclein [175]. Anions and cations act separately as their charges shield different species, but their effects are cumulative [171]. Anions typically have a greater effect than cations as they are more strongly hydrated and so have a greater polarising effect [170, 176]. The mechanism through which salts impact peptide aggregation may also be dependent on the concentration of the ions. An example of this behaviour is islet amyloid polypeptide (IAPP), for which at pH 8 and low salt concentrations the rate of fibril formation depends strongly on the salt anion and scales with the electroselectivity series, suggesting the ionic binding mechanism [176]. However, this anion dependence is reduced at higher salt concentrations and follows a

Hofmeister series, whilst at pH 5.5, the ionic mechanism is maintained at higher salt concentrations, likely due to the charge of the peptide being higher at the lower pH [176]. This emphasises how the salt effects on aggregation are both pH and peptide dependent. Independent of the mechanism, the screening of intra- and intermolecular electrostatic interactions and weakening of hydrogen bonding due to increasing salt concentrations results in hydrophobic interaction becoming the more dominant forces and so the rate of aggregation increasing. Therefore, at sufficiently high salt concentrations, the structure of the monomer may be affected and so the conformation of the aggregate. For example, β_2-microglobulin forms amyloid fibrils at moderate anion concentrations, whilst at high concentrations amorphous aggregates are formed due to the clustering of exposed hydrophobic surfaces [160]. Conversely, amyloid β peptide (Aβ40) unfolds in increasing salt concentrations to form a more mature amyloid fibril conformation [177].

2.3.6.2. Surfactants

Surfactants act to stabilise peptide and protein structures in aqueous solutions from aggregation via two mechanisms: firstly by occupying hydrophobic interfaces, thus preventing protein adsorption or increased localised concentration at such sites or secondly, by interacting with hydrophobic sites of the peptide, inhibiting further interaction between peptide monomers [155, 178, 179]. Both mechanisms aid in protecting the peptide structure against interfacial stresses and the related degradation and aggregation processes that may follow. However, the role of surfactants is not entirely stabilising. Ionic surfactants may also disrupt peptide conformations making them more susceptible to aggregation by binding to the peptide surface [180]. At higher concentrations the surfactants themselves conglomerate to form micelles which in turn may aggregate to form micelle-peptide agglomerations [180]. Conversely, non-ionic micelles are suggested to stabilise peptide/protein structures against denaturation by occupying temporarily hydrophobic regions, enabling them to restructure back to their native state before interaction with another species leading to aggregation [178]. However, even non-ionic surfactants still

present drawbacks, for example polysorbates. They are hydrolysed or oxidised in solution [181], leading to the potential oxidation of the peptide structure itself or the formation of peroxide species which may act to chemically degrade the peptide [182].

2.3.6.3. Osmolytes

Osmolytes are organic molecules that fall into three primary classes: polyols including sugars, amino acids and smaller methyl ammonium species [183]. The mechanism through which osmolytes stabilise the protein and peptide folding is still not fully understood. However, the leading theory is that they are preferentially excluded from the peptide-water interface due to the peptide backbone being "osmophobic" [184]. Therefore, the unfolded state is destabilised to a greater extent than more folded or compact structures promoting folding to minimise the number of sites and surface area available for osmolyte-peptide interactions. Additionally, due to the weak/repulsive interactions between the peptide structure and the osmolyte, there is less competition for peptide-peptide interactions and so the probability of peptide intramolecular interactions resulting in folding is increased [185]. However, in the case of some smaller peptides and proteins with less organised secondary and tertiary structures, the promotion of a more compact structure, may promote aggregation to minimise peptide-osmolyte interactions [175]. In this case though, the osmolyte may inhibit fibrillation and promote the formation of smaller oligomer or amorphous aggregates [184]. Conversely, some amino acids and methylamines may bind strongly to peptide surfaces promoting peptide-peptide interactions resulting in aggregate formation, although of a potentially different nature to that in the absence of the osmolyte [184, 186]. Urea, for example acts to disrupt the hydrogen bonds within the peptide backbone promoting reversible peptide unfolding, but simultaneously inhibiting irreversible aggregation [187]. An additional theory, proposed by Politi and Harries, is that the osmolyte acts by altering the solvation environment to reduce the enthalpy loss of folding [188]. They found that this was in relation to the size of the osmolyte and not its structure, suggesting that osmolyte-peptide interactions were uncommon and not the cause [188].

2.4. Peptide Stability Measurement Techniques

Current techniques used to investigate peptide aggregation have been reviewed previously [189-191]. This Section aims to briefly discuss the most prevalent techniques before focussing on those most pertinent to this study. Each of the methods discussed henceforth characterise peptide aggregation in solution through measurement of the size of the species or analysis of its structure. Static (SLS), Dynamic (DLS) and Multi-Angle (MALS) light scattering methods are in the former of these categories. DLS is discussed in more detail in Section 2.4.1. SLS and MALS both function to estimate the size of polymeric species through measurements of the intensity of scattered light at one fixed angle and multiple angles, respectively. Both techniques have previously been applied to determine the size distribution and molecular weights of oligomers of GLP-1 analogue species in solution [192-194]. MALS is often used in combination with size exclusion chromatography (SEC), another macromolecule size analysis technique, which is discussed in greater detail in Section 2.4.2. Small angle X-ray scattering (SAXS), relative to other scattering techniques, is able to provide structural analysis of the macromolecular species being analysed. For example, SAXS has been applied to characterise semaglutide oligomers determining their radius of gyration and internal structures [195]. The primary drawback to each of these methods is that they are not predicative and rely on allowing the peptide to aggregate, thereby increasing their time-cost.

Circular Dichroism (CD) spectroscopy functions by measuring the relative difference in absorption of left- and right-circularly polarised light by a species. Therefore, in peptide aggregation analysis, it is applied to determine or observe any changes in the secondary structure of peptide. For example, GLP-1 was observed via CD to lose its α-helicity during a 9-day incubation period, with low-molecular weight oligomers formed being characterised by disordered structures and amyloid fibrils by increased β-sheet content [196]. Fourier Transform Infrared (FTIR) and Raman spectroscopy among other vibrational spectroscopy techniques are also used to investigate changes in the secondary structure of peptides which are indicative of

aggregation. For example, both techniques have been used to observe shifts in the amide I band during insulin aggregation from approximately 1650 cm^{-1} indicative of α-helical conformations to 1670 cm^{-1} representing β-sheet formation [99, 197]. Nuclear magnetic resonance (NMR) spectroscopy is another spectroscopic technique that may also be utilised in peptide aggregation studies to detect oligomeric intermediates and characterise fibril structures in solution [198]. Increases in hydrophobic residue exposure or β-sheet structural conformations during peptide fibrillation may also be detected via fluorescence assays. These either detect increases in the intrinsic fluorescence of the more exposed aromatic residues or employ a dye, the fluorescence of which is dependent on its interactions with β-sheet peptide structures. For example, fluorescence assays using Thioflavin T as the dye on liraglutide showed that the peptide fibrillated at more acidic pHs [193]. The drawback with the use of dyes in this strategy though is that their presence in solution may impact the rate of aggregation.

The methods selected for use within this work were chosen in part due to their accessibility within the Surfaces and Particle Engineering Laboratory (SPEL) at Imperial College London. Additionally, DLS has a short measurement duration, allowing for higher throughput for a wide array of conditions relative to other techniques, whilst SEC, being a high-performance liquid chromatography (HPLC) method, allows for automated data collection of samples after their initial preparation. Self-interaction chromatography (SIC) is a method that has not commonly been applied with respect to the study of peptide physical stability and as such is the primary technique of interest within this part of the thesis work. The rest of this Section will focus on the methods that were employed in this study.

2.4.1. Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) is a technique used to measure the diffusion coefficient and size distributions of particulate species in solution [199]. It operates based on the theory of Rayleigh Scattering, wherein monochromatic light is scattered by particles of a diameter smaller than the

wavelength of the incidental light, with smaller species creating more scattering. Unlike SLS, which measures the average intensity of scattered light over a period of time, DLS measures the short term fluctuations in the intensity of this light to provide a velocity distribution of these species [200]. Macromolecules in solution demonstrate Brownian motion, with smaller species moving faster than larger ones, and so creating more rapid fluctuations in intensity. Therefore, this intensity profile may be correlated to the diffusion coefficients of the species and so an estimation of their hydrodynamic radii. As a result, DLS has been commonly applied when studying protein and peptide aggregation to measure the change in size of the species. For example, the method was able to detect the formation of large aggregates of an aged liraglutide analogue at only micromolar concentrations [196].

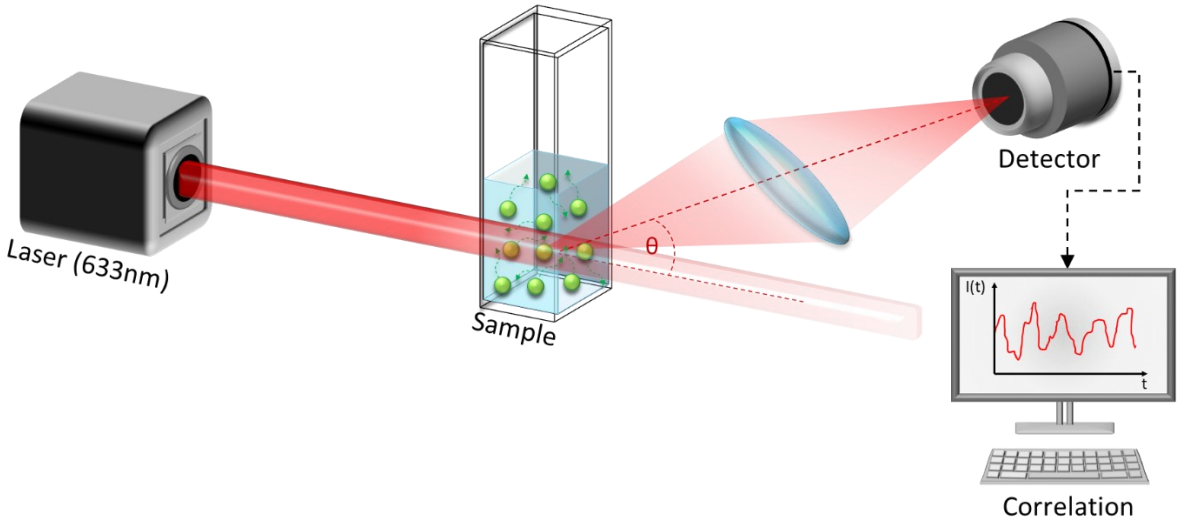


Figure 2.3: Schematic of Dynamic Light Scattering Experimental Setup

A measurement is taken by irradiating a pre-filtered, sample, measured at a specified temperature with visible monochromatic light, typically 633 nm [199]. The time-dependent variations in intensity measured by the detector may then be correlated numerically to the diffusion coefficient via the an autocorrelation function [201]:

$$G(\tau) = 1 + be^{-2D_t q^2 \tau} \qquad \textit{Equation 2.1}$$

Where b is a setup dependent constant, τ the characteristic decay time and q the scattering vector given by the equation below [201]:

$$|q| = \frac{4\pi n_0}{\lambda_0 sin(\theta/_2)} \qquad \textit{Equation 2.2}$$

where n_0 is the refractive index of the solvent, λ_0 the wavelength of monochromatic light (typically 633 nm) and θ the scattering angle. As the peptides are dimensionally over an order of magnitude smaller than λ_0, there is not expected to be an angular dependence of the scattered intensity, so measurements at a single angle should suffice [199]. Finally, the diffusion coefficient is correlated to the hydrodynamic radius, R_h, by assuming the species is a hard sphere and using the Stokes-Einstein relationship [202]:

$$D_t = \frac{k_B T}{6\pi\eta R_h} \qquad \textit{Equation 2.3}$$

where k_B is Boltzmann's constant, T is the temperature and η the dynamic viscosity of the solvent. The primary drawbacks of DLS measurements include the inability to resolve the spectra of particles differing in size by a factor less than five [203]; the susceptibility of the size determinations to be impacted by the presence of large contaminant dust particles, thereby requiring extra care during sample preparation; the assumption of spherical particles, and its poor accuracy for species of sizes <1 nm in diameter, the size range of some smaller peptide monomers [204].

The primary benefit of DLS is that it is non-invasive technique, allowing for the sample to be recovered for further experimentation, which does not require an expert user. Additionally, since large aggregates cause significant scattering, it is able to detect their presence even at low protein concentrations and estimate their size with an uncertainty less than 10% for monodisperse samples [205]. It is also poor at resolving two similarly sized species (requiring a minimum difference in species size of approximately a factor of five) [203].

2.4.2.Size Exclusion Chromatography (SEC)

Size-exclusion chromatography functions by separating biomolecules according to their hydrodynamic size. It employs a non-interacting chromatographic stationary phase of particles with a well-defined range of pore sizes and a suitable aqueous (non-denaturing) mobile phase. Smaller species in the sample analyte may access deeper into the porous volume and so have a longer mean path and so retention time through the column relative to larger, including aggregated, species [206] (see Figure 2.4). Therefore, SEC may provide good separation of different sized species requiring only minimum volume/amount of sample.

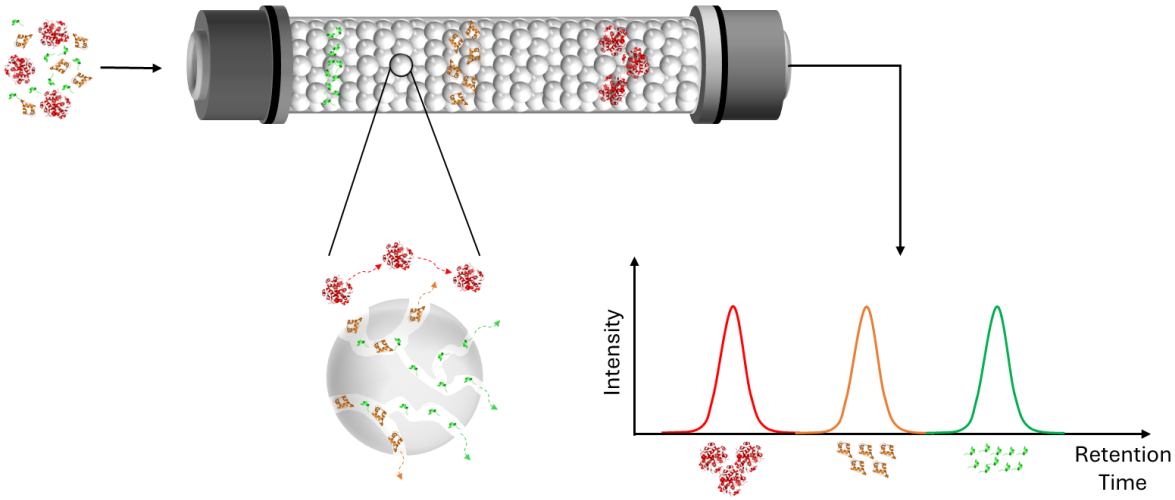


Figure 2.4: Representative schematic of the process of separation in Size Exclusion Chromatography

The separation process is therefore assumed to be entropically controlled based on the size differences of the species, and so follows a Gibbs equation [206]:

$$\ln K_D = -\frac{\Delta S_0}{R} \qquad \textit{Equation 2.4}$$

where K_D is a thermodynamic retention factor, which represents the fraction of the pore volume accessible to the species. SEC columns are calibrated using species of known molecular weight to form a third order polynomial between log M_w and the retention volume with a linear region of highest resolution [207]. However, calibration curves can contain inaccuracies due to the non-exact correlation between the molecular weight and the Stokes radius caused by the non-

spherical nature of peptide conformations [206]. Additionally, whilst the resin is assumed to be non-interacting, weak adsorption of the peptide analyte results in longer retention times and so less accurate molecular weight estimations can be observed [208]. For high resolution determinations between oligomers and monomers, smaller particles with narrower pore distributions or longer columns are required [206]. However, this change will result in higher column pressures and increased shear forces which may damage sensitive peptide structures and conformations [209]. This creates a difficultly with fibrillating peptides where the aggregate size may be several orders of magnitude larger than the monomers and oligomers, and so either resolution is lost between the monomers and oligomers or there is risk of blocking the column [210]. Therefore, size-exclusion chromatography has been most commonly applied in the study of peptide oligomerisation rather than aggregation or fibrillation. For example, the method was applied in the study of the physical stability of a GLP-1 amide, accurately identifying low molecular weight oligomers and isolating them from its monomers [211]. Furthermore, size-exclusion chromatography was also useful in the identification of a pH-induced switch in liraglutide oligomer formation, being able to separate and resolve the 13-mer adopted at pH <7 and the 7-mer adopted at pH >7 [193].

2.4.3. Self-interaction Chromatography (SIC)

Self-interaction chromatography (SIC) is an experimental chromatographic method utilising a stationary phase, with the investigated peptide being covalently immobilised onto its surface. The same peptide is injected as the analyte into a mobile phase of interest and eluted isocratically (see Figure 2.5). The measured retention volume, as detected by a UV-Vis detector, is a chromatographic measurement of the strength of the interaction between both the mobile phase and immobilised peptide species. Longer retention volumes indicate more attractive peptide-peptide interactions. Therefore, if it is assumed that the peptide has been immobilised with a random orientation or at least maintains a high level of flexibility to orientate itself on the

surface, the retention volumes are indicative of the overall average peptide-peptide interaction strength.

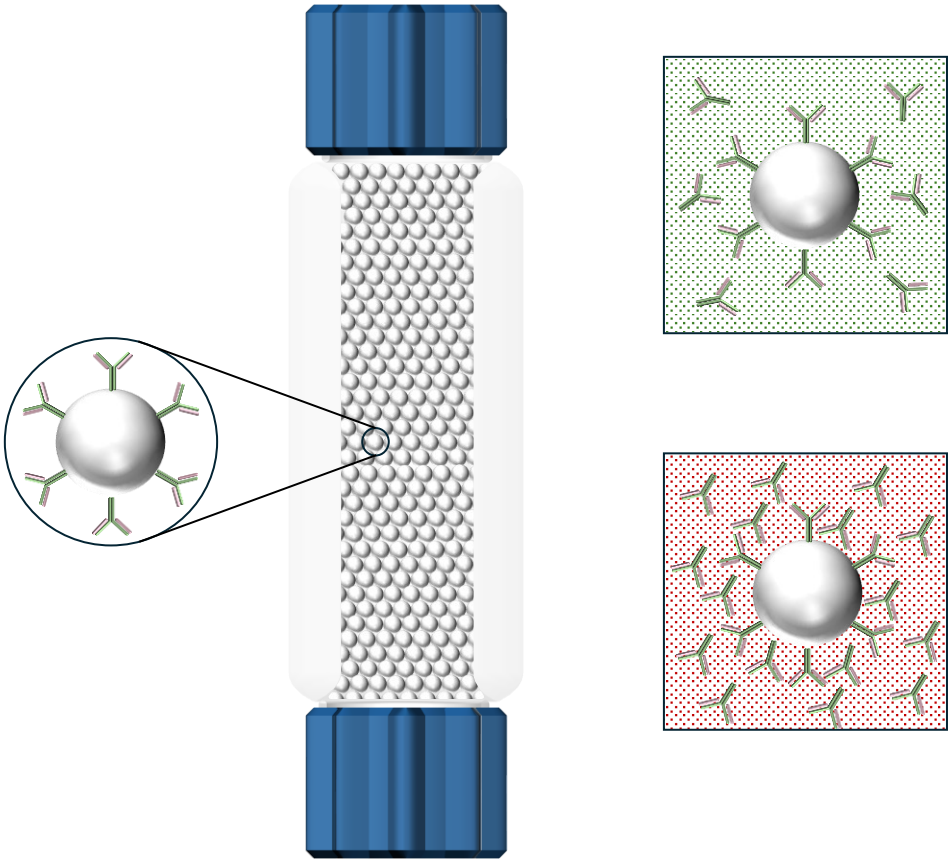


Figure 2.5: Schematic of idealised self-interaction chromatography packed column setup
Inserts: Top/Green and bottom/red images represent weak and strong monomer-monomer interactions respectively

Figure 2.5 shows an illustration of an idealised SIC setup for which all sites of the immobilised species are available for interaction and the mobile phase species is in monomer form. For peptides, the immobilised species may be conjugated by multiple sites or by only one meaning that the number of possible interactions with mobile phase species may be limited. Furthermore, the mobile phase species may also be in an oligomer or aggregate form, limiting interactions with the immobilised species further. Tessier et al. developed a statistical model including the Van der Waals volume of exclusion, the phase ratio and the immobilised protein surface density to

correlate self-interaction chromatography retention volumes to the second osmotic virial coefficient of the analyte (B_{22}) [212]. The full relationship and calculation procedure is outlined in Section 5.2.5.1.5. Tessier et al. also validated this method by correlating their SIC B_{22} values with those collected via SLS. However, the method requires an accurate estimation of the resin surface area per unit volume. Therefore, whilst antibodies have been successfully immobilised onto various resins [213], to reduce the necessary repeated characterisation of the stationary phase resin, those previously characterised by DePhillips et al. are often used by researchers [214].

The relationship between the 2nd virial coefficient of the van der Waals equation (B) for a non-ideal solution and protein solubility has already been well theorised, modelled and documented [215, 216]. In McMillan-Meyer solution theory, the 2nd virial coefficient describes the net force from the sum of all interactions (hydrophobic, electrostatic and exclusion volumes) between two protein molecules in a dilute solution [217]. Therefore, it can provide a semi-quantitative measurement of the strength of the interactions between peptide molecules relative to the strength of interactions with solvent molecules, and so can be used as a measure of their stability in solution [217]. Highly negative values indicate strongly attractive forces leading to rapid aggregation or precipitation. Slightly negative values represent weak attraction forces and so mild instability. Finally, positive values signify net repulsion between peptides, thereby implying the monomeric conformation is stable in solution [217]. This relationship between B_{22} and aggregation was further established in the only self-interaction chromatography study of a therapeutic peptide to date on enfuvirtide, for which the osmotic second virial coefficient values measured correlated well with molecular weight, which in turn represents the level of association [218].

A drawback of SIC is the lack of clarity for negative B_{22} values close to 0, which suggest potentially slow aggregation. This uncertainty is due to the approximation made through immobilisation, as the calculation of B_{22} assumes that all peptide/protein orientations are available for analyte

interactions, which is not the case for the covalently bound species, and is less likely in the case of peptides with fewer immobilisation sites [212]. Additionally, there is the risk that the column ceases to function as designed either through denaturation or degradation of the immobilised species or contamination of the stationary phase. Nevertheless, there are several benefits to SIC relative to other B_{22} measurement techniques. For example, since the immobilised protein columns may be used for several measurements over the course of an extended period of time, the consumption of the sample protein or peptide is significantly reduced relative to static light scattering measurements [219]. Furthermore, since the technique is a chromatographic method, it may be highly automated and so relatively high-throughput. Finally, it is also less sensitive to impurities relative to light-scattering techniques and so requires less stringent sample handling and preparation.

CHAPTER 3

3. GASTROINTESTINAL PROTEOLYSIS & CONDITIONS: A REVIEW

3.1. Structure of the Gastrointestinal Tract

The gastrointestinal tract (GIT) is a smooth muscle-lined, hollow tube extending from mouth to anus functioning to ingest, digest, absorb and excrete dietary nutrients and waste (as illustrated in Figure 3.1). Its structure is primarily divided into six sections: the mouth, for the mechanical initiation of food digestion; the oesophagus, for the transport of ingested material from the mouth to the stomach; the stomach, for the acid- and enzyme-catalysed degradation of nutrients; the small intestine (often referred to as the small bowel), for the further proteolytic and bile-assisted degradation of dietary material and the primary absorption site of nutrients; the large intestine (often referred to as the colon), for the absorption of water and electrolytes and the storage of faeces, and finally the anus, from which waste material is expelled. The oral administration of drugs targets their absorption into the systemic circulation via the gastrointestinal mucosal lining. For peptide drugs, the typical target absorption site for trialled oral formulations is the small intestinal epithelium, such as for example Cyclosporine A [220]. However, absorption via the gastric mucosa has been observed to be weakly effective for the oral delivery of semaglutide [221]. Therefore, design of the formulation strategy is dependent on avoiding or mitigating the effects of the harsh conditions of the target absorption site within the GIT. As such, for experimental techniques aiming to mimic GIT conditions, a detailed understanding of their components and the effects of the relative nutritional status (i.e. fasted and fed state) is required to develop simulated gastric (SGFs) or intestinal (SIFs) fluids. Highly detailed reviews discussing these factors have been carried out by Mudie et al. [39] and Fuchs and Dressman [222]. The

remainder of this subsection will outline briefly a review of the literature discussing the characteristics of the conditions in the stomach and the small intestine. Illnesses affecting GIT conditions such as gastric cancer [223], ulcerative colitis [224] and Crohn's disease [225] are known but are not discussed here as they are outside the scope of the work in this thesis.

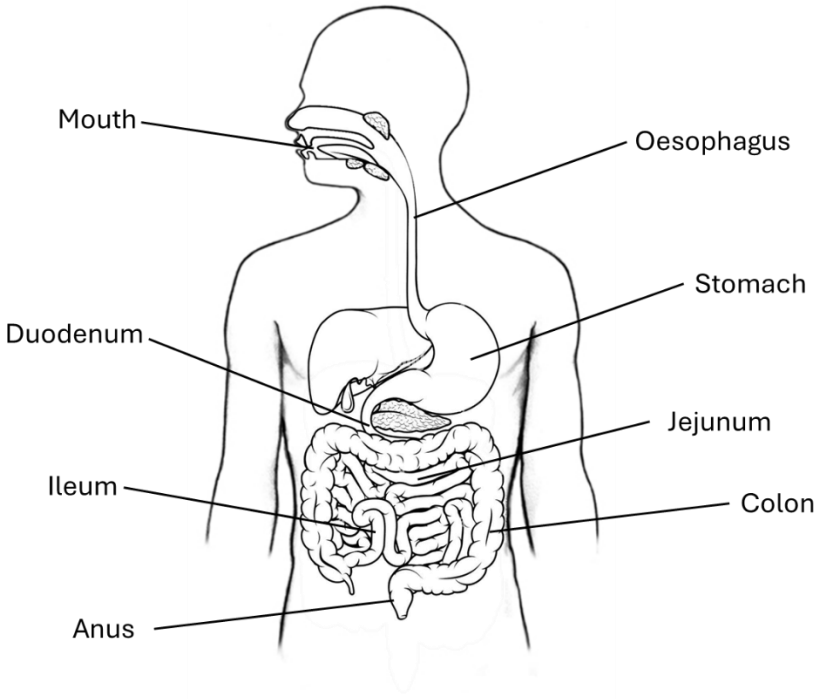


Figure 3.1: Illustration of the human gastrointestinal tract with labelled relevant sections

3.1.1. Gastric Conditions

The stomach acts as the initial major site of retention within the GIT. It is not the primary delivery site for oral drug absorption partially on account of its harsh conditions but principally due to the thickness of the mucosa lining and the low surface area of approximately only 0.1 m^2, resulting from the lack of any specialised absorption structures such as microvilli [226]. Under fasted conditions, it typically has a volume of less than 50 mL [227], but this may increase to over 500 mL with the consumption of a large meal [228]. In fact, gastric conditions, which are known to be highly acidic, vary significantly depending on the relative nutritional status of the individual. For example, a quantitative meta-analysis performed by Abuhelwa et al. found that the average

gastric pH was 1.88 ± 0.16 (1.56-2.20, 95% CI) increasing to pH 4.98 ± 0.16 (4.66-5.31, 95% CI) under fed-state conditions [229]. Furthermore, the fed state also affects the gastric emptying time and so the available absorption time of any potential oral drugs targeting gastric absorption. As described in a different quantitative meta-analysis of 28 studies by Abuhelwa et al., the emptying time of the stomach is estimated to be 1.37 h ± 0.17 h when fasted, but this may range drastically depending on the level of food consumption from 1.68 h ± 0.17 h after the consumption of 100 cal to 5.94 h ± 0.32 h after the consumption of over 1000 cal [230]. Furthermore, these variations differ for different groups, for example the rate of gastric emptying of solids has been observed to be significantly longer in women than men [231]. Therefore, typically values for the concentrations of species in gastric juice are only reported for fasted state conditions due to the high variation with the diet of the individual. The most prominent of these being chloride and bile acids, being measured to be 176.3 mM ± 30.0 mM and in the range 0-0.62 mM, respectively, at the fasted state but 35 minutes after drinking water [232]. Furthermore, a significant increase in buffer capacity and reduction in surface tension of human gastric fluid has been observed postprandially, whilst pepsin concentration doubled [233].

3.1.2. Intestinal Conditions

The small intestine is comprised of three major subsections with different functions. The first of these subsections is the duodenum, approximately 30 cm in length [226], which has three primary functions: release of bicarbonate-rich pancreatic fluids into the GIT to neutralise the acidic chyme from the stomach; further digestion of lipids, carbohydrates and proteins aided by pancreatic enzymes and liver-secreted bile, and thirdly, the initialisation of nutrient absorption through its epithelial surface. However, due to its low absorption area of only approximately 0.1 m^2, it is evident that absorption is not its primary function [226]. The jejunum is the following section of the small intestine. It is approximately 3 m in length and has a large surface area of approximately 60 m^2 for the absorption of nutrients, owing to the microvilli structure adopted by its epithelial surface, which is primarily comprised of densely-packed enterocytes [226]. The final

section of the small intestine is the ileum, the primary function of which, similar to the jejunum, is the absorption of nutrients, although primarily bile and vitamin B12 [234]. The ileum is characterised by a thinner mucosal wall than the jejunum, with a similar surface area of approximately 60 m^2 but being slightly longer at roughly 4 m [226]. Together they have an overall transit time of approximately 3.49 h ± 0.17 h according to Abuhelwa et al. [230], although there is some degree of variance in the literature with Bennink et al. reporting roughly 5 h [231].

Through the use of pH-sensitive capsules, pH profiles of the GIT have previously been recorded, showing a general increase in pH along the small intestine [38, 235]. The duodenum, due to its proximity to the stomach and site of pancreatic ducts for the release of bile salts, has the lowest pH range from 5.5-6.5 under fasted conditions, which increases to pH 6.1-7.1 in the jejunum and to pH 7-8 in the ileum [226]. However, whilst the jejunum and ileum pH profiles remain relatively unchanged by the relative nutritional status of the individual, the pH of the duodenum drops to pH 4.9-6.0, due to the increased release of gastric juices [226]. Predominantly, the concentrations of species within the small intestines are not divided by subsection, therefore there is often relatively high variance in the data reported. However, chloride and sodium concentrations of approximately 120 mM and 100-150 mM are often reported and do not vary significantly with fasted and fed state [39, 232]. Conversely the lipid content of the small intestine has been found to correlate with the nutritional status of the individual being observed to be as low as approximately 0.66 mg/mL to as high as 10.96 mg/mL on average in fasted and fed states, respectively [236].

3.1.3. Simulated Gastrointestinal Fluids

The direct measurement of human GI fluid is a cumbersome process and provides inconsistent solubility data, dependent on the aspiration protocol [237]. As an alternative, simulated gastrointestinal fluids (SGIFs) aim to mimic the gastrointestinal conditions during *in-vitro* experimentation. Therefore, it would be assumed that the more detailed and accurate the SGIF,

the better representation of the GI environment it would. However, a complex mixture may not be required to evaluate specific properties of the peptide, and in fact, a more simple SGIF may simplify and so clarify analysis. However, this will be sample- and property-dependent, for example, it has been demonstrated that the solubility of poorly soluble species in human intestinal fluid was well-predicted by *in-vitro* experiments using complex SIFs, whilst buffer solutions underestimated their solubility [238]. Surface tension was considered to be the most important parameter to match to GI conditions when designing a SGIF for dissolution studies [238]. Conversely, the solubility of cyclosporin A was found not to vary significantly between in water, at pH 1.2 and at pH 6.6, whilst being inversely correlated with temperature [239].

Markopoulos et al. recommend four tiers of increasing biorelevance (Level 0 – Level III) for application in formulation development [240]. Level 0 media control only for pH; Level I adjusts for buffering capacity and so allows for initial distinction between fasted and fed states; Level II additionally accounts for GI lipids, bile components and osmolytes; finally, Level III introduces more complexity to Level II, including dietary proteins and digestive enzymes [240]. Jantratid et al. outline the development process of detailed fasted and fed state simulated gastric and intestinal fluids, including early, middle and late-stage digestion in the proximal intestines, which were stable for 72 h under agitated conditions [241]. Fasted and fed state SIFs have been applied to study the effects of the concentration of various bile salts on octreotide and desmopressin, two orally-administered, cyclic peptide therapeutics across an artificial membrane [242]. The bile salts were found to reduce permeability but increase stability against pancreatin proteolytic degradation.

3.2. Gastrointestinal Enzymes

Within the gastrointestinal environment there are also a number of enzymes that catalyse and facilitate the digestion and absorption of dietary nutrients. Of these, the digestive hydrolases, which break down complex nutrients into simple molecules, may be divided into four categories:

glycosidases, which digest carbohydrates into simple sugars; lipases, which digest fats into fatty acids and glycerol; nucleases, which digest nucleic acids into nucleotides, and finally proteases, which break down proteins and large peptides into amino acids and simple peptides [40]. For the oral delivery of peptide therapeutics, the primary enzymatic barrier in the GIT is unsurprisingly the proteolytic enzymes. Proteases may be further divided into two categories by their mechanism of action: endopeptidases, which act to cleave internal peptide bonds, and exopeptidases, which cleave amide bonds at either end of the primary sequence of the protein to release amino acids or dipeptides [243]. A list of the primary proteolytic enzymes present in the gastrointestinal tract and their specificities are outlined in Table 3.1, and have been reviewed thoroughly by Pereira de Sousa et al. [40] and Bernkop-Schnurch [244]. Whilst all these proteases are present within the GIT, their function is not necessarily for dietary protein digestion. For example, the primary function of enteropeptidase (otherwise known as enterokinase) is for the activation of trypsinogen to trypsin [245]. Nevertheless, the presence of these enzymes should be considered in the case of oral peptide therapeutics with primary sequences analogous to their substrate recognition sites.

Table 3.1: Gastrointestinal proteolytic enzymes and details of their secretion site and specificity *[243, 246, 247]*

Secretion Site	Peptidase Classification	Enzyme Class	Specificity	Cleavage Site
Gastric Chief Cells	Endopeptidase	Pepsin	Phe, Tyr, Trp, Leu	Carboxyl Side (Preferentially)
Pancreas	Endopeptidase	Trypsin	Arg, Lys	Carboxyl Side
		Chymotrypsin	Tyr, Phe	Carboxyl Side
		Elastase	Ala, Gly, Ile, Leu, Val	Carboxyl Side
	C-terminal Amino Acid Exopeptidase	Carboxypeptidase A	Tyr, Phe, Ile, Thr, Glu, His, Ala	Amino Side
		Carboxypeptidase B	Arg, Lys	Amino Side
Brush-Border Membrane	Endopeptidase	Endopeptidase 24.11	Hydrophobic Amino Acids	Amino Side
		Endopeptidase 24.18	Aromatic Amino Acids	Carboxyl Side (Preferentially)
		Enteropeptidase	$(Asp)_4$-Lys	Carboxyl Side
	C-terminal Amino Acid Exopeptidase	Carboxypeptidase M	Arg, Lys	Amino Side
		Carboxypeptidase P	Pro, Gly, Ala	Amino Side
		γ-Glu Carboxypeptidase	γ-Glu	Amino Side
	N-terminal Amino Acid Exopeptidase	Aminopeptidase A	Asp, Glu	Carboxyl Side
		Aminopeptidase G	Ala, Leu, Met	Carboxyl Side
		Aminopeptidase N	Ala, Leu	Carboxyl Side
		Aminopeptidase P	Pro	Carboxyl Side
		Aminopeptidase W	Trp, Tyr, Phe	Carboxyl Side
	C-terminal Dipeptide Exopeptidase	Peptidyl-peptidase A	His-Leu	Amino Side
	N-terminal Dipeptide Exopeptidase	Dipeptidyl Peptidase IV	Pro, Ala	Carboxyl Side

3.2.1. Protease Mechanism of Action

The irreversible proteolytic hydrolysis of the peptide bond of a single protein substrate has been observed to follow the Michaelis-Menten mechanism, in which the substrate binds to the enzyme reversibly before hydrolysis of the peptide bond occurs [248]:

$$A + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EA \xrightarrow{k_2} E + P$$

This leads to the realisation of the Michaelis-Menten kinetic rate law:

$$\frac{d[P]}{dt} = k_2[EA] = \frac{V_{max}[A]}{K_m + [A]} \qquad \textit{Equation 3.1}$$

where K_m is the Michaelis constant, which indicates the substrate affinity, and V_{max} is the maximum turnover rate of the enzyme (i.e. when all enzyme species participate in the reaction). They are both calculated as follows:

$$K_m = \frac{k_{-1} + k_2}{k_1} \qquad \textit{Equation 3.2}$$

$$V_{max} = k_2[E]_0 \qquad \textit{Equation 3.3}$$

Therefore, the rate of hydrolysis with respect to substrate concentration shows two predominant relationships. The first is an apparent 0th order rate law when the substrate is in excess, and the rate of proteolysis is dictated by the concentration of the enzyme as the limiting reagent. As such, the rate law is approximately that of V_{max}. The second is an apparent 1st order rate of reaction, when the concentration of the substrate is the limiting factor and as such K_m becomes the dominant term of the denominator of the rate law.

3.2.2. Condition Dependence of Enzyme Activity

Since enzymes are proteins, their charge and conformation, and so their substrate affinity and activity, are highly dependent on the solution conditions they are exposed to. The most studied of these variables are pH and temperature, which will be outlined briefly here. For a comprehensive review of the environmental factors affecting enzyme activity the books published by Bisswanger and Buchholz et al. are recommended [249, 250].

The activity of most enzymes with relation to pH follows a bell-shaped relationship. Typically, this maximum activity corresponds to a physiological pH decreasing to zero at more extreme values [249]. However, there are notable exceptions such as pepsin, which exhibits peak activity under highly acidic conditions. This behaviour is caused by two simultaneously occurring effects

with changing pH. Firstly, variations in the relative acidity of the environment result in changes in the protonation state of the active site or specific ionic groups involved in the catalytic mechanism, as are common with acid-base catalytic enzymes [249]. The result of this effect is a change in k_2, which may be mathematically related to the $[H^+]$ via the pK_a of the active site [250]. For example, under more acidic conditions (pH <5.5) the imidazole group of trypsin His-57 will be predominantly charged and so will not stabilise the proton transfer from the Ser-195 group, which is the initial step in the cleavage of the amide bond by the serine-protease. The change in activity due to change in protonation state of the active site is typically reversible, since the overall conformation and substrate affinity is not greatly impacted [250]. The second effect of changing pH causing the Bell-shaped curve relationship with activity is the change in charge state of residues that stabilise the conformation of the enzyme [249]. By contrast, these changes may be irreversible as, beyond a certain charge, changes in secondary and tertiary structure result in permanent loss of substrate affinity and potentially lead to aggregation, as described in Section 2.3.4. Therefore, with this loss in substrate binding, a significant decrease in k_1 and so large increase in K_m is observed and hence there is a decrease in rate of proteolysis [250].

The relationship between enzyme activity and temperature also follows a bell-shaped curve due to similar competing factors (as illustrated in Figure 3.2). Initial increases in temperatures result in higher enzyme activity with k_2 increasing in accordance with Arrhenius kinetics [249]. However, this may also be accompanied by a decrease in selectivity of the site of proteolytic action [250]. Furthermore, the maximum of the bell-shaped curved indicates the temperature at which irreversible denaturation of the enzyme initiates which typically occurs in the 40-50 °C range [249]. Enzyme thermal denaturation is a complex multi-step process, however, loss of enzyme activity due to thermal inactivation has been observed to follow first order kinetics, indicating that a single step is responsible for activity loss [249]. However, since physiological temperatures have been observed to remain relatively consistent throughout the gastrointestinal tract (between 34 °C and 38 °C), this is not a relevant factor to in-vivo experimental work [38].

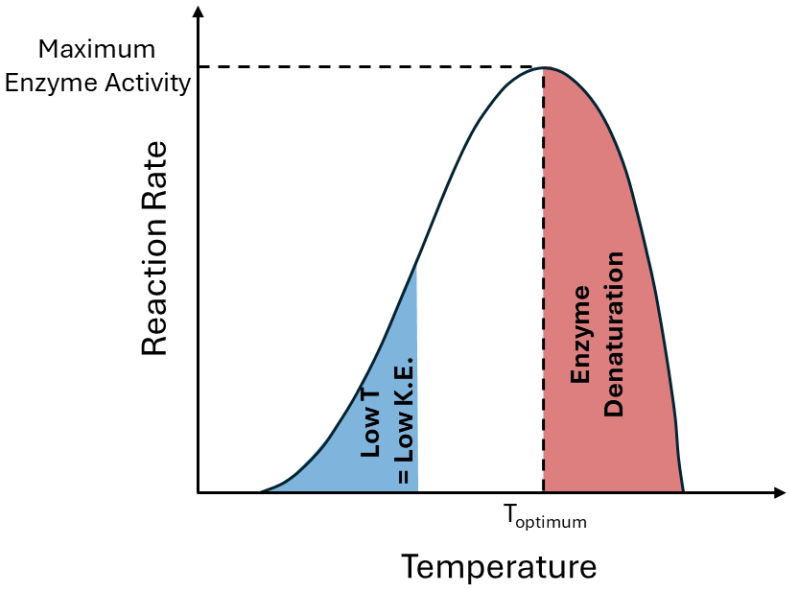


Figure 3.2: Illustration of the relationship between Proteolytic activity and Temperature

3.2.3.Enzyme Inhibitors

In the field of oral protein or peptide therapeutic delivery, a strategy that has been attempted to circumvent proteolytic degradation within the GIT, is the use of enzyme inhibitors within the formulation. Enzyme inhibitors are compounds that directly or indirectly act to alter the structure of the enzyme or block the active site, thereby inhibiting the rate of reaction [250]. Those which reduce enzyme activity by binding directly to the enzyme structure may be classified as irreversible or reversible, based on whether they bind covalently or not. The latter of these groups may be further divided into three groups: competitive, which bind at the active site thereby blocking it; non-competitive (allosteric), which bind to an alternative site on the enzyme that causes a conformational change and reduces activity, and thirdly uncompetitive, which bind to the enzyme-substrate complex, thereby preventing the release of the product [251]. The effects of these species on enzyme activity and their associated rate laws have been thoroughly reviewed by Bisswanger [249] and Buccholz et al. [250]. Indirect inhibitors are commonly species such as chelating agents, which bind to metal ions in solution that are the co-factors for several enzymes, or pH-modulators [252]. Various inhibitors which have previously been co-formulated with peptides and proteins, such as aprotinin, soybean trypsin inhibitor and

Ethylenediaminetetraacetic acid (EDTA), along with their mechanisms have been reviewed by Woodley [253], Werle et al. [252] and Bernkop-Schnürch [244]. The drawback to co-formulating peptide therapeutics with enzyme inhibitors are the potential toxic side effects associated with their long-term use such as mucosal damage, poor digestion of dietary proteins and an enlarged pancreas [244]. However, relative to alternative formulation strategies and technologies, they are cheap and scalable and additionally have previous approval by regulatory bodies [252]. As a result, co-formulation of enzyme inhibitors has been observed in several oral peptide formulation strategies, for example the use of soybean trypsin inhibitor in ORMD-0801, an experimental insulin analogue developed by Oramed Pharmaceuticals targeting oral delivery [254].

3.2.4. Effect of peptide structure

As was discussed previously with respect to aggregation in Section 2.3.1, the structure and primary sequence of the peptide therapeutic may stabilise it greatly from proteolytic degradation. In a study subjecting 17 peptide drugs to simulated gastrointestinal fluids, their stability to proteolytic degradation was found to vary significantly depending on the molecular weight of the peptide, its amino acid sequence and the rigidity of the conformation [41]. Therefore, peptide modification strategies, which maintain its pharmacological effect but stabilise the peptide against enzyme hydrolysis, are commonly implemented and have been thoroughly reviewed [255]. However, only three pertain to the work carried out in this thesis: amino acid substitution, cyclisation and lipidation.

Amino acid substitution, whilst in theory the simplest stabilising strategy, has the most variety. For example, substitution with proline is known to stabilise the primary sequence against proteolytic attack, but this is primarily due to its ability to inhibit helix formation and so alter the enzyme-binding site [256]. A more common and targeted approach is the substitution of a specific amino acid at the target cleavage site to prevent the action of a particular enzyme via steric hinderance, minor conformational changes or altered charge states. An example of this is

the substitution of Ala-8 in the GLP-1 structure for amino-isobutyric acid (Aib) in semaglutide to prevent the action of dipeptidyl peptidase IV [140]. An alternative to substituting with a different amino acid is to substitute the *D*-amino acid enantiomer, thereby increasing the rigidity of the structure and altering its secondary structure, thereby reducing their binding affinity [257]. However, this technique when taken to the extreme results in a significant decrease in functionality of the peptide or recognition at its target site. Alternative substitutions typically result in the formation of peptidomimetics. These include substitution with aza-amino acids, which are amino acids for which the α-carbon has been substituted with nitrogen, which result in minimal overall conformational changes and changes in the electronic and steric properties at the peptide bond, thereby inhibiting its hydrolysis [258].

Cyclisation of the primary structure of the peptide also increases its proteolytic resistance through a reduction in its conformational flexibility and so its ability to bind to the active site of the enzyme [259]. Backbone cyclisation of a linear peptide sequence has the additional benefit of removing the C- and N-terminal binding sites of the peptide, which are susceptible to hydrolysis via exopeptidases. However, the increased rigidity provided by disulfide bond cyclisation has also been shown to provide proteolytic resistance [41]. Lipidation of the peptide structure has also been observed to increase peptide proteolytic resistance [260, 261], although this is suggested to be primarily on account of increased steric hinderance or the formation of more complex conformations inhibiting enzyme binding.

3.3. Proteolytic Activity Measurement Techniques

For the various classifications of enzymes there are a range of different techniques which have been summarised by Copeland [262] and Bisswanger [263]. However, in the case of peptide proteolysis studies, real-time monitoring techniques such as colorimetry, fluorimetry or electrochemical methods such as polarography are not typically applied since proteolysis does not generally liberate coloured, fluorescent or charged species. Whilst the substrate may be

labelled with fluorescent dyes to enable real-time monitoring of the progress of the reaction [264], this functionalisation would change the structure of the substrate peptide and so may affect the overall kinetics of the reaction. Therefore, the most common technique involves timepoint sampling of the reaction mixture followed by immediate inactivation of the enzyme either via precipitation, typically using trichloroacetic acid (TCA) [265] or ice-cold organic solvents [259], or a change in pH via the addition of acid/base [41]. The samples are then typically centrifuged before the supernatant is then analysed, often via chromatographic methods, including measuring the concentration change of the substrate via area under the curve methods [259] or mass spectrometry [266].

3.3.1. Immobilised Enzymes

Enzyme immobilisation is defined as the confinement of enzyme species to a fixed location [267]. The principle behind the immobilisation of enzymes is to achieve the same benefits as those of heterogeneous catalysts: a simplified post-reaction separation procedure reducing enzyme loss and improved potential applicability in a continuous process [250]. As a result, immobilisation reduces costs and allows for better reaction control. However, it also results in the improved stability and so an increased storage time of 3-6 months [250] of the enzyme whilst additionally allowing for the co-immobilisation of multiple enzymes to enable multistep reactions without autolysis [268]. Furthermore, when implemented into an *in-situ* digestion reactor setup coupled with reversed-phase high-performance liquid chromatography (RP-HPLC) and mass spectrometry, a trypsin-functionalised mesoporous silicate nanoreactor allowed for the detection of proteins at concentrations as low as 2 ng/μL [269]. Additionally, increased enzymatic activity and substrate selectivity have also been observed [270]. However, this increase in activity is not typical as the increased rigidity and reduced mobility of the enzyme is generally accompanied by a reduction in catalytic activity, although this is partially on account of mass transfer and diffusion limitations associated with heterogeneous catalysts [267]. Immobilisation may be performed via one of four distinct methods: adsorption, covalent-

bonding, entrapment or encapsulation and cross-linking. Each of these methods have their benefits and their drawbacks which will be briefly summarised here, but for further information see the reviews by Jiang et al. [271] and Mohidem et al. [268].

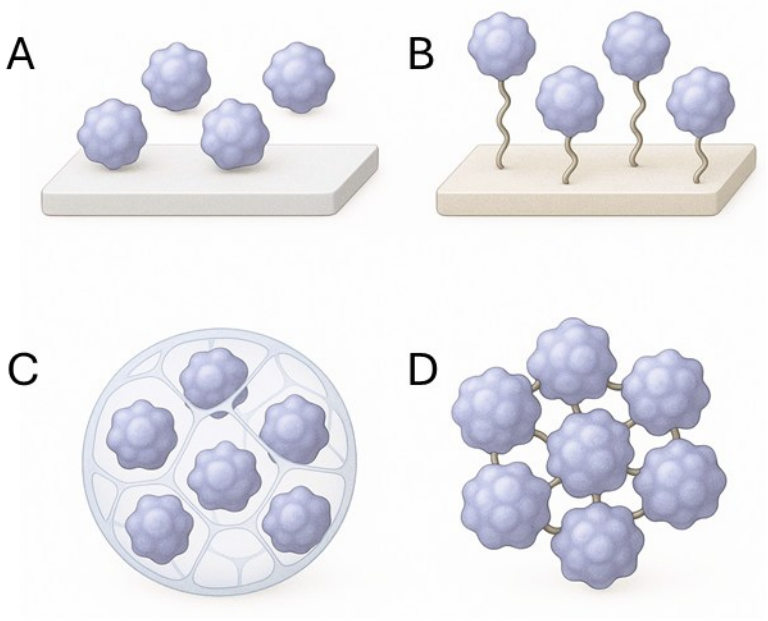


Figure 3.3: Illustration of the four primary methods of enzyme immobilisation
(A) Physisorption, (B) Covalent Bonding, (C) Entrapment or Encapsulation and (D) Cross-linking Enzyme Aggregates (Generated by Microsoft CoPilot [272])

3.3.1.1. Adsorption immobilisation

Adsorption immobilisation of enzymes utilises physical interactions such as van der Waals forces, hydrogen bonding and weak ionic forces to bind the enzymes weakly to a carrier or support surface. Therefore, the immobilisation procedure is relatively quick and cheap [273]. Despite adsorption immobilisation being weak relative to alternative methods, the primary advantage of this technique is that the weak interactions result in minimal conformational changes to the enzyme structure and minimal impact on its native activity [274]. Additionally, the method is highly flexible as various carriers may be used depending on the target enzyme to be immobilised and the conditions under which the final catalysis is to be taken. Therefore, the physicochemical properties such as the polarity and charge of the enzyme dictate the method of adsorption and the required surface properties of the support. Furthermore, in the event the

desired support is not suitable for physisorption of the investigated enzyme, surface modifying agents such as glutaraldehyde may be utilised to improve adsorption [275]. The typical support materials may be divided into inorganic and organic materials. Of the former of these, silicates are the most common since they are cheap, highly porous, easy to produce and are highly stable, although metal carriers such as titanium and aluminium have also been utilised [273]. The support materials may be further optimised. For examples, mesoporous silicates may also be synthesised with tailored pore sizes optimised for biomolecule adsorption to minimise steric hinderance of the substrate and reduce the rate of leaching from the resin surface [276]. For example, chloroperoxidase immobilised onto a periodic mesoporous organosilane, PMO- PA-40, exhibited little loss in activity after 20 cycles [277]. Organic supports such chitin or cellulose have higher biocompatibility but have a narrower operating window due to their reduced durability and their increased selectivity. Nevertheless, due to the intrinsically weak interaction strengths, enzymes immobilised via adsorption may leach enzyme from the surface over time and as such may operate under a narrower range of conditions to prevent breakdown of their interactions and rapid leaching [278].

3.3.1.2. Covalent immobilisation

Covalent immobilisation of enzymes is similar to adsorption and requires a support surface onto which the enzyme is fixated via covalent bonds. The immobilisation procedure utilises surface-available with reactive amino-acid R-groups in the enzyme primary sequence to react with functional groups on the support surface. Nucleophiles such as cysteine thiols and amines from lysine, arginine (guanidino groups) and the N-termini are generally the most reactive groups endogenous to the enzyme structure. Conversely, carboxylic acid R-groups required activation with species such as carbodiimides to enable their covalent fixation to simple surface chemistry groups [279]. The various different covalent immobilisation strategies and carrier activation procedures have been thoroughly reviewed by Robescu & Bavaro [280]. Hydroxy, amino and epoxy surface functional groups are typical surfaces for covalent immobilisation, although the

latter two require pre-activation. Glutaraldehyde activation of the support for Schiff's base formation is a highly versatile procedure functioning with several surface chemistries, whilst remaining selective to enzyme amine groups [281]. The orientation of the enzyme bound to the carrier surface may be controlled either through the selection of a unique functional group on the enzyme surface [282] or via the use of engineering techniques to introduce chemical functionalities into the enzyme structure which enable site-specific immobilisation [279, 280]. Either strategy maintains the availability of the active site of the enzyme and so minimises any reduction in average activity. Nevertheless, the primary benefits of covalently immobilising the enzyme instead of via physisorption, is the increased stability to denaturing conditions such as extreme pHs, temperatures or (in the case of organic solvents) hydrophobicities due to the increased rigidity provided through covalent fixation. Furthermore, enzyme leaching is minimal, enabling the use of a wider range of operating conditions in which the heterogeneous catalyst may function [283]. However, covalent immobilisation is more time intensive to implement than physisorption and, due to the alteration in the enzyme structure and increased rigidity, has been correlated with a loss in activity of the enzyme, which may only be observed after repeated cycles of use [284]. Nevertheless, the balance between stability and activity may be controlled through the use of linker arms to reduce surface interactions and minimise steric hinderance [285] or multipoint covalent immobilisation which increases rigidity and stability [286].

3.3.1.3. Enzyme encapsulation/entrapment

Encapsulation and entrapment of the enzyme both involve its protection from the external environment via physical confinement, but the former is within a structure with a semipermeable membrane such as vesicle, whilst the latter, a porous support material. There are fewer examples of encapsulation in the literature and as such entrapment techniques appear to be often mistaken for encapsulation. However, they differ significantly in their preparation. Encapsulated enzymes are extruded into liposomes or vesicles comprised of amphiphilic copolymers [287]. Conversely, entrapment results from the dissolution of the enzyme into a

gelating or fibre-forming mixture before addition of a cross-linking, polymerising or precipitating agent which results in a porous structure forming around the enzyme [288]. The benefits of both strategies is that the enzyme remains in its native conformation due to minimal interaction with the support matrix, whilst also remaining protected from the external environment, thereby preserving its activity [280]. However, both also suffer from mass-transfer limitations through the semi-permeable membrane or porous network [271], which in-turn makes them ill-suited for the study of large substrate species. Additionally, in the case of enzyme encapsulation, this limited rate of diffusion may lead to membrane bursting due to the accumulation of high concentration of substrate degradants [287]. If the pore size of entrapped enzyme carriers is increased to reduce mass-transfer limitations, this may result in increased enzyme leaching from the support [280].

3.3.1.4. Enzyme cross-linking

Cross-linked enzymes, unlike other immobilisation techniques, benefit from not requiring a support carrier onto which the enzyme is entrapped or fixated. Rather, the enzymes are first precipitated, using a non-denaturing precipitant such as ammonium sulfate, before being covalently reacted together through the addition of a cross-linking agent such as glutaraldehyde [289]. The resultant species are known as cross-linked enzyme aggregates (CLEAs). The properties of the CLEA may be adjusted significantly by the selection and the amount the cross-linking agent utilised [290]. The primary benefit of the use of CLEAs is that they are cheap, requiring no support carrier for immobilisation and being able to use crude enzyme as a raw material since precipitation is inherently a purification technique, whilst providing an increase in stability [291]. Furthermore, the lack of a support increases the relative catalytic density of the heterogeneous catalyst in reaction. However, the preparation of CLEAs is much more complex than alternative techniques requiring careful consideration of the conditions of each reaction step to ensure the desired properties [292]. Additionally, the structure of CLEAs shows lower

activity retention whilst also being relatively shear labile, necessitating the use of more gentle separation processes [292].

3.3.1.5. Novel immobilisation strategies

Magnetic nanoparticles have been applied as the solid support for enzyme immobilisation as they allow for more homogeneous catalytic behaviour whilst enabling reusability and easy separation via their magnetism [293]. The nanoparticles serve as the solid support for enzyme immobilisation as the surface may be functionalised to create sites for immobilisation [294]. Nano-sized magnetic materials, typically iron oxide due to its large surface area and exhibiting superparamagnetism at ambient temperatures. The technique has been applied to proteolytic enzymes with mutant T70I functionalised magnetic silica nanoparticles maintaining >78% of their proteolytic activity after 7 reuses [293].

Nanoflowers are organic-inorganic hybrid nanostructures synthesised via biomimetic mineralisation, wherein metal ions coordinate with enzyme electron-donating groups to form a hierarchical flower-like structure [295]. They typically display very high catalytic activity due to the increased surface area for catalytic accessibility and preservation of the enzyme conformation. For example, trypsin functionalised $Zn_3(PO_4)_2$ nanoflowers with 6.5% enzyme content and a surface area of 68.35 m^2/g showed superior activity to free trypsin, maintaining >80% activity after 12 reuse cycles [296]. Additionally, nanoflowers improve the stability of the enzyme to solution conditions with zinc-hybrid collagenase nanoflowers showing higher activity than the free enzyme across a range of pHs, temperatures and organic solvents [295].

Stimuli-responsive smart polymers are materials for which the properties change in response to an external stimulus including light, pH or temperature and have been comprehensively reviewed by Balcerak-Wozniak et al. [297]. They are typically hydrogel structures which swell or contract in response to a stimulus to expose or protect the functionalised-enzyme and control diffusion pathways. Therefore, they exhibit great improvements in operational storage and stability to free

enzymes; for example, Eudragit-100 conjugated α-amylase exhibited a half-life >10 times longer than the free enzyme [298].

Positioning and orientation of the immobilised enzyme may be controlled through complementary base-pairing with single-stranded deoxyribonucleic acid (DNA) tags functionalised to the solid support surface [299]. This technique, referred to as DNA-directed immobilisation or DNA hybridisation-drive attachment is a non-covalent immobilisation strategy with high affinity and as such enables low levels of leaching and avoids harsh chemistries that may reduce activity. The strategy allows for proteolytic activity of the immobilised enzyme; for example, horseradish-peroxidase conjugated to a phase-transitioned lysozyme modified TiO_2 solid support via DNA linker achieved 90.52% recovered activity [300].

3.3.1.6. Immobilisation efficiency

Typically, during a reaction, it would be appropriate to measure its yield to estimate the effectiveness of the procedure. However, in the case of enzyme immobilisation, and in particular CLEAs where yield is significantly more difficult to measure, it is perhaps more accurate to evaluate a protocol via the immobilisation efficiency [280]:

$$Efficiency = \frac{Immobilised\ Enzyme\ Activity\ (^{U}/_{g})}{Total\ Free\ Enzyme\ Used\ Activity\ (^{U}/_{g})} \qquad \textbf{Equation 3.4}$$

Whilst this may be a useful way of evaluating the immobilisation procedure, it is not necessarily optimal for comparing the finalised immobilised enzyme to the free enzyme. Therefore, what is often reported is the activity recovery, with which the immobilisation efficiency should not be confused [280]:

$$Efficiency = \frac{Immobilised\ Enzyme\ Activity\ (^{U}/_{g})}{Free\ Enzyme\ Used\ Activity\ (^{U}/_{g})} \qquad \textbf{Equation 3.5}$$

CHAPTER 4

4. RESEARCH GAP AND THESIS OBJECTIVES

4.1. Research Gap

There are various formulation strategies being researched to improve the oral bioavailability of peptide therapeutics. Several of these propose encapsulation of the peptide aiming to shield the drug-payload from the GI environment in addition to improving its intestinal permeability. These strategies include solid lipid nanoparticles [301], self-emulsifying drug delivery systems [302], nanoparticles [303] and microneedle injectors [49, 304]. However, other strategies aim to improve oral bioavailability by increasing the GI stability and epithelial permeability of the peptide therapeutic. These may include structural modifications of the peptide active pharmaceutical ingredient (API) such as cyclisation [305], amino acid substitution [306] or conjugation with cell penetrating peptides [307] or lipids [308]. Additionally, formulation strategies to improve peptide therapeutic GIT stability and permeability have also been investigated including the use of proteolytic enzyme inhibitors, such as in the formulations clinically trialled by Oramed [254], or most notably the use of permeation enhancers, as in the oral formulations of semaglutide [221] and Octreotide [309]. Unlike formulation strategies that encapsulate the peptide therapeutic payload, these techniques do not avoid the changing solution conditions and proteolytic enzymes present within the GI environment, and as such require experimental screening for their initial optimisation prior to animal testing and clinical trials.

There are several techniques typically applied to analyse peptide conformation and measure their physical stability, outlined in Section 2.4, and which have previously been applied in predicting the propensity to aggregate or denature under GI conditions. However, these

strategies generally have at least one of three major drawbacks with respect to their application in the study of peptide aggregation propensity: they have low throughput, are peptide-sample intensive or poorly mimic the GI environment. Self-interaction chromatography (SIC), the theory and application of which was also reviewed in Section 2.4.3, respectively, is a technique that has been previously applied to estimate the second virial coefficient of proteins [212] and adapted for use as a screening tool to predict the aggregation propensity of monoclonal antibodies [219]. The technique has seen limited use for studying peptides [218]. Therefore, there is opportunity to assess its feasibility in this field as it may be automated, allowing for high throughput, it is not sample intensive, lower costs; and it better simulates the ingestion of oral dosages into the GI tract environment than current techniques.

In the study of the proteolytic degradation of peptides by GI enzymes, typically reaction timepoint sampling of the solution reaction is the only strategy applied, the reasons for which have been outlined in Section 3.3. There are several drawbacks to this method: the peptide is required to be separated from the enzyme post reaction; the separation procedure may result in loss of sample peptide, reducing the accuracy of the final results; the method consumes a large amount of substrate peptide and enzyme, whilst also being time intensive. Immobilised enzyme reactors, also reviewed in detail in Section 3.3.1, despite being a well-understood field, have yet to see widespread application in assessing the proteolytic stability of peptides. The application of a packed-bed, immobilised enzyme reactor would allow for automated study of the proteolytic stability of peptides in a high-performance liquid chromatography (HPLC) set-up, therefore not requiring any additional investment in equipment. Additionally, it removes the requirement for post-reaction substrate-peptide separation and better simulates the GI transit of the orally ingested peptide.

4.2. Thesis Objectives

The primary target of this thesis is to develop chromatographic tools to understand the physical and proteolytic stability of peptides under various gastrointestinal conditions. A specific focus will be the use of self-interaction chromatography and immobilised enzyme reactors to achieve this goal.

The first objective in appropriating self-interaction chromatography for peptides was to establish whether the procedures and equipment previously applied for monoclonal antibodies are applicable to peptides. This objective will involve assessing the stationary phase choice and its end-capping to ensure minimal interaction of the mobile phase sample peptide with the stationary phase under various conditions. The effect of stationary phase end-capping on the sample elution in SIC has seen little previous study and so will provide a novel aspect to the work. Additionally, the flowrate to column volume ratio will require optimising to ensure optimal separation of the mobile phase peptide for a given column volume. The lessons learned through these studies could then be applied to the development of both immobilised enzyme reactors and SIC columns.

The second objective was to establish the accuracy and consistency of the self-interaction chromatography method with regards to estimating the physical stability of selected peptides, semaglutide, liraglutide and insulin, under GI conditions. SIC has seen limited previous application with respect to studying peptides, and not at all with respect to studying sample stability under gastrointestinal conditions. Therefore, using the technique for these samples and these conditions represents a novel aspect to the work in this thesis. The work in this section of the thesis will necessitate establishing an immobilisation procedure that ensures high stationary phase coverage and testing the method for model peptides under an array of simulated GI conditions. The resultant calculated second virial coefficient measurements may then be correlated and compared to the peptide size distribution data gathered using more established

techniques, such as dynamic light scattering and size exclusion chromatography to determine the accuracy of the method.

The third and final objective was to determine the feasibility of an immobilised pepsin reactor to study the proteolytic stability of oral therapeutics under gastric conditions. This work will initially necessitate the selection of an appropriate resin and immobilisation procedure that allows for high enzyme coverage on the support. It will require comparison of solution enzyme reaction data with that of the immobilised enzyme reactions to establish whether immobilisation results in a change in enzyme selectivity, whilst varied sample concentrations will allow for quantitative analysis of any changes in activity. The application of a covalently-bound pepsin packed-be reactor to study protein degradation rates has previously been explored, however, the effectiveness of the method to study peptide degradation under gastric conditions has not. Therefore, this aspect of the work will be novel within this scope.

CHAPTER 5

5. MATERIALS & METHODS

5.1. Materials

All reagents and species outlined in the following section were used as received from the suppliers unless otherwise stated.

5.1.1. Insulin

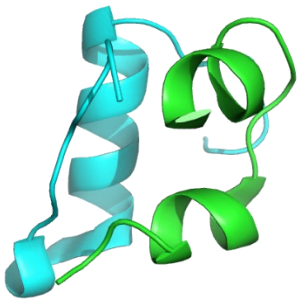


Figure 5.1: Visual representation of the three-dimensional structure of human insulin *(PDB ID:3I40) visualised using PyMOL. Chains A and B are represented in green and cyan respectively*

Insulin is a pancreas-secreted peptide hormone responsible for the regulation of blood-glucose levels. Its primary structure consists of two peptide chains, chain A and B being 21-mers and 30-mers [310], respectively, with three disulfide bridges, namely between A6 and A11, A7 and B7, and A19 and B20, linking the chains [4] (see Figure 5.2). Therefore, since its structure consists of 51 amino acids, it is sometimes referred to as a protein rather than a peptide. In its native state, the A-chain forms two α-helix regions A1-A8 and A12-A19 towards the N- and C-terminals, whilst the B-chain is composed of two β-turns, B6-B9 and B20-B23, separated by an α-helix, B9-B19, with a B-turn conformation at the C-terminal, B24-B28, whilst the initial sequence at the N-terminal has a more random coil structure [311] (see Figure 5.1). The most physiologically

prevalent conformation insulin takes is as a hexamer, consisting of three dimers coordinated with two Zn^{2+} ions, which dissociates to form dimers upon a decrease in Zn^{2+} concentration, and eventually monomers [156]. Additionally, insulin may also form stable tetramers at elevated concentrations in a Zinc-free environment. Despite these stable oligomeric structures, insulin does still form fibrils. The partial unfolding of the α-helix structures has been suggested to be the initiating factor in fibrillation [311]. The monomers then agglomerate through B-sheet interactions, reportedly between parallel B11-B17 sections of monomers [312] until they form ring-like oligomers with inner and outer diameters estimated to be 2 nm and between 6 nm and 7 nm, respectively [313]. It has been reported that these are the nuclei of fibrillation, with two intertwining to form a protofibril with a diameter of 10 nm [314]. Insulin appears to have relatively regular fibrillation kinetics, with its lag time correlating linearly with peptide concentration [311]. However, the shortest lag times were observed at very mildly basic pHs 7-7.8, not at the isoelectric point, which is suggested to be due to protonation of the His-B10 residue increasing electrostatic interaction strength between monomers initiating aggregation [315]. The insulin utilised in these experiments was human recombinant insulin (>98% purity, pI ≈ 5.4, molecular weight 5807.57 Da) purchased from Sigma Aldrich.

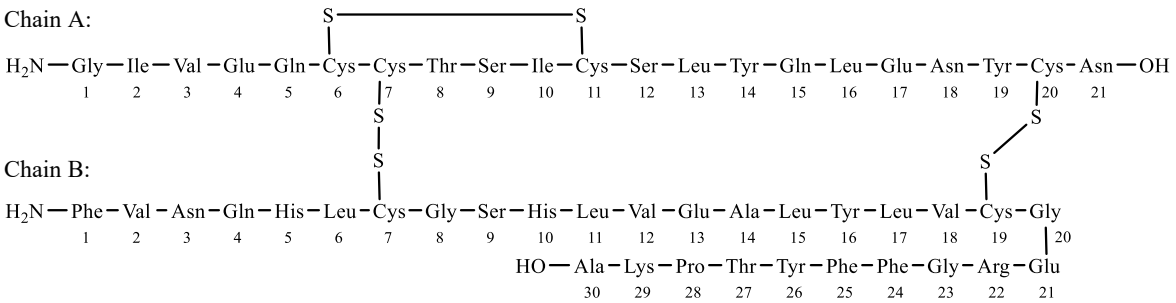


Figure 5.2: Representation of the human insulin amino acid sequence including cysteine bridges

5.1.2. Glucagon-Like Pepide-1 (GLP-1) Analogues

Glucagon-like peptide-1 (GLP-1) is a 31-mer peptide hormone secreted in response to food-intake to regulate blood-glucose levels. It functions via three mechanisms: promoting pancreatic

insulin secretion, supressing the release of glucagon from the liver and increasing satiety by slowing gastric emptying [2, 316]. However, endogenous GLP-1 (7-37) has limited therapeutic use due to it being rapidly degraded in the blood by the enzyme dipeptidyl peptidase IV (DPP-4), cleaving the bond between Ala-8 and Glu-9 [74], resulting in a very short half-life in the plasma of only 1-2 minutes [317]. To overcome these issues, pharmaceutical companies have developed several GLP-1 analogues for clinical use with structural modifications which improve stability but maintain their therapeutic effect. However, the focus of this work will only include semaglutide and liraglutide, so only those structural modifications will be analysed (see Figure 5.3).

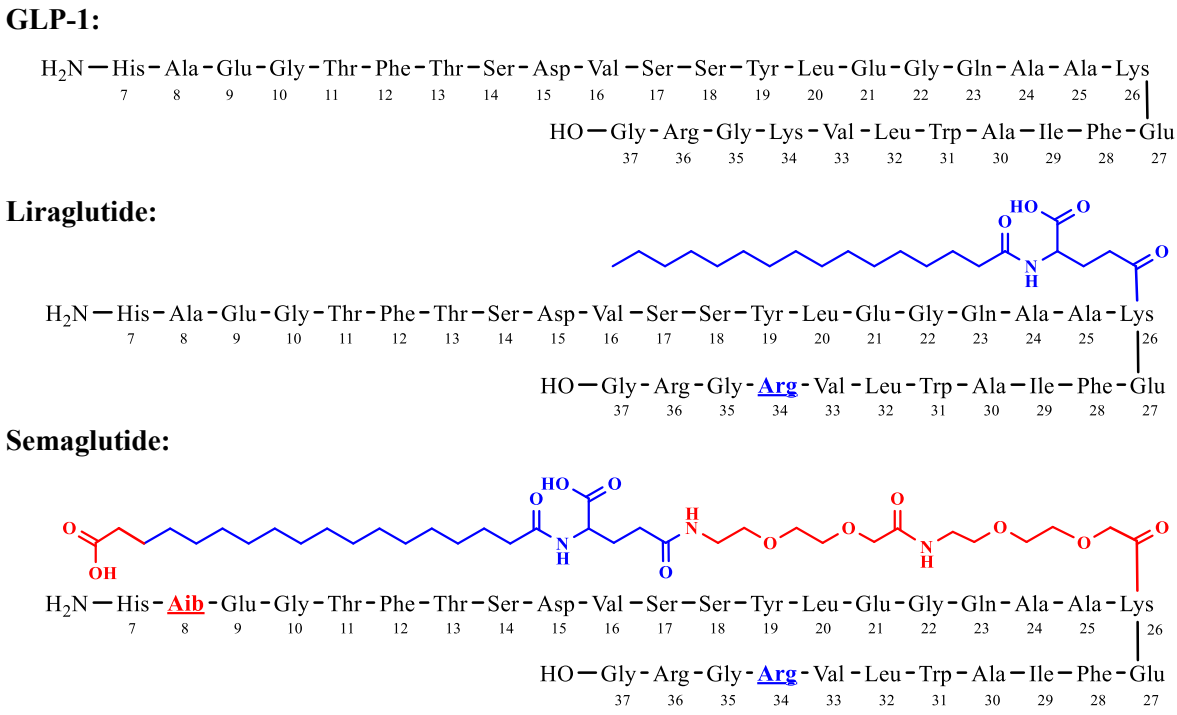


Figure 5.3: Amino acid sequences of GLP-1 and its analogues liraglutide and semaglutide *with their differences highlighted in Blue and Red*

Under more highly charged conditions, GLP-1 exhibits relatively regular fibrillation formation kinetics with increasing peptide concentration reducing the lag time of fibril formation [318]. However, at pHs closer to the isoelectric point of the peptide, with increased α-helix formation, GLP-1 forms two competing aggregate structures, an oligomer precursor to the fibril and an off-

pathway oligomer [158]. This results in an increased lag time with increasing peptide concentration, as the off-pathway oligomer species preferentially forms, thereby inhibiting fibril formation [158]. This analysis does not contradict the assertion that peptides generally aggregate more quickly at conditions closer to their isoelectric point, only that more smaller aggregates are formed more quickly and larger fibrils more slowly under these conditions. Since this transition between the two behaviours occurs between pH 7.5 and pH 8.5, it has been suggested that protonation of the N-terminus, with a pKa ≈8, is key in the increased α-helix structure responsible for off-pathway oligomers and increased lag-time in fibril formation [158].

5.1.2.1. Liraglutide

Liraglutide is a synthetic analogue of GLP-1, administered subcutaneously to manage type-2 diabetes and first approved under the brand name Victoza in 2009 and 2010 by the European Medicines Agency EMA and the US Food and Drug Administration (FDA), respectively. The structure of liraglutide is homologous to that of GLP-1 with the addition of two key modifications. Firstly, a palmitoyl moiety is conjugated to the amino R-group of Lys-26 via a γ-glutamic acid spacer [319] to enable the peptide to bind to Human Serum Albumin (HSA) and so reduce its rate of renal excretion and DPP-4 degradation, thereby increasing its half-life in the blood plasma [320]. The second modification is the substitution of Lys-34 with Arg-34, the purpose of which is to prevent the palmitoyl side chain being functionalised to the wrong site [140]. Liraglutide has been observed to exhibit more complex aggregation behaviour than GLP-1. Above ~ pH 7, it has been observed to form off-pathway, octameric micelles, whilst when the acidity is reduced to pH 6.9 and below, these transition to dodecamers [192]. Nevertheless, within the pH range 6.4 to 8.1, an increase in β-sheet structural confirmation has been observed after a 43 hr lag time, indicating fibril formation [193]. Experiments were performed using liraglutide (pI~4.9, molecular weight 3751.20 Da) supplied by Eli Lilly & Company.

5.1.2.2. Semaglutide

Similar to liraglutide, semaglutide is also a synthetic analogue of GLP-1. It was first approved as subcutaneous injection for treatment of type-2 diabetes under the brand name Ozempic in 2017 and 2018 by the FDA and EMA respectively. However, unlike liraglutide, an oral formulation, Rybelsus, has since been approved and a formulation of semaglutide has also been approved for weight management under the brand name Wegovy. Its structure incorporates the same modifications to GLP-1 that liraglutide does with three additional changes. First, the lipid chain conjugated is not a palmitoyl group but instead a stearic acid group with the carboxylate group enabling stronger interactions with HSA and so a further extended half-life [5]. The second difference to liraglutide is the incorporation of two additional ethylene glycol groups, linked by amide bonds, to the γ-glutamic acid spacer thereby increasing the mobility and flexibility of the conjugated lipid side chain [319]. The final difference to liraglutide is the substitution of Ala-8 for an amino-isobutyric acid (Aib) group to protect against DPP-4 degradation of the peptide, as the Ala-8–Glu-9 bond specifically is the site susceptible to proteolytic cleavage (see Figure 5.4 for Aib structure). Semaglutide despite its structural similarity to liraglutide has been observed to display more regular fibrillation kinetics, with a shorter lag time at slightly acidic pHs (closer to the isoelectric point of the peptide) relative to at slightly basic pHs, with very long lag times on the order of weeks for β-sheet formation to be observed [320]. Experiments were performed using semaglutide (pI~5.4, molecular weight 4113.64 Da) supplied by Eli Lilly & Company.

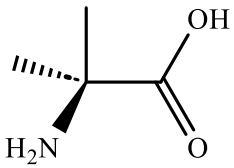


Figure 5.4: Structure of modified Alanine residue Amino-isobutyric acid (Aib)

5.1.3. Peptide Hydrophobicity Library

A set of six peptides, P1-P6 (see Table 5.1), have been proposed as retention standards and hydrophobicity indices for reversed-phase liquid chromatography (RPLC) by Krokhin and Spicer [321]. The use of these peptides in proteolytic activity studies has previously been limited. However, due to the increasing mass fraction of leucine in these standards, which is susceptible to proteolytic cleavage from pepsin, their consistent molecular weights and the limited number of cleavage sites present, this peptide library would also make a very good library for the qualitative measure of pepsin activity. P6 with the highest leucine content would be expected to be the most susceptible to pepsin proteolytic cleavage and P1 and P2 the least. The peptide sequences (purity ≥ 95%) were ordered from Genscript Biotech, Oxford, United Kingdom.

Table 5.1: RPLC peptide retention standard sequences and molecular weights [321]

Designated Name	Sequence	Molecular Weight (Da)
P1	$H_2N-Leu-Gly-Gly-Gly-Gly-Gly-Gly-Asp-Gly-Ser-Arg-OH$	888.89
P2	$H_2N-Leu-Gly-Gly-Gly-Gly-Gly-Gly-Asp-Phe-Arg-OH$	891.94
P3	$H_2N-Leu-Leu-Gly-Gly-Gly-Gly-Asp-Phe-Arg-OH$	890.99
P4	$H_2N-Leu-Leu-Leu-Gly-Gly-Asp-Phe-Arg-OH$	890.05
P5	$H_2N-Leu-Leu-Leu-Leu-Asp-Phe-Arg-OH$	889.11
P6	$H_2N-Leu-Leu-Leu-Leu-Leu-Asp-Phe-Arg-OH$	1002.26

5.1.4. Other Peptides

A library of other peptides was also studied to examine their susceptibility to enzymatic degradation. In this Section, they have been grouped into various categories depending on their structure and purpose, since part of the reason of the study was to analyse the effects peptide structure had on the rate of proteolysis.

5.1.4.1. Taspoglutide

Taspoglutide is another GLP-1 analogue intended to manage type-2 diabetes, but unlike semaglutide and liraglutide its development was halted in Phase III clinical trials due to adverse gastrointestinal side effects and allergic reactions [322]. Unlike semaglutide and liraglutide, it does not have a conjugated lipid chain to improve its stability nor a substitution of Lys-34 to Arg. Instead, like semaglutide it has a methylation of Ala-8 to form amino isobutyric acid, thereby inhibiting DPP-4 cleavage [323]. However, it additionally substitutes Gly-35 and Gly-37 for amino-isobutyric acid and an amide group respectively to increase C-terminal alpha-helicity and improve its stability to plasma Serine proteases [323] (see Figure 5.5). The taspoglutide used in these experiments with a pI~4.1, molecular weight 3339.76 Da and ≥95% purity was purchased from Sigma Aldrich in the form of taspoglutide trifluoroacetate.

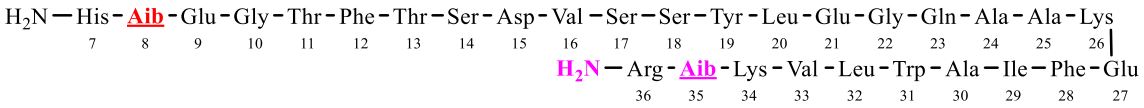


Figure 5.5: Amino acid sequence of taspoglutide with differences to GLP-1 highlighted in pink and red

5.1.4.2. Bradykinin-derived Fragments

Bradykinin is a nonapeptide derived from the enzymatic degradation of kininogens functioning primarily as a vasodilator and inflammatory mediator within the human body [324]. [Lys-des-Arg^9]-bradykinin is a modified synthetic form of bradykinin employed in the study of B1 receptors. Its structure differs from native bradykinin through the addition of a Lysine group at the N-terminal and the deletion of the Arginine at the C-terminal. Therefore, the structure of bradykinin (2-9) and Lys-des-Arg^9-bradykinin may show application in the study of Trypsin activity, due to the similarity in their sequences other than the locations of Arg and Lys groups, to which Trypsin is particularly active. Additionally, the locations of the Pro groups are of use to study their inhibitory effects on Trypsin activity. Bradykinin (2-9) with a pI~10.0, a molecular weight of 904.02 Da and a

purity ≥97%, and [Lys-des-Arg^9]-bradykinin with a pI~11.0, a molecular weight 1032.21 Da and a purity ≥95% were used in these experiments and were purchased from Sigma Aldrich.

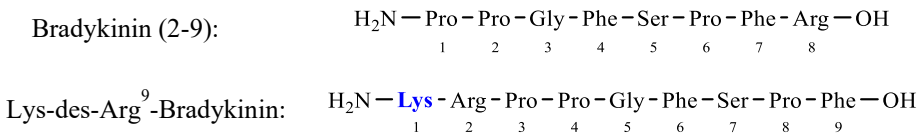


Figure 5.6: Amino acid sequences of bradykinin (2-9) and [Lys-des-Arg^9]-bradykinin *with the additional Lysine group highlighted in Blue*

5.1.4.3. Substance P

Substance P is an undecapeptide from the tachykinin family of neuropeptides, produced by neurons and as such functioning as a neurotransmitter and neuromodulator [3]. It is of interest in the study of oral peptide proteolysis due to its amphipathic, cell-penetrating structure being similar to strategies employed to improve peptide delivery. Experiments were conducted using substance P with a pI~10.6, molecular weight of 1347.63 Da and a purity ≥95% which was purchased from Sigma Aldrich.

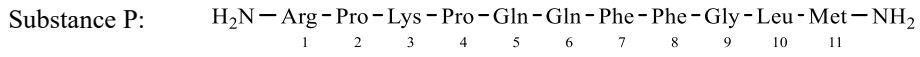


Figure 5.7: Amino acid sequence of substance P

5.1.4.4. Oxytocin and Vasopressin Analogues

Oxytocin and vasopressin are both cyclic, nonameric neuropeptides produced in the hypothalamus and functioning in social bonding [325] and kidney regulation [326], respectively. Their peptide sequences are homologous, including the disulfide bridge between Cys-1 and Cys-6, but there are two key differences at the third and eighth residues where vasopressin has Phe-3 and Arg-8 whilst oxytocin has Ile-3 and Leu-8, respectively. [Lys^8]-vasopressin is a hormone natural to pigs whilst terlipressin is a synthetic analogue of vasopressin prescribed for hepatorenal syndrome [327]. For both analogues Arg-8 is substituted for Lys, whilst for

terlipressin there is additional triglycine sequence conjugated to the N-terminal (see Figure 5.8). Each of these structures were applied to study the effects slight changes in sequence and cyclicity may have on the rate of a peptide proteolysis. Oxytocin acetate with a pI~7.6, molecular weight of 1007.19 Da, and purity ≥ 97%; [Lys^8]-vasopressin with a pI~10.9, molecular weight of 1056.22 Da and purity ≥ 98%, and terlipressin acetate with a molecular weight of 1227.37 Da and purity ≥98% were used in experiments. All were purchased from Sigma Aldrich.

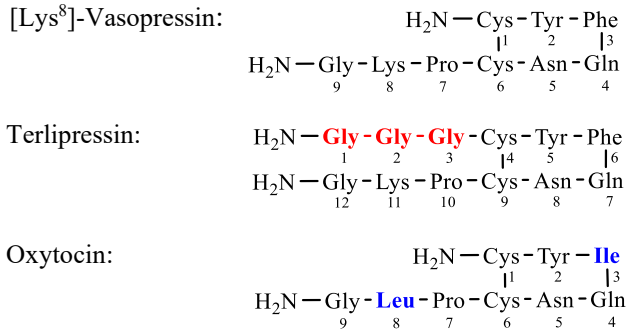


Figure 5.8: Amino acid sequences of [Lys^8]-vasopressin, terlipressin and oxytocin *with their differences in structure from Lys^8-vasopressin being highlighted in Red and Blue, respectively*

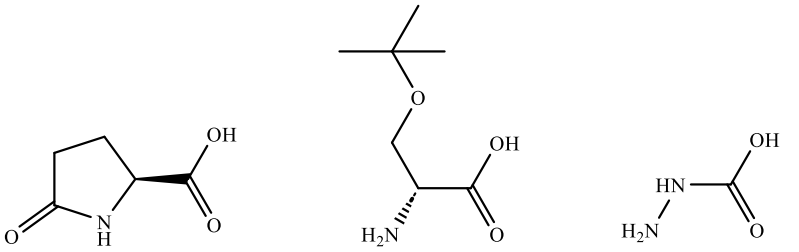
5.1.4.5. Gonadotropin-releasing hormone analogues

Gonadotropin-releasing hormone, also known as luteinising hormone-releasing hormone (LHRH), is a decapeptide secreted by the hypothalamus that acts to stimulate the release of luteinising hormone and follicle-stimulating hormone [328]. Goserelin is a synthetic analogue of LHRH, prescribed for the treatment of hormone-sensitive cancers and gynaecological conditions [329]. It has two modifications to the structure of LHRH, with Gly-6 and Gly-10 being substituted for tert-butyl D-Serine and Aza-glycine (a modified glycine residue in which the α-Carbon is substituted for nitrogen), respectively (see Figure 5.10). The structures are relevant to the study of pepsin activity due to the modifications increasing hydrophobicity but also increasing the steric hinderance to enzyme binding. LHRH acetate with a molecular weight of 1182.29 Da and purity ≥98%, and Goserelin acetate with a molecular weight of 1269.41 Da and a purity ≥99% were utilised in these experiments. Both were purchased from Sigma Aldrich.

LHRH: Pyr – His – Trp – Ser – Tyr – Gly – Leu – Arg – Pro – Gly – NH_2
(1 2 3 4 5 6 7 8 9 10)

Goserelin: Pyr – His – Trp – Ser – Tyr – D-Ser(tBu) – Leu - Arg - Pro – AzGly – NH_2
(1 2 3 4 5 6 7 8 9 10)

Figure 5.9: Amino acid sequence of Luteinising Hormone releasing hormone (LHRH) and its synthetic analogue Goserelin



Pyroglutamic Acid (Pyr) Tert-Butyl-D-Serine (D-Ser(tBu) Aza-Glycine (AzGly)

Figure 5.10: Structures of modified amino acids present in gonadotropin-releasing hormones

5.1.5. Pepsin

Pepsins are aspartate protease active in the acidic gastric environment and are the major proteolytic enzyme involved in the gastric digestion of dietary proteins through peptide bond hydrolysis. They are endopeptidases which are secreted as pepsinogens, their inactive proenzyme precursors, from gastric chief cells in the mucosal layer of the stomach lining. The nomenclature of different pepsinogens varies in the literature and so will be briefly summarised here. Pepsinogens may be classified into the following five groups according to the significant differences in their primary structures: pepsinogens A, B and F, progastricsin (also known as pepsinogen C) and prochymosin [330]. In the human gastric environment, seven pepsinogens have been detected and may be numbered accordingly as pepsinogens 1 through 7 (Pg1-7) in the literature [331]. These may be categorised into two groups with Pg1-5 in pepsinogen group I (PGI), which are pepsinogen A (PGA) isozymogens, and Pg6-7 in pepsinogen group II (PGII), which are pepsinogen C (PGC) isozymogens (precursors to gastricsin) [331]. Pepsinogen A is the predominant of these species and undergoes an auto-catalytic activation upon exposure to the acidic gastric environment (pH <5), during which a 47 amino acid peptide is released (often as

two peptides) from the N-terminal, to form pepsin A [332]. The most common isozyme of pepsin A is pepsin 3 (PGA3) [333] and as such this will be the human pepsin structure used here as a reference.

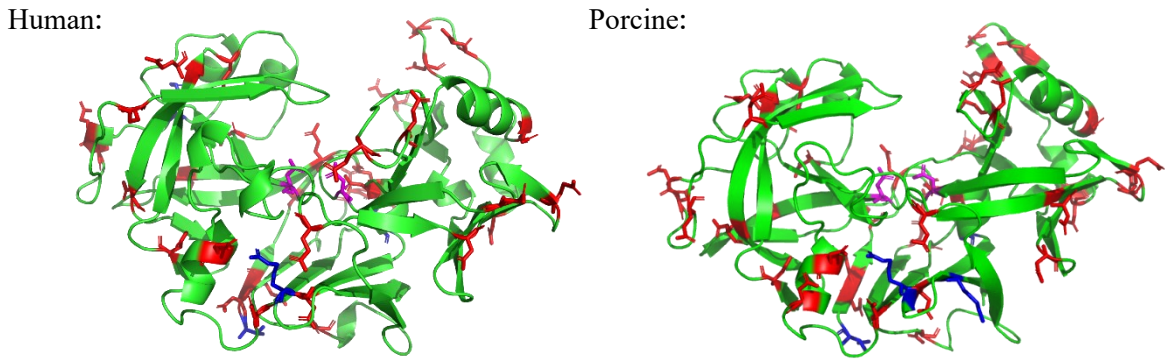


Figure 5.11: Schematic representation of three-dimensional structure of human and porcine pepsin

(PDB ID: 1PSO and 4PEP, respectively) visualised using PyMOL. The active site residues Asp32 and Asp215 are highlighted in pink, whilst Asp and Glu groups are shown in Red and the N-terminal, Lys and Arg groups in blue

Despite having 47 differences in their 326 residue primary structures, both human and porcine pepsin form a bi-lobal tertiary structure consisting largely of β-sheets (see Figure 5.11). In both structures, between the two lobes, is the substrate-binding cleft at the centre of which are the two aspartate residues Asp-32 and Asp-215 which constitute the catalytic site [330]. The polarity of the aspartate groups allows it to form hydrogen bonds with water and the substrate peptide bond allowing for the correct alignment of the reagents before cleaving the target bond through an acid-base catalytic mechanism (see Figure 5.12) [334]. Access to the active site is hypothesised to be controlled by two antiparallel beta strands, comprised of residues Leu-71 through Gly-82 in porcine pepsin (same primary sequence in human pepsin 3 other than Val-71), which is commonly referred to as the "flap" [335]. The Lys-320 residue, through the positive charge of its amino R-group, is suggested to interact with "flap" residues influencing its motility and the overall rate of catalysis [336]. The surface of the cleft is comprised of seven subsites, S_4

through S_1 and S'_1 through S'_3, which may accommodate seven corresponding substrate residues, P_4 through P_1 and P_1' through P_3' (where peptide bond hydrolysis occurs between P_1 and P'_1), through intermolecular interactions [337]. The residues which may interact with an ideal heptapeptide substrate at each porcine pepsin subsite were tabulated by Sielecki et al. [337]. The proposed residues comprising each subsite are listed in Table 5.2, with any residue differences to human pepsin noted to establish the similarity in their proteolytic behaviour. For example, the S'_1 subsite exhibits similar specificities for both human and porcine pepsin, given the hydrophobicity of each contributing residue and the similarity of Leu-291 and Val-291, the only residue difference between human and porcine pepsin, respectively, for that subsite [338]. In accordance with this, porcine pepsin has been observed to exhibit similar site specificity to human pepsin, both showing a preference for hydrophobic and aromatic substrate residues at the P_1 and P'_1 sites, respectively [338].

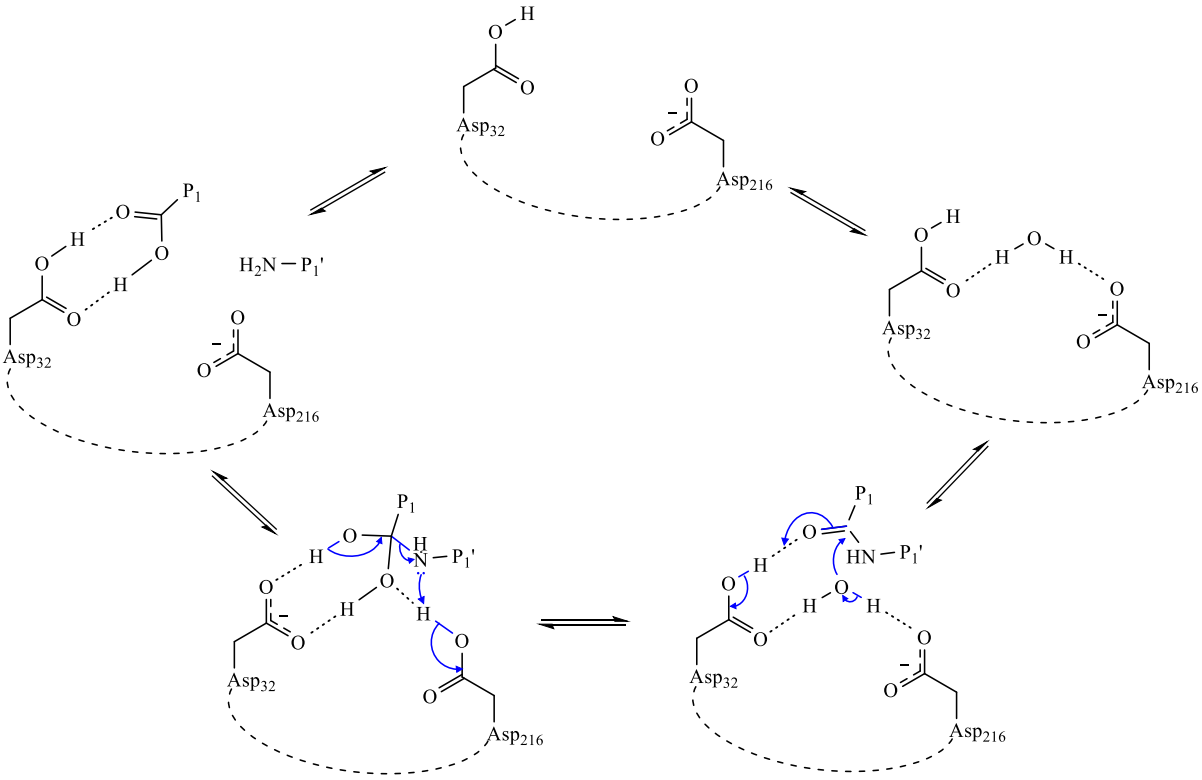


Figure 5.12: Acid-base nucleophilic mechanism of peptide bond hydrolysis by pepsin

Table 5.2: Porcine pepsin residues interacting with residues of an ideal heptapeptide substrate *with differences to human pepsin 3 highlighted.[337]*

Subsite	Residues	Human Pepsin Differences
S_4	**Thr^{12}**, Ser^{219}, Leu^{220}	Met^{12}
S_3	Glu^{13}, **Ile^{30}**, Phe^{111}, Phe^{117}, Ser^{219}	Val^{30}
S_2	Gly^{76}, Thr^{77}, Thr^{218}, Thr^{222}, **Glu^{287}**, Met^{289}	Gln^{287}
S_1	**Ile^{30}**, Asp^{32}, Tyr^{75} Thr^{77}, Phe^{111}, Leu^{112}, Phe^{117}, Ile^{121}, Gly^{217}	Val^{30}
S'_1	Tyr^{189}, Ile^{213}, **Val^{291}**, Thr^{293}, Leu^{298}, Ile^{300}	Leu^{291}
S'_2	Gly^{34}, Ser^{35}, Asn^{37}, Ile^{73}, Tyr^{75}, **Ala^{130}**	Ser^{130}
S'_3	Ile^{128}, Gly^{188}, Tyr^{189}	-

The effectiveness of proteolytic cleavage at a specific site is primarily determined by the ability of the substrate residues to interact with the corresponding enzyme subsites. The strongest influence over proteolytic specificity is considered to be the P_1 residue or S_1 subsite, since the probability of cleavage at this site is highest [330]. Keil, based on their experimental observations, suggested general rules for the probability of peptide bond hydrolysis based on the amino acid sequence in proximity to the cleavage site [339]. These have been tabulated in addition with findings from Powers et al. [340] and Hamuro et al. [341] in Table 5.3 to provide an approximate guide on whether a site would be expected to undergo peptide bond hydrolysis. The data collated is in accordance with what would be expected based on residue interaction strength with the subsite residues. For example, given the number of hydrophobic residues contributing to the S_1 subsite, it is unsurprising that aromatic and hydrophobic amino acids show the greatest cleavage prevalence. By contrast, proline inhibits pepsin binding due to the pyrrolidine ring in its R-group, resulting in the absence of an amide hydrogen reducing its hydrogen bonding capacity and rigidly constraining the N-C^{α} bond to a rotational angle of -58° [342]. Notably, the P_4 position has been found to have little influence over the cleavage probability, simplifying analysis.

Table 5.3: General effect substrate residue location has on likelihood of proteolytic activity of pepsin
[339-341]

Residue	Promoting	Inhibiting
P_4	-	-
P_3	-	Arg, Lys, His
P_2	-	Pro
P_1	Phe, Tyr, Trp, Leu, Met	Lys, His, Arg, Pro
P'_1	Phe, Tyr, Trp, Leu	-
P'_2	-	Pro
P'_3	-	Pro

As a protein, the conformation, stability and therefore proteolytic activity of pepsin are highly pH-dependent. Optimal catalytic behaviour has been observed at approximately pH 2.0, although under such conditions pepsin will also exhibit autolytic behaviour [343]. The proteolytic activity begins to decrease significantly in the pH 4.5-6.0 range, at which point a native-like, catalytically-inactive conformation starts to be adopted [344]. Nevertheless, peptide bond hydrolysis is observed up to pH 6.5, at which point catalytic behaviour is no longer observed, but the structure remains stable such that it may be reactivated upon a reduction in pH [343]. However, pepsin is irreversibly denatured in the range pH 7.0 to 8.0, at which point proteolytic activity is permanently lost. The pepsin studied in subsequent experiments described here was pepsin from porcine gastric mucosa (EC 3.4.23.1, molecular weight 35 kDa, specific activity ≥ 400 units/mg protein) and was purchased from Sigma Aldrich.

5.1.6. Column Packing Resins

Dry AF-Tresyl-650M (0014472) particles were purchased from Tosoh Bioscience and refrigerated at 4 °C prior to use. The surface of the HW-65 methacrylate polymer resin is conjugated with tresyl groups that are highly reactive towards amine and thiol groups, and so are ideal for protein or peptide immobilisation under neutral conditions. The relevant properties of the resin according to the provider's documentation are tabulated in Table 5.4.

Table 5.4: Toyopearl AF-Tresyl-650M properties

Property	Value
Mean Pore Size	100 nm
Mean Particle Size	65 µm
Ligand Density	80 µmol/g (dry)

Reagent grade 3-Aminopropyl-functionalised silica gel was purchased from Sigma Aldrich to function as an irregular silica particle platform with reactive amine sites ideal for peptide or protein immobilisation. The properties of the resin relevant for its use in peptide or protein functionalised packed-bed systems are tabulated in Table 5.5, as provided in the provider's documentation.

Table 5.5: Propylamino Silica resin properties

Property	Value
Mean Pore Size	60 Å
Mean Particle Size	52 µm
Ligand Density	1 mmol/g

5.1.7. Other Materials

The following chemicals were purchased from Sigma Aldrich grouped according to their chemical purity gradings: BioXtra grade dibasic and monobasic sodium phosphate; American Chemical Society (ACS) grade potassium phosphate monobasic and dibasic, hydrochloric acid (37%), sodium hydroxide, tris(hydroxymethyl)aminomethane (henceforth referred to as Tris), citric acid and ethanolamine; Reagent grade sodium cyanoborohydride, which was stored in a dark, dry place and extensively wrapped in parafilm between uses ; and food grade sodium citrate dihydrate. Additionally, Grade I, 25% glutaraldehyde solution in water was also purchased from Sigma Aldrich. However, upon its arrival, it was aliquoted into 1 mL samples, which were

wrapped in aluminium foil and stored at −20 °C prior to use, to minimise any polymerisation effects. Soybean L-α-Lecithin (≥97% purity, MQ100) was also purchased from Sigma Aldrich and stored at -20 °C in a freezer, prior to its use. Analytical Reagent (AR) grade acetone was obtained from Fisher Scientific. Sodium chloride of ACS grade was from VWR.

Buffer solutions were prepared using deionised water processed by a CENTRA water purification system from ELGA. The pH was measured using a Mettler Toledo FiveEasy pH meter and adjusted using 1 M hydrochloric acid or 1 M sodium hydroxide before being filtered using through 0.22 µm bottle top filters from Millipore to remove any remaining particulates. 1 M hydrochloric acid and 1 M sodium hydroxide solutions were in turn prepared with pre-filtered deionised water to prevent any contamination with filter degradation products due to the extreme pHs of those solutions.

5.2. Methods

5.2.1. High-Performance Liquid Chromatography (HPLC) Equipment

Self-interaction chromatography (SIC) experiments were performed using one of two HPLC setups using either a LC-20 series or Nexera series HPLC systems from Shimadzu (Shimadzu UK Ltd, Buckinghamshire, UK).

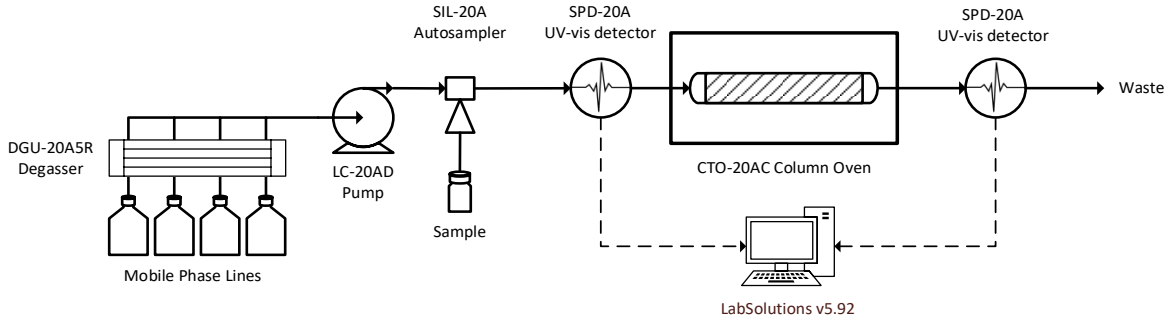


Figure 5.13: Schematic of LC-20 HPLC system setup

The LC-20 series was comprised of a modular set up including a SIL-20A total injector autosampler; a LC-20AD, parallel-type double plunger pump; a five-channel on-line degasser,

the Shimadzu DGU-20A5R; SPD-20A, dual-wavelength UV-Vis detectors with a temperature-controlled flow cell fitted inline pre- and post-column, and a thermostatted, forced air-circulation, CTO-20AC column oven. The solvent inlets were fitted with GLC suction filters (Shimadzu). The system was controlled by a CBM-20A communications module with data collected using Shimadzu's LabSolutions software version 5.92.

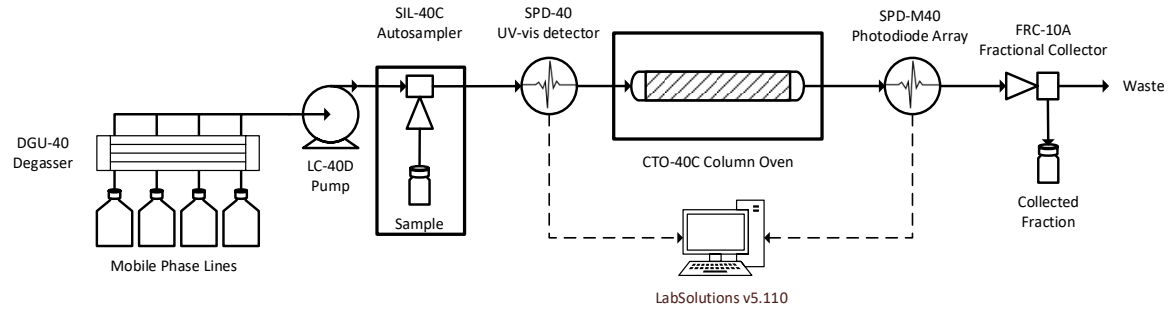


Figure 5.14: Schematic of Nexera HPLC system setup

The Nexera series had a similar setup including a SIL-40C temperature-controlled, total injector autosampler; a LC-40D, parallel-type double plunger pump; a five-channel on-line degasser, the Shimadzu DGU-40; a SPD-40, dual-wavelength UV-Vis detector with a temperature-controlled flow cell, fitted inline pre-column; a temperature controlled, forced air-circulation, CTO-40C column oven, and a SPD-M40 photodiode array (PDA) inline post-column detector. The solvent inlets were likewise fitted with GLC suction filters (Shimadzu). Additionally, the Nexera series had 6 column lines, controlled by a switching valve, allowing for multiple columns to be stored and run in parallel (although not simultaneously). Furthermore, it was also fitted with an FRC-10A fractional collector, for sample collection and further analysis. This system was controlled by a CBM-40lite communications module with data collected using Shimadzu's LabSolutions software version 5.110.

5.2.1.1. Elution Volume vs Retention Volume

Throughout the chromatographic experimental section of this work, the terms elution volume and retention volume will be used. The difference between them is slight and so often they are

used synonymously. In this work, elution volume is used to describe the volume of mobile phase eluted since the start of the run, whilst retention volume is used to describe the required volume of mobile phase from sample injection to cause its elution from the column. Therefore, in this work elution volume will be used to describe chromatograms, elution profiles and any pre-column detector measurements. Conversely, retention volume will be applied in the discussion of measurements and individual data points. Additionally, retention time and retention volume for size exclusion chromatography (SEC) and self-interaction chromatography (SIC) methods exhibit identical trends as they are all performed isocratically at constant flowrate. However, in the case of the reversed-phase methods, since neither of these are the case, retention time will be exclusively used throughout the discussion of that data.

5.2.2. UV-Vis spectroscopy

UV-Vis spectroscopic analysis of samples was performed using an Implen Nanophotometer. Measurements were made at 280 nm and were made across 5 sample pedestals to minimise errors caused by blanking variation and poor sample application. Before each measurement began the sample pedestals were first wiped with an ethanol-wetted tissue and allowed to dry for approximately 15 seconds. 1.5 µL of the background buffer solution was then pipetted onto each sample pedestal and a blank measurement recorded. For each measurement 1.5 µL of sample was applied to each sample pedestal in use and the mean absorption value was taken as the final value. Between measurements, the sample pedestals were wiped with deionised water, then ethanol and allowed to dry for approximately 15 seconds to minimise sample cross-contamination. All measurements (including calibration curve measurements) were performed in increasing concentration order to reduce the effects of sample build-up on the sample pedestals. Sample concentration calibration curves were generated by the same procedure using triplicate samples.

5.2.3. Resin preparation procedures

5.2.3.1. Peptide immobilisation

For each SIC experiment, the peptide or protein of interest was immobilised onto a Toyopearl AF-Tresyl-650M resin from Tosoh. This material is a hydroxylated methacrylic polymer structure (HW-65) that swells in aqueous conditions to form particles of 65 μm diameter and with a 100 nm pore size. The surface of the HW-65 polymer resin has been functionalised with tresyl groups making it highly reactive to amine and thiol groups. Therefore, the relevant peptide was immobilised via its available surface amine groups following the mechanism outlined in Figure 5.15. The experimental method adopted was that previously described by Jacobs et al. [345] but incorporating changes made by Hedberg et al. [346].

Analysing the mechanism in Figure 5.15, it must be noted that it potentially allows for peptides with multiple amine groups to be conjugated at multiple sites. This would reduce their surface mobility and minimise the number of potential orientations that they may adopt, thereby greatly affecting the possible interactions with mobile phase peptides. The method proposed does not provide any mechanism through which to ensure only one surface conjugation site for peptides with multiple surface amine groups. This was for two primary reasons: firstly, since peptide immobilisation was a slow reaction with relatively low yield it was deemed more optimal to maximise the amount of peptide immobilised than control the reaction sites; secondly, since random orientation is a base assumption of SIC, reducing the number of possible orientations through reaction-controlling measures seemed counter-intuitive, even if it provided for increased mobility of the immobilised peptide.

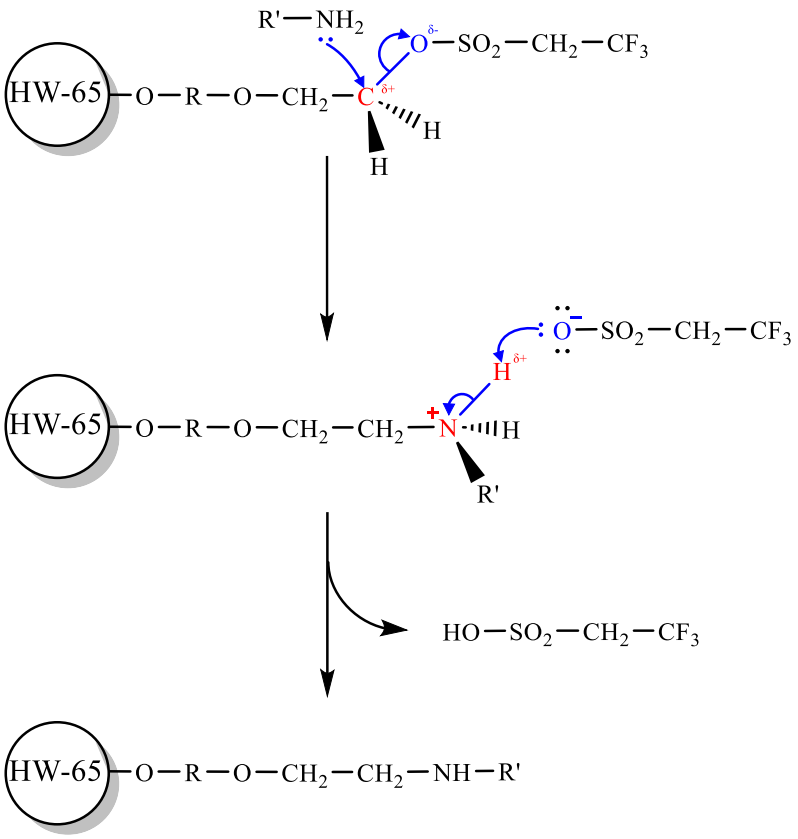


Figure 5.15: Nucleophilic substitution mechanism of the tresyl functionalised surface groups with primary amine groups
(HW-65 is the methacrylic polymer of the Toyopearl AF-650M resin)

The dry resin particles were first weighed out and washed with approximately 50 mL/g resin using deionised water before being allowed to swell for 30 minutes. Once the resin had settled, the supernatant was pipetted off leaving a slurry of approximately 3.5 mL/g resin. If the resin was to be activated with glutaraldehyde, then 1.0 mL/mL.resin of a 25% glutaraldehyde water solution was added and mixed with the resin before being washed 10 times with 15 mL/g resin pH 7.5, 0.1 M potassium phosphate. Glutaraldehyde has been theorised to activate the tresyl surface by Hedberg et al. by undergoing nucleophilic substitution of the tresyl groups leaving an even more weakly bound hemiacetal group. However, here an alternative mechanism is suggested in Figure 5.16, given that in the method performed by Hedberg et al. the glutaraldehyde was added before the resin is equilibrated with neutral/mildly basic buffer [346].

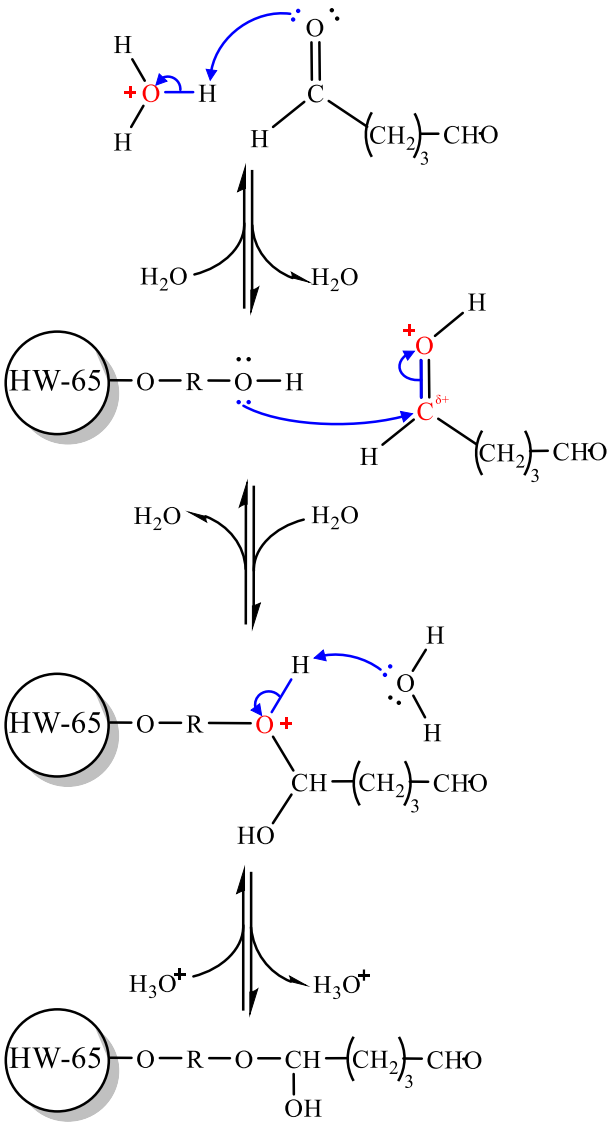


Figure 5.16: Proposed nucleophilic addition mechanism of glutaraldehyde to unfunctionalised surface alcohol groups
(HW-65 is the methacrylic polymer of the Toyopearl AF-650M resin)

Therefore, the glutaraldehyde may form hemiacetal groups with any remaining surface alcohol groups, thereby resulting in additional surface aldehyde groups that may provide immobilisation sites via Schiff's base formation as shown in Figure 5.17 and Figure 5.18. After each wash, the resin was left to settle and the supernatant removed. 30 mg/g resin peptide was weighed out and dissolved at 2.0 mg/mL in pH 7.5 0.1 M potassium phosphate before being added to the resin with 60 mg/g resin sodium cyanoborohydride to catalyse and increase the overall coverage of the reaction. For the support particles without peptide functionalisation (for dead column

preparation), a 0 mg/mL peptide solution was used instead, and the procedure was carried out similarly. The mixture was then placed on a Stuart Scientific Roller Mixer SRT2 overnight. Samples were taken at the start and end of the reaction for UV-Vis spectroscopic analysis using an Implen NanoPhotometer to establish the amount of peptide immobilised (using calibration curves outlined in Figure 12.1 in the Supplementary Materials).

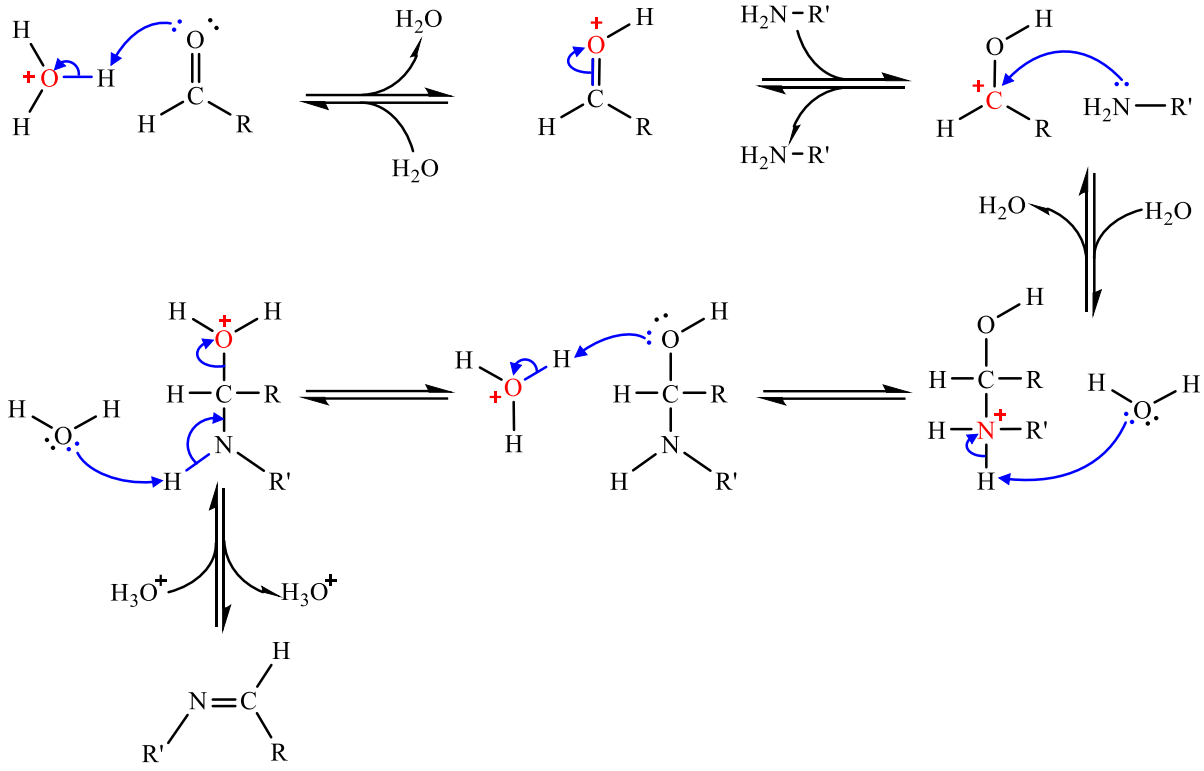


Figure 5.17: Mechanism for Schiff's base (imine) formation between aldehyde and primary amine groups

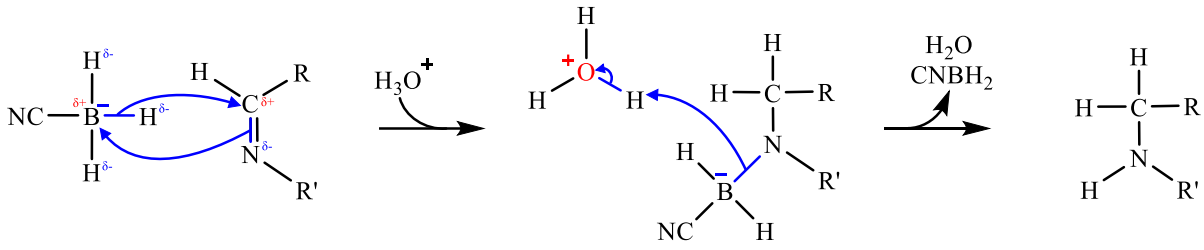


Figure 5.18: Mechanism of Schiff's base (imine) reduction to a Secondary amine via a Cyanoborohydride reducing agent

5.2.3.2. Peptide Resin End-capping and Dead Column Preparation

Upon completion of peptide immobilisation (see Section 5.2.3.1), if the resin was being end-capped with monoethanolamine (MEA), it was washed with 15 mL/g resin 1.0 M pH 8.0 MEA solution. However, if the remaining resin surface tresyl groups were being functionalised with Tris, it was washed with 15 mL/g 0.1 M pH 8.1 Tris-HCl, 0.5 M NaCl then stored in the solution overnight at 4 °C. For both end-capping techniques, the resin was then washed with 20 mM sodium phosphate 50 mL/g resin before being stored at 4 °C prior to column packing. In the case of the preparation of unfunctionalised dead columns, the resin was prepared exactly as in Section 5.2.3.1, instead using 0 mg/mL peptide concentration solution to ensure consistency in the properties between the unfunctionalised and peptide-surface-decorated column resins.

5.2.3.3. Pepsin Immobilisation

Pepsin was immobilised onto a 1 mM/g propylamino silica resin via a two-step procedure, first activating the resin with glutaraldehyde and before immobilising the enzyme. The first step was adapted from work performed by Monsan [347]. The required amount of propylamino silica resin was first weighed out (0.5 g to 1.0 g) before being thoroughly washed with 50 mL/g.resin deionised water to ensure the deprotonation of any silanol groups and some amino groups and allowed to dry for several hours at room temperature and pressure. 0.5 mL/g.resin 25% glutaraldehyde solution was then diluted 1:19 (i.e. by a factor of 20) in 50 mM pH 8.0 potassium phosphate solution to form a pH 8.0 1.25% glutaraldehyde solution which was then added to the resin. The solution was allowed to react following a similar mechanism to that seen in Figure 5.17, whilst being gently mixed on a Stuart Scientific Roller Mixer SRT2 overnight (~15 h). If the resin washing phase was not performed, the resin instantaneously turned a deep red upon exposure to the glutaraldehyde solution, even at very low concentrations, indicating the rapid formation of imine bonds. However, this resulted in poorer future immobilisation likely due to the high surface coverage resulting in polymerisation and cross-linking of the immobilised glutaraldehyde groups.

The washing procedure resulted in a light golden-orange resin colour indicating lower coverage but significantly reducing the potential of cross-linking and polymerisation.

After activation with glutaraldehyde the resin was thoroughly washed with 150 mL/g.resin deionised water. Each individual wash involved the centrifugation of the resin mixture at 10000 rpm for 10 minutes pipetting off the supernatant, then the addition of 15 mL/g.resin of deionised water and thoroughly shaking. If bubbling of the wash solution was still observed upon shaking or pipetting after 150 mL.g resin deionised water, then additional washes were carried out until it was not.

The second step was required to be carried out within 24 hours of the previous being completed to minimise any reactions between glutaraldehyde and the environment causing surface contamination. Prior to the immobilisation reaction, the resin was washed with 50 mL/g.resin 50 mM pH 4.4 sodium citrate solution in order to equilibrate the silica surface with the reaction buffer. The immobilisation procedure was adapted from work performed by Han et al [348] and follows the mechanisms outline in Figure 5.17 and Figure 5.18. A 200 mg/mL porcine pepsin stock solution in pH 4.4 50 mM sodium citrate was prepared and refrigerated at 4 °C. 10.0 mL/g.resin of pepsin stock solution was mixed with 1.75 mL/g.resin of 1.5 M sodium cyanoborohydride ($NaCNBH_3$) solution before being added to the resin and mixed. 17.5 mL 1.5 M sodium sulfate (Na_2SO_4) is then added dropwise to the mixture before the reaction mixture was sealed and incubated at 4 °C under 1000 rpm agitation to prevent resin settling in an Eppendorf ThermoMixer F2.0 for 24 hours. Simultaneously a t=0 h sample solution is prepared with 500 µL pepsin stock solution, 87.5 µL of the $NaCNBH_3$ solution and 875 µL Na_2SO_4 and incubated in the same mixer for measurement of the initial concentration and to account for any autolysis. Meanwhile a control solution using the same volumes of each solution but replacing the pepsin stock solution with pH 4.4 50 mM sodium citrate solution was also prepared.

After 24 hours, the agitation of the reaction was stopped allowing the resin to briefly (approximately 2 minutes) settle such that three individual 0.5 mL samples of the reaction solution may be aliquoted to establish the t=24 hour pepsin concentration. The t=0 and t=24 hour samples were diluted by a factor of 20 with 50 mM pH 4.4 sodium citrate so their concentration was more appropriate for UV-Vis analysis. The amount of pepsin immobilised was then estimated via UV-Vis spectroscopy of the initial reaction mixture and final reaction mixture samples and compared to previously prepared calibration curves (see Figure 12.2 in the Supplementary Material).

5.2.3.4. Pepsin resin end-capping

The pepsin immobilisation reaction was then stopped through the addition of 5 mL/g.resin of 100 mM ethanolamine solution dissolved in pH 4.4 sodium citrate solution before being re-incubated at 4 °C and 1000 rpm overnight (approximately 15 hours). The reaction was completed by ending the agitation and allowing the resin to settle and removing the supernatant before washing the resin with 100 mL/g.resin of pH 4.4 50 mM sodium citrate until the supernatant shows no observable peak versus a control solution of the same buffer during a UV-spectroscopic scan using a NanoPhotometer from Implen. The resin is then stored in the pH 4.4 50 mM sodium citrate solution and refrigerated at 4 °C prior to column packing.

5.2.3.5. Raman Spectroscopy

The surface composition of the peptide and enzyme immobilised resins were analysed via Raman spectroscopy using a Senterra II Raman Microscope (Bruker Nano Analytics, Coventry, UK). The resin slurry was first washed with 50 mL/mL resin deionised water to remove buffer from the slurry. The resin-water slurry was then pipetted onto a 76x26x1 mm glass microscope slide to form an evenly-distributed layer of resin slurry. The slide was then placed in a Medline Scientific™ Jeio Tech 665 L Vacuum Oven (OV-12) overnight at 70 °C ±0.1 °C and 50 kPa ±10 kPa to thoroughly dry the resin (Medline Scientific, Rotherham, UK). An individual sample resin bead

was focussed on using a 20x objective lens through a 50 µm aperture (see an example image in Figure 5.19). Raman spectroscopy measurements were performed across a 4x4 grid of points on the resin bead using a 532 nm wavelength laser with a power setting of 25 mW with a 3000 ms integration time. The spectral resolution was set to 4 cm^{-1} and to reduce the signal to noise ratio, 24 coadditions were gathered for each individual measurement point. The measurements were repeated in triplicate (on three separate resin beads). The data was then processed and analysed using OPUS 8.7 software.

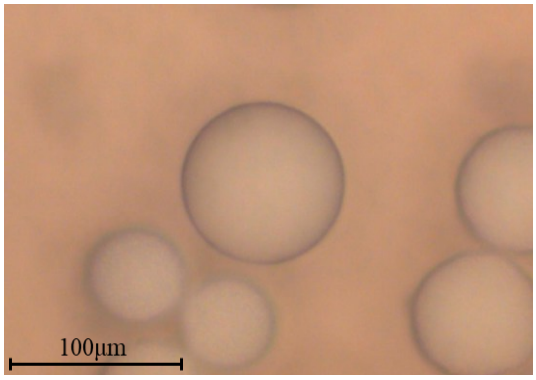


Figure 5.19: Senterra II Raman Microscope 20x objective lens focussed image of AF-Tresyl-650M resin bead

5.2.4.Column Packing

5.2.4.1. Self-interaction Chromatography

The end-capped resin stationary phases were slurry packed into empty glass microcolumns from Omnifit of 25 mm in axial length with a 3 mm internal diameter, The packing procedure was then finalised by eluting the resin storage solution through the column, progressively increasing the flowrate from 0.05 mL/min to 10x the maximum experimental flowrate (with a maximum column pressure of 50 bar). This flowrate was maintained for 10 minutes. Prior to any experiments being carried out, 10 µl injection of 2% v/v acetone in deionised water was eluted through the resin packed glass column at the desired flowrate to establish the quality of the resin packing. A well-packed column is indicated by a Gaussian elution peak detected in the post-column UV-Vis

spectrophotometer. Fronting and tailing denoted an over- and an under-packed column, respectively. If significant fronting, tailing (as calculated in Section 5.2.4.3) or peak splitting were observed the column was repacked.

5.2.4.2. Pepsin functionalised resins

The procedure for enzyme functionalised resins is almost identical to that of the SIC resins, but with a few minor differences. Firstly, the columns used were metal 30 mm length (L) x 4.6 mm internal diameter (ID) from VWR International Ltd, which have an empty volume of 0.499 mL. Secondly, due to the higher column volumes and their higher-pressure ratings, they were packed at a flowrate of 5 mL/min.

5.2.4.3. Determination of Peak Tailing and Peptide Loss

An SIC set-up which implements a pre- and post-column UV-Vis detector improves the estimation of the retention volumes by accounting for the majority of the pre-column tubing and the peak spread. However, it additionally allows for the determination relative amounts of protein/peptide injected and eluted from the column, provided that a calibration curve between the peak area and sample concentration for both detectors has been established [349]. Peak area and not peak height should be calibrated to sample concentration as there is relatively more sample spreading for the post-column peak resulting in a wider, shallower peak. The peak area is calculated through the following equation:

$$Area = \int_{t_{5\%,1}}^{t_{5\%,2}} Absorbance\ dt \qquad \textbf{Equation 5.1}$$

where $t_{5\%,1}$ and $t_{5\%,2}$ denote the times at which the absorbance reaches 5% of the peak maximum height before and after the maximum point, respectively. Therefore, the benefits of this experimental set-up are that it can account for degradation or leakage of the immobilised peptide/protein from the stationary phase and any retention of the injected sample, as an

increase or decrease, respectively, in the post-column detector peak area will be observed relative to the pre-column peak area [349].

An additional factor that may impact elution peak analysis is peak tailing. The causes of peak tailing are numerous [350], however, those associated with poor injection technique or column packing such as the presence of additional dead volume should have been highlighted by the pre-column detector or during the initial acetone injection described in Section 5.2.4.1, respectively. The factors of primary concern and the purpose of this experiment are those associated with ensuring that the solute retention in the stationary phase exhibits linear behaviour and is sufficient to achieve a dynamic equilibrium. The US Pharmacopeia recommends characterising peak tailing using the following equation [350]:

$$Tailing\ Factor = \frac{t_{5\%,2} - t_{5\%,1}}{2(t_{max} - t_{5\%,1})} \qquad \textbf{Equation 5.2}$$

where $t_{5\%,1}$ and $t_{5\%,2}$ are the retention times for the peak to reach 5% of the peak maximum height, which occurs at t_{max}, before and after the peak maximum, respectively (as shown in Figure 5.20).

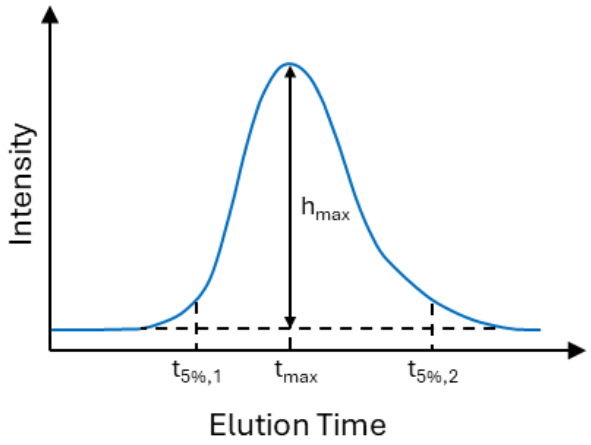


Figure 5.20: Diagrammatic chromatographic elution curve to represent the parameters required for peak tailing calculations

5.2.5. Peptide Aggregation Determination Methods

5.2.5.1. Self-interaction Chromatography

5.2.5.1.1. Sample Preparation

Samples of the peptide under investigation were prepared in triplicate by reconstituting the lyophilized powder in deionised water at 0.1 mg/mL and gently mixing until the peptide fully dissolved to leave a transparent solution. Deionised water rather than the mobile phase condition, which is typical for self-interaction chromatography, was used to ensure that any sample loss on account of instability in solution resulting in aggregation would be consistent across all conditions and potentially reduce losses of the limited amount of sample relative to the number of potential conditions. The samples were then filtered through low protein binding, 0.22 µm cellulose acetate luer syringe filters. The concentration was then verified against a calibration curve (as may be seen in Figure 12.3 in the Supplementary Materials) via UV-Vis spectroscopy using a NanoPhotometer from Implen, before being stored at 4 °C prior to use for a maximum period of 30 minutes. Initially, the aggregation state of the sample peptide was checked via dynamic light scattering at hour intervals to establish the point at which oligomers or aggregate structures begin to be observed. Once this time interval was established, the sample peptide solution was changed at this time henceforth in future experiments for a freshly prepared sample.

5.2.5.1.2. Setup and Experimental Procedure

The experiments were performed in a temperature-controlled laboratory at 20 °C +/-1 °C with mobile phase solvents and injection samples stored at laboratory temperature. All measurements were performed in triplicate. The column oven was set at 37 °C to simulate the gastrointestinal environment. 220 nm and 280 nm were set as the wavelengths for both UV-Vis spectrophotometers to enable the detection of the peptide bonds and aromatic residues, respectively.

Prior to each new mobile phase condition, the column was equilibrated for 10 column volumes (or until the post-column UV-Vis detector's baseline stabilised) of the buffer with the pH and salinity to be studied acting as the mobile phase. Each individual measurement of a solution condition was performed in triplicate and was performed by injecting 10 µL of peptide solution through the column with the pre- and post-column detectors detecting its retention volume. The average of the three retention measurements and their standard error are reported for a range of pHs and salt concentrations.

During initial trials, pH 6.0 and pH 8.0 sodium phosphate buffer mobile phases were utilised, representing the extremes of the range of intestinal solution conditions [39], to establish the stability and efficacy of the resins. A 20 mM buffer concentration was selected, as is typical for SIC [351], to provide sufficient buffering capacity and not require excessively long stationary phase equilibration times, whilst not being too strong resulting in ion charges dominating peptide interactions. Furthermore, to establish the effects of variations in salt concentration, NaCl at 0 mM, 5 mM, 20 mM, 100 mM and 250 mM concentrations were also applied within the mobile phase buffered to the same pHs. 0 mM and 5 mM NaCl conditions provided a clear distinction between conditions in the absence of NaCl and a low NaCl concentration relative to the buffer concentration. 20 mM and 100 mM NaCl conditions were similarly to test the effects of the sodium chloride being at or significantly beyond the buffer concentration, whilst 250 mM NaCl was the salt concentration beyond which the buffer became unstable with increasing concentrations of NaCl, once raised to 37 °C.

5.2.5.1.3. Determination of the Retention Volume

The retention volume, as mentioned in Section 2.4.3, has typically been determined in SIC as the elution volume of the peak maximum. However, the elution volume ($\bar{V}$) of the peak centre of mass has exhibited improved correlation with the second virial coefficient [349] and is defined by the following equation [352]:

$$\bar{V} = \frac{\int_0^\infty cV\ dV}{\int_0^\infty c\ dV}$$ ***Equation 5.3***

Where *c* is the concentration of the peptide/protein being eluted as a function of the elution volume, *V*, of mobile phase. For an elution peak profile without fronting or tailing, that is to say a Gaussian peak, the peak maximum and centre of mass would have the same retention volume.

5.2.5.1.4. Calculation of Theoretical Plate Height

Upon first packing a column for SIC, it is necessary to determine the height equivalent to a theoretical plate (HETP) to establish the column efficiency. An HETP represents a physical length of column required for the mobile and stationary phase to reach equilibrium, essentially acting as a theoretical single separation step. Therefore, a shorter HETP implies more efficient column separation. As such, the calculated values allow for the ability of different stationary phases to separate a given species to be compared, for the determination of the optimum mobile phase flowrate and analysis of the column packing.

The theoretical plate height was calculated using the half peak height method [353]. In this procedure, first the elution time of peak maximum (t_{max}) and the peak width at half the peak height ($W_{1/2}$) are measured as illustrated in Figure 5.21. These values are then related to the number of theoretical plates (*N*) for the given column length by the following equation:

$$N = 5.54\left(\frac{t_{max}}{W_{\frac{1}{2}}}\right)^2$$ ***Equation 5.4***

The number of theoretical plates is then related to the HETP (*H*) by the column length (*L*), a known parameter of the column:

$$H = \frac{L}{N}$$ ***Equation 5.5***

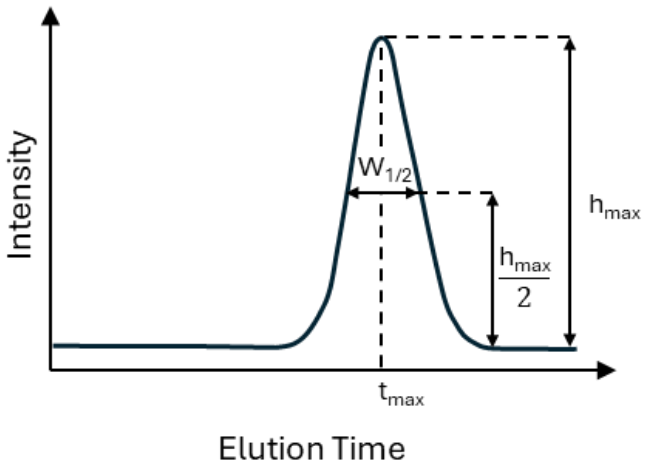


Figure 5.21: Illustrative chromatographic elution curve depicting the parameters required to calculate the column HETP

5.2.5.1.5. Determination of the Osmotic Second Virial Coefficient

In self-interaction chromatography the calculation of the osmotic second virial coefficient (B_{22}) requires an accurate measurement of the column void (dead) volume, V_0. The dead volume is primarily comprised of two regions: the interstitial space between resin beads as well as their porous volume and secondly the volume of the tubing in the system. Due to the porosity of the resin, the void volume is partially dependent on the size of the eluted mobile phase species, since larger species are not able to access as deeply into the porous network. Therefore, it is inaccurate to assume a dead volume equal to the retention volume of a small non-interacting species, such as acetone [212]. Consequently, end-capped, self-interaction chromatography resins without peptide functionalisation (as described in Section 5.2.3.2) were similarly packed into columns to form peptide-free "dead" columns. The dead column retention volumes of acetone (V'_a) and the peptide of interest (V'_p) were measured under a range of conditions. Similarly, the retention volume of acetone through the peptide immobilised column (V_a) was measured under the same conditions. These values in combination with the volume of immobilised peptide (V_i), calculated with the species radius and assuming a spherical species, were used to estimate the peptide immobilised column peptide dead volume [212]:

$$V_0 = V_a\left(\frac{V'_p}{V'_a}\right) - V_i \qquad \textit{Equation 5.6}$$

Using this estimated void volume and the elution volume of the peptide through the peptide functionalised column (V_p), an apparent retention factor (k') was calculated [212]:

$$k' = \frac{V_p - V_0}{V_0} \qquad \textit{Equation 5.7}$$

To establish, the apparent osmotic second virial coefficient, from the apparent retention factor, a number of other parameters are required. Firstly, the peptide volume of exclusion from the Van der Waals hard sphere contribution term (B_{HS}), similar to V_i was calculated assuming a spherical species, but using twice the peptide hydrodynamic radius. Additionally, the immobilised peptide surface density (ρ_s) was calculated from the concentration of immobilised peptide per resin volume (C_i), the porosity of the resin beads (P), which in the case of the AF-tresyl-650M is 0.811, and the phase ratio (ϕ_{1r}):

$$\rho_s = \frac{C_i}{P\phi_{1r}} \qquad \textit{Equation 5.8}$$

The resin phase ratio data was calculated, assuming a monomeric peptide species, from inverse size exclusion data for dextran molecules of various sizes through the same resin HW-65 resin, obtained by DePhillips and Lenhoff [214]. An effective phase ratio (ϕ_{3r}), assuming the accessible surface area for a species with triple the hydrodynamic radius of the peptide (or three coeluting peptide monomers), was likewise calculated from the same data set. Therefore, an apparent second virial coefficient may be calculated as follows:

$${B_{22}}^{app} = \frac{N_A}{{M_w}^2}\left(B_{HS} - \frac{k'}{\rho_s\phi}\right) \qquad \textit{Equation 5.9}$$

where N_A is Avogadro's constant and M_w is the molecular weight of the peptide of interest [212].

5.2.5.2. Dynamic Light Scattering

5.2.5.2.1. Peptide Sample Preparation

Peptide samples were prepared in triplicate by reconstituting the lyophilized powder with 20 mM sodium phosphate at the desired pH at 2.5 mg/mL and gently mixing until the peptide fully dissolved to leave a transparent solution before being filtered through low protein binding, 0.22 µm cellulose acetate luer syringe filters. The concentration was then verified against a calibration curve via UV-Vis spectroscopy using a NanoPhotometer from Implen, before the high concentration sample was stored at 4 °C prior to use for a maximum period of 30 minutes. 0.5 mg/mL samples were then prepared immediately prior to analysis by diluting samples with a combination of the same pH 20 mM sodium phosphate buffer containing either 0 mM NaCl or 500 mM NaCl both stored at 37 °C to achieve the desired concentration of peptide and NaCl.

5.2.5.2.2. Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) was carried out using a Zetasizer Nano ZS (Malvern Instruments, Worcestershire, UK), with the temperature controlled to 37 °C +/- 0.1 °C with a Peltier temperature controller. The system was equipped with a 4 mW, 632.8 nm He-Ne laser and measurements were made using a 173° scattering angle. 50 µL of 0.5 mg/mL peptide solution at the desired buffer conditions was prepared in 1.5 mL Eppendorf tubes as described in Section 5.2.5.2.1, before being immediately transferred via micropipette to a UV-transparent disposable cuvette (ZEN0040, Malvern Instruments, Worcestershire, UK) for analysis. The cuvette was delicately tapped to remove any bubbles formed before being wiped with ethanol to remove any surface dust. The sample was equilibrated in the system for 2 minutes before 3 measurements were taken comprising of a minimum of 12 runs to ensure convergence. Each condition was repeated in triplicate. The data was analysed using Zetasizer 7.12 software using a 0.6 nm cutoff.

5.2.5.3. Size-Exclusion Chromatography (SEC)

5.2.5.3.1. Peptide Sample Preparation

The peptide samples were prepared in triplicate in the same manner as for the DLS experiments as outlined in Section 5.2.5.2.1.

5.2.5.3.2. Setup and Procedure

Size-Exclusion Chromatography (SEC) was performed for the same sample conditions as DLS and SIC. A TSKgel G3000SWXL column was used to study both insulin and liraglutide aggregation. The methodology and setups are described henceforth. An SWXL guard column and a TSKgel G3000SWXL column (Tosoh Bioscience, Reading, United Kingdom) in series were used with a mobile phase consisting of 50 mM sodium phosphate and 150 mM NaCl at pH 7.0. The mobile phase flowrate was set to 1.0 mL/min, and the injection volumes were 10 µL. This column was used in the Nexera Series HPLC system, temperature controlled to 25 °C, with the pre-column UV detector absorbance wavelengths set to 220 nm and 280 nm. The column was first calibrated with a protein standard mixture from Tosoh Bioscience [354] to establish the relationship between sample molecular weight and retention time (see Figure 12.5 in Supplementary Materials). 0.5 mg/mL peptide samples were prepared in triplicate as described in Section 5.2.5.2.1 and t=0 measurements were taken directly thereafter. The temperature of the sample compartment in the SIL-40C autosampler was set to 37 °C ±0.5 °C to store the samples until the next injection timepoint.

5.2.6. Proteolytic Reactions

5.2.6.1. Immobilised Pepsin Reactions

5.2.6.1.1. Sample Preparation

Prior to injection, the sample peptides were reconstituted in the same manner as in preparation for self-interaction chromatography samples as outlined in 5.2.5.1.1, except using the reaction mobile phase as the solvent instead of deionised water.

5.2.6.1.2. Immobilised Pepsin Mobile Phase

Unless otherwise stated, immobilised pepsin reactions were carried out using a pH 2.0 / 0.6 M formic acid mobile phase to increase buffering capacity relative to HCl, thereby controlling the reaction pH. This has the effect of reducing the presence of any additional Cl^- adduct peaks to simplify liquid chromatography mass-spectrometry (LC-MS) analysis. For reactions at other acidities 50 mM ammonium formate was used as the mobile phase.

5.2.6.1.3. Setup and Procedure

Immobilised enzyme reactions were conducted using the Nexera Series HPLC from Shimadzu with both the peptide samples and the immobilised enzyme columns incubated at 37 °C. Unless otherwise stated, 10 µL of the peptides, dissolved at 0.1 mg/mL in the same solvent as the mobile phase, were injected in triplicate and eluted through the column for the reaction at a flowrate of 0.25 mL/min. The reacted peptides were fractionally collected using the FRC-10A fractional collector into 3 mL vials controlled by detection of a 220 nm elution peak as illustrated in Figure 5.22. The maximum vial collection volume was controlled to 0.5 mL with multiple vials being used in the case of broader peaks or in the event of multiple peaks. The mobile phase selected for each proteolytic column (as outlined in section 5.2.6.1.2) were selected to ensure compatibility with mass spectrometry.

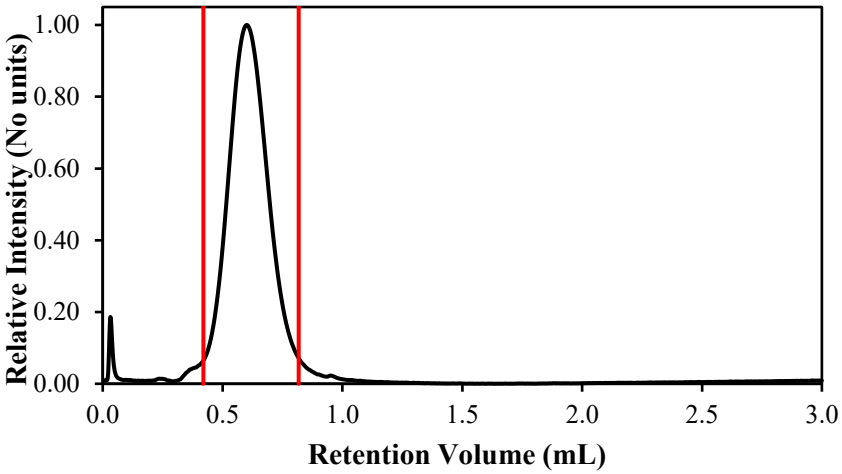


Figure 5.22: Example peptide elution profile through a packed immobilised enzyme column *with red lines highlighting the beginning and end of the sample fractional collection (10 µL semaglutide injection into an immobilised pepsin column at 0.1 mL/min and pH 2.0*

5.2.6.1.4. Column Ageing Protocol

For each study, the immobilised pepsin reactors were used for 60 cycles before being emptied and repacked with resin stored at 4 °C in pH 4.4 50 mM sodium citrate post-immobilisation reaction to minimise the effects of activity loss after repeated use. This number was selected as it allowed for 20 triplicate injections and so minimised enzyme and resin waste. However, in the case of the column ageing studies, each column underwent 57 injection cycles per measurement batch, before being washed in pH 4.4 50 mM sodium citrate and stored at 4 °C prior to its next use. The resin was not changed between each use as it was intended to observe the behaviour of the column as if for consistent repeated use similar to that a commercial column may undergo. Reactions were carried out in batches of 57 cycles, beginning immediately after immobilisation, then repeating 57 cycles each week for 4 weeks (week 1 – week 4), before repeating twice more after intervals of 4 weeks (week 8 and week 12). This resulted in the immobilised enzymes performing 342 injection cycles by the end of week 12.

5.2.6.2. *Pepsin Solution Reactions*

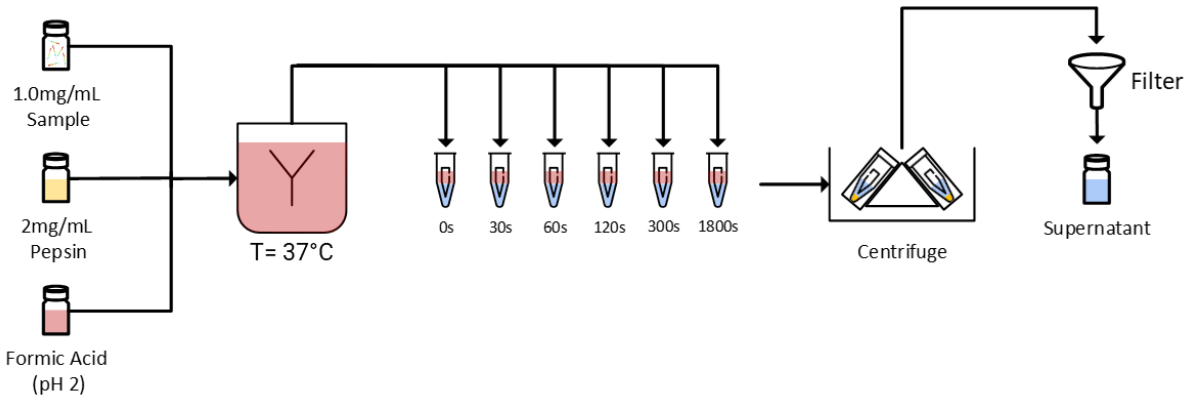


Figure 5.23: Illustration of the pepsin solution reaction workflow

5.2.6.2.1. *Reaction Procedure*

All solution reactions were carried out in triplicate in an Eppendorf ThermoMixer with the solution maintained at 37 °C whilst being agitated at 1000 rpm and each reaction was carried out in triplicate. Stock solutions of 1.0 mg/mL peptide and 2.0 mg/mL enzyme in the relevant buffer solution for the condition tested were prepared and warmed to 37 °C immediately prior to reaction preparation and initiation. All reactions were conducted in pH 2.0 HCl unless otherwise stated. Enough peptide stock solution was added to a 1.5 mL Eppendorf such that it would produce the desired initial concentration at a 500 µL total reaction volume before additional buffer was added to make up the volume to 375 µL. If the concentration of peptide is not stated it was prepared to 100 µM. 7.5 µL of enzyme stock solution was then added to and mixed with 70 µL of the relevant reaction stopping solution before 22.5 µL of the prepared 375 µL initial peptide concentration solution. This constituted the t=0 reaction timepoint to ensure that no reaction had occurred. 117.5 µL of enzyme stock solution was then added to the remaining peptide solution to initiate the reaction and giving an enzyme concentration of 0.5 mg/mL. 30 µL samples of the reaction mixture were added to and mixed in 1.5 mL Eppendorfs containing 70 µL of reaction stopping mixture at 30 s, 60 s, 120 s, 300 s and 1800 s timepoints. For samples containing no substrate peptide or enzyme, only t=0 s and t=1800 s timepoints were sampled.

5.2.6.2.2. Enzyme Separation

After completion of the reaction, the enzyme was precipitated from the sampled timepoint solution mixtures with 400 µL of -20 °C ethanol to produce a final solution of 80% ethanol. The precipitated solutions were then centrifuged in a Heraeus Biofuge Pico Microcentrifuge from Thermo Scientific at 10000 rpm. The supernatant was then filtered through the 0.22 µm Polytetrafluorethylene (PTFE) syringe filters prior to LC-MS analysis to prevent the solution from resulting in any column blocking and reduce the number of background enzyme peaks.

5.2.6.3. Liquid Chromatography – Mass Spectrometry (LC-MS)

The setup included a Dionex Ultimate 3000 HPLC system from ThermoFisher Scientific with 20 mm L x 0.075 mm ID and 150 mm x 0.075 mm ID PepMap 100 C18 trap and analytical columns packed with 3 µm particles to concentrate and separate the sample, respectively. Sample elution was initially confirmed via UV-Vis spectroscopy using an Ultimate 3000 variable wavelength detector (VWD-3400RS) configured with an in-line post-analytical-column. The detector has four measurement channels, two in the UV and two in the visible range, which were set to 200 nm and 280 nm, and 540 nm and 700 nm wavelengths, respectively. Mass spectrometry of the separated sample was conducted in sequence using an Orbitrap Velos Pro from ThermoFisher Scientific. The reversed-phase liquid chromatography separation was conducted via 50-minute method with 5 primary stages. Initially during sample injection, the sample was loaded into the trap column by the loading pump at 5 µL/min in a 3% acetonitrile mobile phase with 0.1% formic acid modifier for 5 minutes. The line is then switched such that the nanocapillary pump flows at 0.5 µL/min through both the trap and separation column in sequence to the mass spectrometer. Simultaneously, the mobile phase gradient change from 3% to 80% acetonitrile over 25 minutes is initiated. Once completed the mobile phase is held isocratically at 80% acetonitrile for 10 minutes, before returning to 3% in 3 minutes and maintained at that concentration until the end of the run to re-equilibrate the system. All samples were ensured to be <1.0 mg/mL and initially, 1 µL of analyte sample was injected into

the system for analysis. If no peak was observed, then the injection volume was increased progressively to the minimum at which clearly discernible peaks were observed.

5.3. Error Analysis

5.3.1. UV-Vis spectroscopy

Since 5 UV-vis measurements were made for each individual sample, the error of an individual UV-Vis measurement was calculated as the standard error ($\sigma_{\bar{x}}$) using the sample standard deviation of the individual sample (s) and the number of measurements (n). First the sample standard deviation was calculated from the individual measurement values (x), their mean ($\bar{x}$) and the number of measurement points:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{n-1}} \qquad \textit{Equation 5.10}$$

From this, value the standard error for the individual measurement point was calculated as follows:

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}} \qquad \textit{Equation 5.11}$$

Since the sample conditions were performed in triplicate the mean (μ) of the measurement mean values was reported as final value, whilst the reported error ($\overline{\sigma_{\bar{x}}}$) was that of the propagated error, which was calculated as follows:

$$\overline{\sigma_{\bar{x}}} = \frac{\sqrt{\left(\sigma_{\bar{x}_1}\right)^2 + \left(\sigma_{\bar{x}_2}\right)^2 + \left(\sigma_{\bar{x}_3}\right)^2}}{3} \qquad \textit{Equation 5.12}$$

5.3.2. Self-interaction chromatography

The error in the final second virial coefficient was a combination of several propagated errors. Firstly, the error in the volume of immobilised peptide (V_i) in Equation 5.6 was calculated from the standard error of the initial ($\Delta\overline{C_i}$) and final concentration ($\Delta\overline{C_f}$) measurements of the peptide

in the immobilisation solution, which were measured by UV-Vis spectroscopy as described in Section 5.3.1 in Equation 5.11, via the following equation:

$$\Delta\overline{V_l} = \overline{V_l}\sqrt{\left(\frac{\Delta\overline{C_l}}{\overline{C_l}}\right)^2 + \left(\frac{\Delta\overline{C_f}}{\overline{C_f}}\right)^2} \qquad \textit{Equation 5.13}$$

Since the measured retention volumes of acetone and the peptide of interest through the dead column, and of acetone through the peptide functionalised column were measured in triplicate, the standard errors of their retention volumes ($\Delta\overline{V_a'}$, $\Delta\overline{V_p'}$ and $\Delta\overline{V_a}$, respectively) were calculated as in Equation 5.11. These values were then used in addition to the error in the immobilised peptide volume ($\Delta\overline{V_l}$) to calculate the error in the estimated void volume ($\Delta\overline{V_0}$) via the following equation:

$$\Delta\overline{V_0} = \overline{V_0}\sqrt{\left(\frac{\Delta\overline{V_l}}{\overline{V_l}}\right)^2 + \left(\frac{\Delta\overline{V_a'}}{\overline{V_a'}}\right)^2 + \left(\frac{\Delta\overline{V_p'}}{\overline{V_p'}}\right)^2 + \left(\frac{\Delta\overline{V_a}}{\overline{V_a}}\right)^2} \qquad \textit{Equation 5.14}$$

Similarly, the standard error of the retention volumes of the peptide through the peptide functionalised column ($\Delta\overline{V_p}$) were calculated from the triplicate measurements as outlined in Equation 5.11. This error value was propagated with the approximated error in the estimated void volume ($\Delta\overline{V_0}$) into the error of the mean of the apparent retention factor ($\Delta\overline{k'}$) as in Equation 5.7:

$$\Delta\overline{k'} = \overline{k'}\sqrt{\left(\frac{\Delta\overline{V_p}}{\overline{V_p}}\right)^2 + \left(\frac{\Delta\overline{V_0'}}{\overline{V_0'}}\right)^2} \qquad \textit{Equation 5.15}$$

The reported errors in the second virial coefficient values were estimated based on the propagated errors from the terms in Equation 5.9. Of these values, Avogadro's constant (N_A) and the peptide molecular weight (M_w) were assumed to be constants and so had no error, whilst the volume of exclusion of the peptide (B_{HS}) and the phase ratio (φ) were based on literature values and so their errors were not included. Therefore, only the error in the apparent retention factor ($\Delta\overline{k'}$), as calculated in Equation 5.15, and the error in the immobilised peptide surface density (ρ)

were required. The latter of these was deemed to be equivalent in terms of percentage error to that of the immobilised volume of the peptide ($\Delta\overline{V_i}$). This was because the surface density (ρ) was calculated as outlined in Equation 5.8, in which the phase ratio (φ_{1r}) and the porosity (P) were literature values and the concentration immobilised (C_i) was directly proportional to the volume immobilised (V_i). Therefore, the reported error in the apparent second virial coefficient ($\Delta\overline{B_{22}^{app}}$) was calculated as follows:

$$\Delta\overline{B_{22}^{app}} = \overline{B_{22}^{app}}\sqrt{\left(\frac{\Delta\overline{V_i}}{\overline{V_i}}\right)^2 + \left(\frac{\Delta\overline{k'}}{\overline{k'}}\right)^2} \qquad \textit{Equation 5.16}$$

Since the second virial coefficient values were reported as surface plots, the error reported in the figure caption is that of the highest percentage error recorded for a specific data point in the individual figure.

5.3.3. Dynamic light scattering

Each individual sample was measured 3 times with each measurement consisting of 12 runs. Unless otherwise stated, the Z-average calculated by the Zetasizer 7.12 software was taken to be the hydrodynamic radius (R_h) for the individual measurement. The individual sample R_h was calculated as the mean of the three measurements, whilst the error sample standard deviation and standard errors of the individual measurements were calculated in the same way as the UV-vis measurements as in Equation 5.10 and Equation 5.11, respectively. Since each sample condition was measured in triplicate, the final reported error was calculated similarly as in Equation 5.12. Since surface plots were used to report the data, the maximum error value was reported in the figure caption.

5.3.4. Size-exclusion chromatography

The mean retention volumes of each triplicate sample were the values used to calculate the reported molecular weight estimates. The error was the standard error and calculated from the sample standard deviation as in Equation 5.11 and Equation 5.10, respectively. The standard

error of the measured retention volume (R_v) was scaled proportionally to the ratio of the estimated molecular weight (M_w) and the mean retention volume of the triplicate samples, as follows:

$$\Delta\overline{M_w} = \Delta\overline{R_v}\frac{\overline{M_w}}{\overline{R_v}} \qquad \textit{Equation 5.17}$$

Since the estimated species molecular weights were reported as surface plots, the maximum error for an individual data point within the specific plot was reported in the figure caption.

5.3.5. Proteolytic Reaction Rates

The error in the calculated Michaelis-Menten constants and maximum rate of reactions were calculated via linear regression from the Lineweaver-Burk plots using the least squares method via the LINEST function in Microsoft Excel. The results provided the standard error in the slope (m) and the y-intercept (b) of the Lineweaver-Burk plot. Since the y-intercept is equal to the reciprocal of the maximum reaction velocity ($b = 1/V_{max}$) and the slope to the Michaelis-Menten constant divided by maximum reaction velocity ($m = K_m/V_{max}$), their errors could be calculated from these values. First the error in V_{max} was easily calculated from the error in b:

$$\Delta V_{max} = \frac{\Delta b}{b^2} \qquad \textit{Equation 5.18}$$

Then the error in K_m was calculated using the error in the slope (m) and the error in V_{max} being propagated forward:

$$\Delta K_m = K_m\sqrt{\left(\frac{\Delta V_{max}}{V_{max}}\right)^2 + \left(\frac{\Delta m}{m}\right)^2} \qquad \textit{Equation 5.19}$$

CHAPTER 6

6. SELF-INTERACTION CHROMATOGRAPHY – STATIONARY PHASE ANALYSIS

6.1. Introduction

Self-Interaction Chromatography (SIC) was first applied as a technique for studying protein-protein interactions [355] before being adapted as a method for determining the second virial coefficient (B_{22}) of lysozyme and chymotrypsinogen [212]. The approach was typically employed as an alternative strategy to static light scattering, Dynamic Light Scattering (DLS) and Size Exclusion Chromatography (SEC) for estimating solubility and aggregation propensity of model proteins via measurement of their second virial coefficients. The technique has exhibited effectiveness as a tool to study monoclonal antibodies [356, 357]. Researchers using SIC have shown that the correlation of the retention volume data to the second virial coefficient and its independence from the mobile phase flow rate were improved when taking the retention volume for the peak centre of mass rather than the peak maximum for the chromatogram (see Section 5.2.5.1.3 for Equation) [349]. The technique has also been adapted to that of cross-interaction chromatography which employs a column loaded with a relevant model protein, removing the requirement to immobilise the specific protein of interest, to study the solubility and aggregation propensity of antibodies [345, 358].

The application of SIC for the study of peptide stability and aggregation has been very limited until now, with only one study on enfuvirtide, a 36 amino acid antiretroviral peptide therapeutic having been carried out using this technique [218]. The method was shown to be effective with the second virial coefficient correlating strongly with solubility as for larger proteins, but a general trend of more negative B_{22} values was reported. This B_{22} trend was hypothesised to be on account

of the significant reduction in protein folding in the smaller peptide structure relative to previously tested proteins, thereby increasing the exposure of hydrophobic residues and so the number of sites for attractive hydrophobic interactions [218]. However, there was little discussion of the effects the increased exposure of hydrophobic residues and the peptide chain may have on resin interactions and how these may change with pH and salt concentration. These interactions were established to potentially have a larger effect in a study on SIC dead columns, in which lysozyme and an IgG1 monoclonal antibody were both shown to interact strongly with a formyl surface group methacrylate resin under more acidic conditions [351]. However, they did not exhibit this same trend for the same polymer resin with amino surface groups. The other important factor in resin selection for SIC is the kinetics of peptide/protein immobilisation, with achievable protein loading being significantly increased on a polymethacrylate resin with tresyl surface groups relative to the same resin with amino and formyl surface groups [346].

The work in this Chapter aims to analyse the effects that the choice of resin and any potential end-capping will have on the SIC retention volumes and elution profile of a lipidated peptide, semaglutide, under various solution conditions. The greater exposure of hydrophobic residues and the peptide backbone in peptides relative to proteins, due to the absence of tertiary protein folding structures, in addition to the stearic acid side chain being present in this peptide's structure, results in their being an increased surface density of potential resin-interacting sites for semaglutide that need to be accounted for.

6.2. Materials and Methods

All experiments in this Chapter were carried out using a tresyl-functionalised methcrylate resin from Tosoh Bioscience as described in Section 5.1.7. Within the same section, descriptions of the monoethanolamine, glutaraldehyde, tris(hydroxymethyl)aminomethane and sodium phosphate used are described as well as the buffer preparation procedure. Additionally, the semaglutide used was that described in Section 5.1.2.2.

Resin end-capping and column packing procedures were performed as outlined in Sections 5.2.3.2 and 5.2.4.1, respectively. Chromatographic experiments were carried out on an LC-20 high-performance liquid chromatography (HPLC) system from Shimadzu the configuration of which is described in 5.2.1. The HPLC experimental procedures were performed as outlined in Section 5.2.5.1, whilst the amount of peak-tailing was determined following the description in Section 5.2.4.3.

6.3. Results & Discussion

6.3.1. Peak Elution Analysis

To optimise any chromatographic methodology, it is desirable for the mobile phase analyte to have minimal interactions with the unfunctionalised stationary phase resin so that any differences in the retention volumes for two conditions or analytes are known to be due to differences in the specific property being measured. Tosoh Toyopearl AF-Tresyl-650M resins were selected to be studied initially, since the porosity and surface area measurements of the HW-65 polymer resin have previously been studied in detail [214] and the tresyl surface group shows the highest activity during immobilisation [346]. However, as described in Section 6.1, the interaction strength between unfunctionalised stationary phase surface and the mobile phase protein or peptide may be strong for specific peptides or proteins under specific conditions [351]. Since these interaction strengths differed for the same resins with different functional surfaces, end-capping of unreacted surface functional groups on the resins should be a means through which these interactions may be mitigated. Therefore, following the end-capping procedures discussed by Hedberg et al, four initial stationary phases were considered: two resins end-capped with tris(hydroxymethyl)aminomethane (henceforth referred to as tris) and two with monoethanolamine (MEA). For each pair, one was activated with glutaraldehyde during the immobilisation procedure and the other was not. The differently end-capped resins shall each be referred to as Tris, Glutaraldehyde-Tris, MEA and Glutaraldehyde-MEA, henceforth.

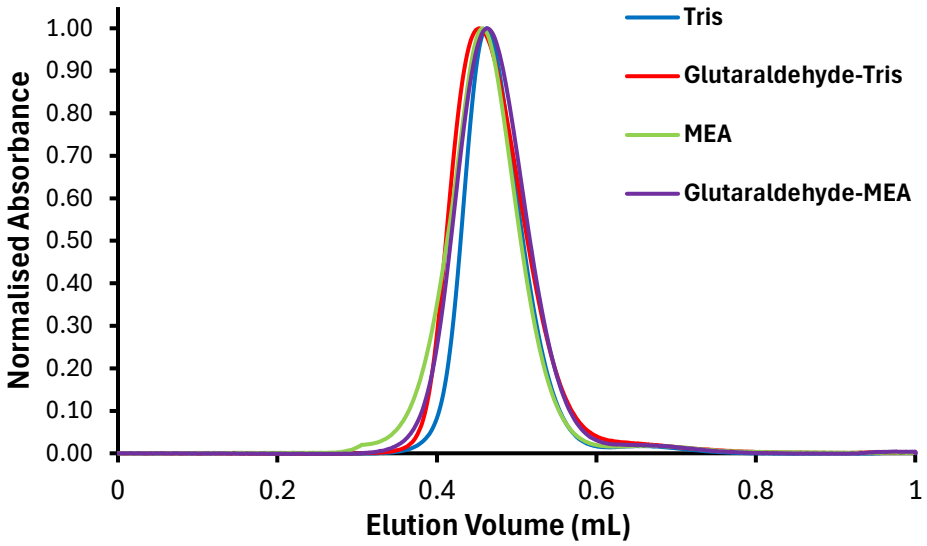


Figure 6.1: Acetone elution profile comparison for four AF-Tresyl-650M column stationary phases *(20 mM sodium phosphate Buffer pH 6.0, flowrate 0.05 mL/min at 37 °C)*

Initially, to establish the quality and consistency of the packing for each stationary phase resin, 2% v/v acetone in deionised water injections were eluted (as stated in Section 5.2.4.1) through the columns. The chromatograms for the non-interacting acetone through each column under a single condition are shown in Figure 6.1. The similarity in shape and retention volume of each peak demonstrates the consistency of the column preparation, since the retention volume for the different stationary phases should be dictated predominantly by the resin mass (which was the same for each column) and quality of packing, with other fractional differences being due to differences in equipment such as the individual columns and column tubing lines. Furthermore, the quality of packing can also be confirmed through calculation of the tailing factor (TF) for each of the acetone void volume marker elution peaks which was less than 1.2 for all four columns and so was acceptable. In theory, end-capping experiments should have a negligible effect on the stationary phase volume. However, glutaraldehyde activation could lead to a reduction of resin swelling, and so void volume, due to the increased cross-linking between surface amine groups, as previously reported by Dmitriev et al. for polyacrylamide hydrogels [359]. Since this

was not observed and due to the absence of glutaraldehyde reactive surface groups, it can be safely assumed that cross-linking reactions do not occur during resin preparation.

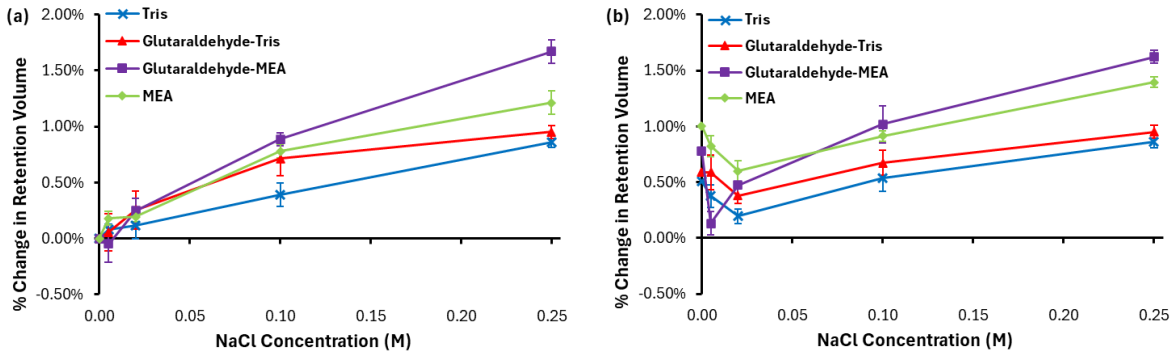


Figure 6.2: Acetone void volumes versus NaCl Concentration at (a) pH 6.0 and (b) pH 8.0 *relative to 0 mM NaCl, 20 mM sodium phosphate Buffer pH 6.0 mobile phase conditions*

After establishing the consistency in resin preparation and packing for the different columns, it was also necessary to confirm whether the stationary phase properties and packing quality would remain consistent across the range of simulated intestinal fluid (SIF) mobile phase conditions to be applied. Therefore, 2% v/v acetone injections were repeated across a range of NaCl concentrations at pH 6.0 and pH 8.0, the minimum and maximum of the range to be tested. The percentage changes in retention volumes relative to that measured at pH 6.0 and 0 mM NaCl (the condition plotted in Figure 6.1) are shown for pH 6.0 and pH 8.0 in Figure 6.2a and 2b, respectively. Overall, there was a general trend observed that the retention volume of acetone slightly increased up to nearly 2% with higher concentrations of NaCl at both pHs tested. A potential cause of this phenomenon was the reduced swelling of the HW-65 resin due to an increase in osmotic pressure of the mobile phase, thereby increasing the void volume and consequently retention volumes. This would be in agreement with findings from Horkal and Basser, who showed that the swollen structure of crosslinked polyacrylic acid gels collapsed upon exposure to increased concentrations of $CaCl_2$ [360]. Since the changes in retention volumes across all conditions were only small, minimal changes in peak tailing with salinity were observed. However, it would be expected that a larger change in resin swelling, and so porosity

and packing, would result in several, non-linear changes in elution behaviour and peak tailing. In fact, there appeared to be a greater change in peak tailing with mobile phase pH, whilst remaining independent of NaCl concentration.

Nevertheless, despite increasing to high salinities relative to those previously tested for SIFs and simulated gastric fluids (SGFs) [39], the change in retention volume remained less than 2%. Therefore, whilst it may have been necessary to factor in any potential reduction in resin swelling with increasing mobile phase salinity (which had not been previously considered in SIC) into calculations for the second virial coefficient, since it was relatively minor, it would not have a significant impact on any initial analyses. Therefore, it would not inhibit the ability of the method to be applied as a high-throughput tool for studying peptide stability under gastrointestinal conditions.

6.3.2. End-capping Comparison

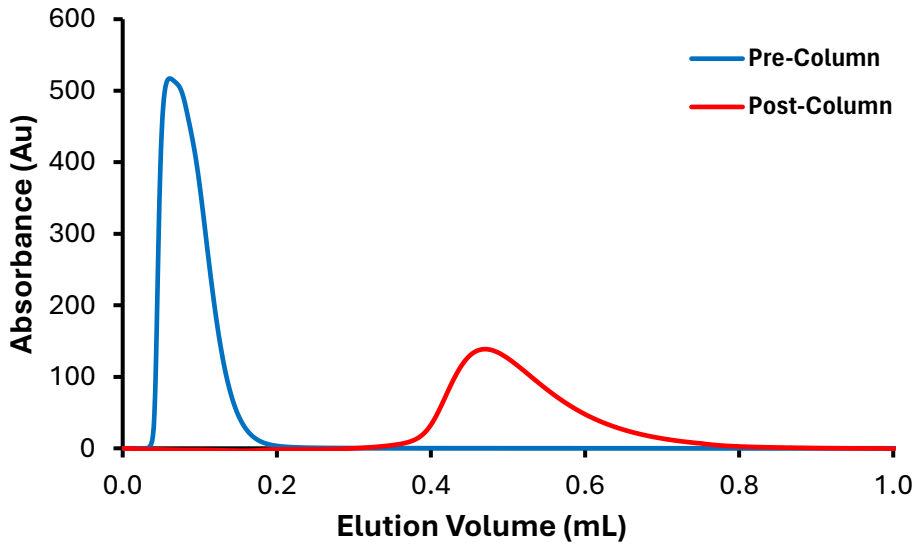


Figure 6.3: Chromatogram comparisons for 10 µL injection of semaglutide in UV-Vis detectors *before and after a Tris end-capped AF-Tresyl-650M packed column (20 mM sodium phosphate pH 8.0, 0.05 mL/min flow rate, 37 °C Temperature)*

To establish the capacity for the various end-capping strategies to function as non-interacting stationary phases for the analysis of semaglutide in SIC, a series of 10 µL injections of 0.1 mg/mL semaglutide in deionised water were performed through each column under the same conditions as the previously analysed 2% v/v acetone injections. Three properties of the elution chromatograms may then be analysed sequentially to establish whether the stationary surface properties were exhibiting suitably low interactions with the mobile phase peptide. Firstly, the ratio of the peak areas for the pre- and post-column detectors should be assessed, as discussed in Section 5.2.4.3. These values may not be exactly 1:1 due to the two UV-Vis detectors being calibrated differently, peak spreading through the column (see example in Figure 6.3) and potential sample structural changes (e.g. aggregation) during elution that may change UV-Vis absorption profile [191]. However, the pre- and post-column absorbance areas should be similar in value or at least at a relatively consistent ratio; if not, then there has either been significant irreversible retention of the sample by the column or a structural change in the peptide sample whilst in the column, resulting in reduced and increased post-column absorbance respectively. If the absorbance areas were consistent, then subsequently peak tailing should be analysed to determine the strength of any weak non-linear interactions disrupting peptide elution. Finally, if both the peak area ratios and peak tailing values were acceptable, the retention volume should be compared to that of 2% v/v acetone under the same conditions to ensure there were minimal levels of chromatographic interactions occurring between the resin and the peptide.

Upon examination of the results outlined in Table 6.1, it firstly became clear that using the ratio of UV-Vis absorption area (measured at 220 nm) was merely an indicative metric of the fraction of sample eluted and not an accurate quantitative value, since it showed relative consistency in its value across a range of salinities but varied with pH. A calibration curve was not prepared as it would have been required for each condition, potentially consuming a large amount of sample, and the analysis required merely needed to be qualitative. At pH 8.0 and for all salt concentrations, the pre- and post-column absorption ratios were consistent showing that

semaglutide was eluted through each column regardless of the end-capping. However, for the MEA column there was a slight reduction, which was perhaps due to higher peak spreading or column surface interactions causing a change in peptide structure and so UV-absorbance. However, due to the lack of significant difference in retention behaviour relative to the other columns and absence of any peptide elution peaks when re-equilibrating, it did not seem as if any peptide was being retained irreversibly for most solution conditions. However, at pH 6.0, for both stationary phases end-capped with MEA, there was significant irreversible retention of the injected peptide, such that its retention time was longer than the duration of a sample measurement, requiring several column volumes to elute the peptide. Therefore, as observed in Table 6.1, the ratio of post- to pre-column UV absorbance peak area, was a small fraction, since the majority of the sample was irreversibly retained. This retention behaviour was likely on account of the fact that pH 6.0 is relatively close to the pKa of semaglutide (pKa ≈ 5.4), suggesting that semaglutide under these conditions has a relatively neutral charge. Therefore, hydrophobic interactions would have been the dominant force at this pH, similar to aggregation-inducing hydrophobic interactions suggested by Raman et al. at acidities close to the peptide isoelectric point [160]. End-capping with MEA, relative to Tris, may have increased the hydrophobicity of the HW-65 resin surface, resulting in stronger interactions with the lipid sidechain of the peptide and so longer, less predictable retention. Therefore, as evidenced by the long retention times, this made an MEA end-capped stationary phase unsuitable for SIC of semaglutide since the neutral stationary phase interactions with the peptide may have been more dominant than the peptide-peptide interactions intended to be measured.

Table 6.1: Pre- and Post-column UV-Vis Absorption area ratios
10 µL 0.5 mg/mL semaglutide in deionised water injections for different end-capping and mobile phase conditions

pH	[NaCl] (mM)	Tris		Glutaraldehyde-Tris		Glutaraldehyde-MEA		MEA	
		Ratio	Error	Ratio	Error	Ratio	Error	Ratio	Error
6.0	0	0.84	0.03	0.91	0.08	0.05	0.03	0.03	0.02
	5	0.95	0.07	0.85	0.03	0.02	0.02	0.02	0.01
	20	0.84	0.06	0.80	0.06	0.02	0.01	0.02	0.01
	100	0.88	0.11	0.97	0.10	0.01	0.01	0.01	0.01
	250	0.83	0.08	0.81	0.06	0.21	0.10	0.03	0.02
8.0	0	1.09	0.09	1.10	0.01	1.12	0.05	0.85	0.03
	5	1.15	0.02	1.07	0.08	0.95	0.06	0.89	0.07
	20	1.08	0.06	0.99	0.05	1.01	0.05	0.82	0.04
	100	1.12	0.07	1.04	0.01	1.21	0.06	0.84	0.01
	250	1.10	0.09	1.04	0.02	1.07	0.05	0.89	0.07

Further analysis of the semaglutide elution data for the Tris functionalised columns may allow for it to be established whether glutaraldehyde activation during peptide immobilisation may alter the surface of the resin. If this was found to be the case then the resultant increase in non-peptide-peptide interactions during SIC would likely negate the benefit of the increased peptide immobilisation rate and coverage that glutaraldehyde allows for, as described by Hedberg et al. [346].

Using the chromatogram exhibited in Figure 6.3 as an example, it was observed that there was significant tailing of the semaglutide elution peak in both columns under all conditions relative to the 2% v/v acetone injections. However, as also observed in Figure 6.3, there was also relatively high tailing of the injection profile detected in the pre-column UV-Vis detector. The sample solution was prepared freshly, and the injector needle was cleaned prior to each measurement. Therefore, it was concluded that this pre-column tailing was likely caused by peptide structural transitions due to the change in solution conditions after mixing with the mobile phase, leaving multiple conformations or aggregates. This is the only suggested cause of peptide peak tailing by BenchChem that is independent of the any stationary phase effects and

as such may cause pre-column tailing [361]. As the pre-column tailing was therefore assumed to be inherent to the experimental set-up, it was factored into the post-column tailing via the ratio: $\frac{Post-column\ Tailing\ Factor}{Pre-column\ Tailing\ Factor}$. The resultant ratios for both the Tris-functionalised columns across all conditions were all consistently in the range 0.9 to 1.3, which is within the range of acceptable tailing factors for HPLC analysis.

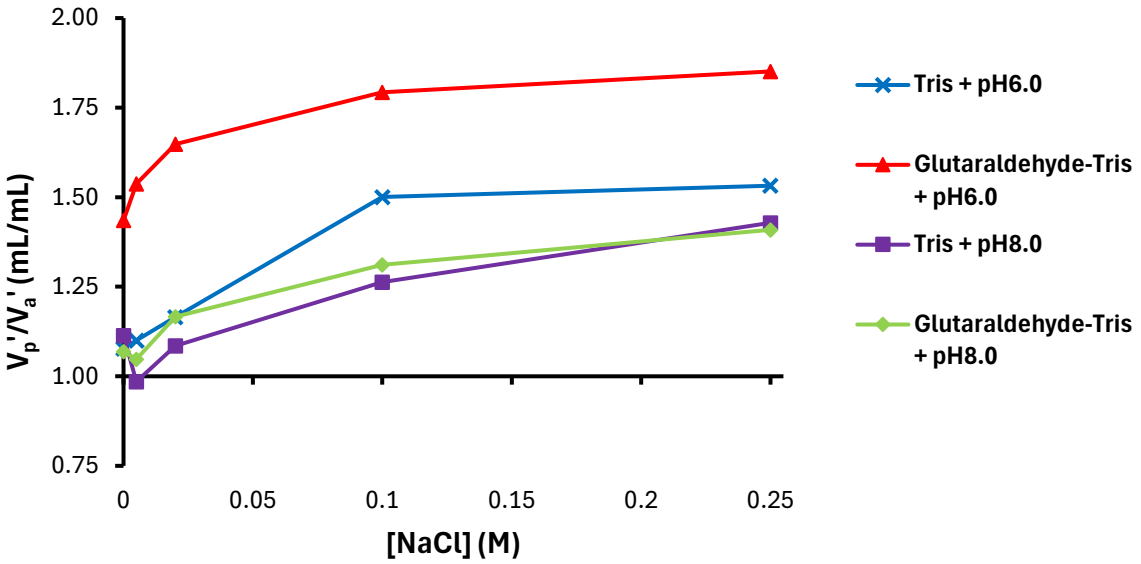


Figure 6.4: Retention volume ratios of semaglutide and acetone versus [NaCl] for tris-functionalised HW-65 AF-Tresyl-650M stationary phases at pH 6.0 and pH 8.0 mobile phases (0.05 mL/min Flowrate, 37 °C Column Oven Temperature)

Figure 6.4 shows the relationship between the concentration of NaCl and the ratio of the retention volumes for semaglutide and acetone injections at pH 6.0 and pH 8.0 for both Tris functionalised stationary phase resins tested. Under ideal conditions where there is minimal interaction between the injected peptide sample and the stationary phase, it would be expected that there would be minimal change in retention volume with changing conditions unless a significant change in sample structure occurred; the volume ratio should therefore be approximately 1.0. Furthermore, since HW-65 polymer resins are traditionally used for size exclusion chromatography, it would be expected, under ideal non-interacting conditions, that

semaglutide would exhibit a very slightly lower retention volume than acetone ($V_p'/V_a' < 1$) due to its larger molecular size and so increased size exclusion from the stationary phase porous structure. However, as exhibited in Figure 6.4, not only was the retention volume of semaglutide much greater than acetone but it was evident that there was an increase in semaglutide-stationary phase interactions with increasing salinity. This was likely on account of the higher ion concentration causing an increase in screening of charge repulsive forces between the resin and the peptide, similar to salt-induced charge screening effects causing aggregation discussed by Raman et al. [160]. This results in hydrophobic interactions with the lipid sidechain of semaglutide and the methacrylic polymer structure becoming more dominant. The trend was less pronounced at pH 8.0, at which semaglutide exhibits a higher charge and greater stability, than at pH 6.0, but there was still an approximately 20% increase in retention volume between 0 mM and 100 mM NaCl concentrations, which is far from ideal for SIC. Nevertheless, since the retention volume ratios, V_p'/V_a', were lower for the stationary phase that was not treated with glutaraldehyde, it was concluded not to incorporate glutaraldehyde activation into the peptide immobilisation procedure as described by Hedberg et al. [346]. Furthermore, unless a more appropriate resin could be found to act as the stationary phase, it would be required to factor these shifts in dead column retention volume with changing conditions into the final calculations for the second virial coefficient. This differed from the work performed by Hedberg et al., who described the use of non-interacting proteins, of a similar size to the protein of interest, as void volume markers during dead column experiments to calculate a more accurate V_p' [351]. However, in that case, sample-resin interactions were not accounted for and removed from the measured retention volumes, which were intended to be indicative of mobile phase peptide interactions with immobilised peptides. As a result, these sample-resin interactions will be factored into the final calculated values for the second virial coefficient, thereby making a correction that will give a sensible final result.

6.3.3. Flowrate studies

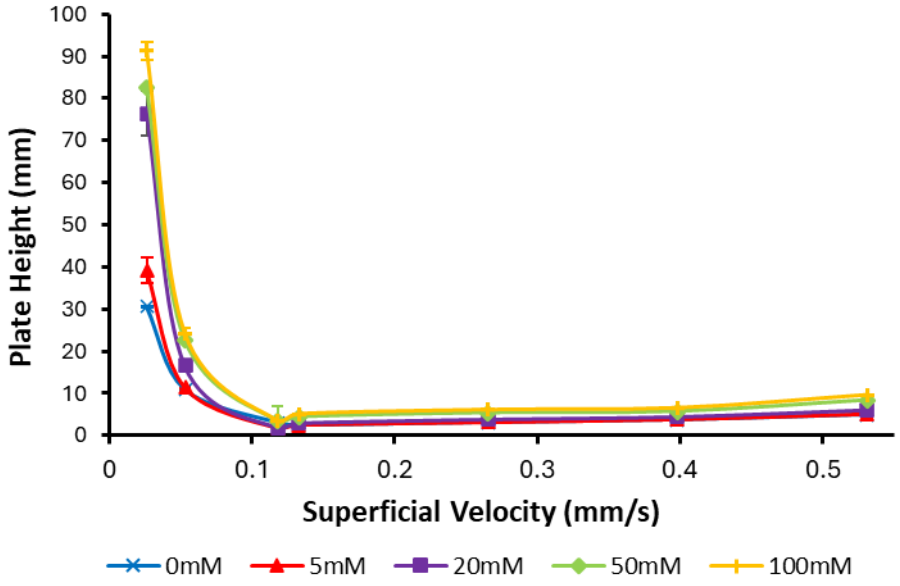


Figure 6.5: Van Deemter Plot of the variation in the efficiency of 0.1 mg/mL semaglutide elution with mobile phase velocity
through columns packed with Tris-endcapped Tosoh Toyopearl AF-Tresyl-650M resin at various concentrations of sodium chloride (pH 8.0, 37 °C)

For an ideal chromatographic setup, a low height equivalent to a theoretical plate (HETP) is desired, as, with respect to SIC in particular, it implies high column efficiency, providing sharp, resolved peaks for a shorter column length. A Van Deemter plot provides a graphical representation of the relationship between the linear velocity of the mobile phase and the theoretical plate height, thereby allowing identification of the optimum flowrate for the specific conditions tested. Typically, this experiment would be performed for SIC after the protein or peptide had been immobilised to the surface such that the flowrate would be optimised for the peptide-peptide interactions. However, in this scenario, due to the potential for size exclusion effects within the resin and surface interactions, it was considered that optimal elution of the peptide under the supposed non-interacting, “dead” column conditions would be a higher priority such that any changes in behaviour would be known to be due to peptide-peptide interactions and could therefore be analysed as such. Therefore, new columns were packed

using Tris-endcapped AF-Tresyl-650M resin and semaglutide injections were performed at various flowrates using a pH 8.0 mobile phase with various NaCl concentrations. As observed in Figure 6.5 and in agreement with data and explanations presented earlier in this Chapter, the theoretical plate height increased with sodium chloride concentration in the mobile phase suggesting poorer column efficiency. However, at superficial velocities that were sufficiently low (<0.1 mm/s) the HETP was very high due to increased longitudinal diffusion of the analyte. This effect may have been exacerbated by the resin itself, since the HW-65 hydroxylated methacrylic polymer structure has a pore size of 100 nm, if there was any aggregation of the analyte peptide then size exclusion effects may have resulted in sufficient sample hinderance, resulting in peak spreading and poorer resolution. At superficial velocities that were too high, the mass transfer effects were reduced as there was insufficient time for the mobile phase peptide to diffuse into the resin pores and interact with the stationary phase. This is in line with theory based on the Van Deemter equation, wherein at higher mobile phase flowrates, the relative effect of the mass transfer term is greater, whilst longitudinal diffusion effects tend to zero, resulting in narrow peaks with little resolution [362]. Therefore, it was ideal that a relatively shallow slope was observed here as it implied that despite semaglutide showing strong interactions with the resin relative to acetone in Sections 6.3.1 and 6.3.2, the interaction with the Tris-functionalised resin was relatively weak. Overall, of the conditions tested, a superficial velocity of approximately 0.12 mm/s was found to have the lowest HETP. This result is equivalent to a 0.05 mL/min flowrate for a 180 µL Omnifit column, which has been the typical flowrate previously applied with these columns.

6.3.4.pH variation study

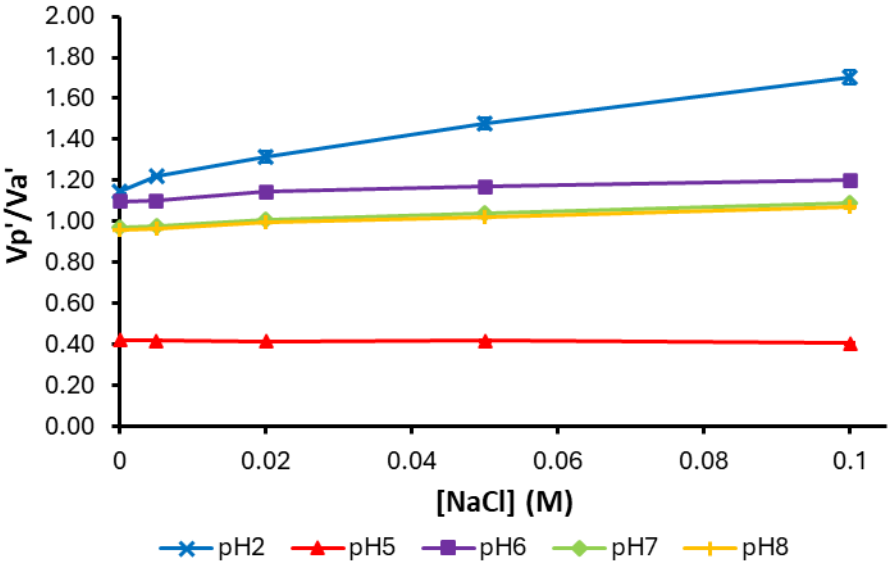


Figure 6.6: Retention volume ratios of semaglutide and acetone *versus [NaCl] with a tris-functionalised HW-65 AF-Tresyl-650M stationary phase at various mobile phase pHs (0.05 mL/min Flowrate, 37 °C Column Oven Temperature)*

To ensure that the AF-Tresyl-650M resin selected would be suitable across a wide range of simulated gastrointestinal fluids (SGIFs) for peptide elution, the previous experiments were repeated under the optimised conditions using a newly prepared and packed 180 µL column. Simple fasted and fed state SGFs were prepared using pH 2.0 HCl and 20 mM pH 5.0 sodium citrate, whilst various SIFs were prepared using pH 6.0, 7.0 and 8.0 20 mM sodium phosphate buffers.

The repeated experiments for pH 8.0 and 6.0, exhibited similar results to previous experiments, demonstrating the reproducibility of the experimental setup. The retention of semaglutide at pH 7.0 being more similar to the behaviour exhibited at pH 8.0 than at pH 6.0 indicated that the transition in the charge state of the His-1 side chain (pKa ≈ 6.0) or N-terminal of semaglutide may have been responsible for the increased retention volume. This change in charge state would result in a dense positive charge at the N-terminal, which would have strong electrostatic interactions with the electron-rich ester groups of the methacrylic polymer and any remaining,

unfunctionalised fluorine groups in the tresyl functionalisation on the resin surface. However, there likely was also a transition in peptide structure between pH 7.0 and pH 6.0, perhaps caused by the protonation of imidazole group of His-1, to a structure with more exposed hydrophobic regions. This idea is supported by findings from Zapadka et al., who previously demonstrated a pH-induced switch in aggregation behaviour in this pH range [158]. This is on account of the relative increase in retention behaviour of the peptide at higher salt concentrations being greater at pH 6.0 than at pH 7.0 and pH 8.0, displaying stronger hydrophobic interactions under the high charge shielding conditions. The trend of increasing retention behaviour of the peptide with higher mobile phase acidities continued at pH 2.0. At low salt concentrations, the peptide at pH 2.0 exhibited similar retention behaviour to that at pH 6.0, suggesting that the protonation of the Histidine group still governed the weak electrostatic interactions with the resin that increased the peptide retention volume relative to acetone. However, the effect of increasing salinity was much stronger on retention volume at pH 2.0, indicating that the protonation of the aspartate, glutamate and C-terminal groups resulted in an increase in hydrophobic structure of the peptide.

At pH 5.0 the elution behaviour was significantly different. Two peaks were observed with a 0.5 mg/mL injection concentration, with the first being the larger one reported in the data in Figure 6.6. (An example elution profile of semaglutide under pH 5.0 mobile phase conditions may be observed in Figure 12.4 in Section 12). For the larger peak the peptide was eluted much more rapidly than acetone, resulting in a significantly lower V_p'/V_a' value than for the other conditions. Furthermore, relative to the other conditions there was very little variation in the retention behaviour with salinity at pH 5.0. This was due to the low solubility of semaglutide in aqueous conditions at pH 5.0, due to the proximity to the isoelectric point of the peptide (pI ≈ 5.4) resulting in a close to net neutral charge. Therefore, upon injection the peptide rapidly aggregated and was size excluded, resulting in limited interaction with resin. Since the pore size of the resin was approximately ~100 nm, these were known to be large aggregate species (>decamer). The second smaller peak also showed little variation with increasing salt concentration, consistently

being eluted with a V_p'/V_a' of approximately one. This indicated that these species were small peptide oligomers that were also non-interacting but were not sufficiently large to be size excluded from the porous resin. Furthermore, given the pre- and post-column UV-Vis peak areas were of similar orders of magnitude, it seemed that limited semaglutide was precipitated and trapped in the column during elution. Upon reducing the injection concentration to 0.1 mg/mL it was the second of the two peaks (the oligomer peak) that was observed to be more dominant.

6.4. Conclusions

In this Chapter a number of simple experiments were performed to establish the protocol and end-capping strategy for the stationary phase resin in the application of peptide SIC. Overall, compared to other reported SIC studies using monoclonal antibodies, peptides exhibit a much more complex solution behaviour in SIC type experiments. The results indicated that fixating the AF-Tresyl-650M resin with ethanolamine resulted in significantly increased peptide retention at mildly acidic conditions. In addition to this, the use of glutaraldehyde to activate the resin surface during immobilisation was found to increase the retention behaviour of semaglutide with increasing salt concentration. As a result, a Tris-functionalised AF-Tresyl-650M resin without glutaraldehyde activation would be used for future experiments.

Flowrate studies were conducted to establish the optimum conditions for the peptide SIC study and ideal elution profiles. Van Deemter plots were created for various sodium chloride concentrations, with a superficial velocity of approximately 0.12 mm/s being found to provide the lowest HETP and best column efficiency.

Experiments were conducted on the selected resin to analyse the effects of the mobile phase on dead column elution of the model peptide. Increasing salt concentrations resulted in poorer peptide elution profiles and larger retention volumes. At pH ≤ 6.0, there was a greater sensitivity to salt concentration in the peptide retention volumes, whilst pH 7.0-8.0 showed the most ideal

behaviour. However, at pH 5.0 solubility of the peptide was found to be a limiting factor in the experimental setup with a significantly reduced peptide concentration being required.

This work was important for future peptide SIC experiments as it firstly ensured that the HW-65 resin that has been studied thoroughly within the context of SIC was applicable for the use of peptide elution under varying SIF conditions. Furthermore, the resin end-capping studied verified that capping the tresyl groups with tris and not using glutaraldehyde to activate the resin surface would be the optimal immobilisation procedure to minimise peptide-resin interactions. Additional studies allowed for the selection of the optimal flowrate for peptide elution in future experiments and a good understanding of how the conditions affect the peptide elution behaviour and how that should be factored into the results from peptide immobilised columns.

The level of complexity observed in these SIC peptide experiments are substantial and unprecedented. It is not year clear whether the benefits of obtaining SIC data for large peptides will outweigh the complexity of extracting understanding from these complex data sets.

CHAPTER 7

7. SELF-INTERACTION CHROMATOGRAPHY OF A MODEL PEPTIDE - SEMAGLUTIDE

7.1. Introduction

Of the several barriers to achieving a high bioavailability for orally administered peptide therapeutics, one of the lesser considered is the impact the gastrointestinal (GI) conditions have on the peptide's physical stability. Throughout the GI tract the local environmental conditions are continuously changing from the highly acidic gastric (pH 1.0-2.0) to the more neutral (pH 7.0-7.5) ileal conditions [363]. In addition, they are influenced by several factors including age, gender, health conditions (e.g. Crohn's disease, Intestinal Bowel Disease etc) and even the recency of the individual's last meal [51]. This complex changing chemical environment increases the likelihood that for orally administered peptide therapeutics aggregation will occur. The low gastric pH increases the probability of peptide unfolding or undergoing conformational changes, while the pH will likely reach the isoelectric point of the peptide as it passes through the GI tract [364]. Typically, the variation in GI conditions is accounted for in oral drug formulations by targeting intestinal delivery through the use of enteric coatings. These are comprised of polymeric materials which maintain their structure in the low pH gastric environment, but dissolve upon exposure to higher intestinal pHs, thereby releasing the active pharmaceutical ingredient (API) [365]. Nevertheless, without additional formulation additives or encapsulation strategies, enteric coatings do not prevent peptide aggregation at the delivery site.

The drawbacks to peptide physical instability are threefold: the peptide aggregate may have an increased risk to induce an immunogenic response [366]; if the conformational change is irreversible there is a loss of API activity [155], whilst in accordance with Lipinski's rule of 5, an

increase in molecular weight in the form of aggregate will result in a reduction in permeability across the intestinal cell membrane [13]. Therefore, all of these three factors combine to decreased overall bioavailability of the therapeutic peptide, thereby requiring higher doses and so a higher cost to the end consumer. The results of reduced bioavailability are exemplified by the difference in dosage of the subcutaneous (Ozempic®) and oral (Rybelsus®) formulations of semaglutide, which are 0.25-2.0 mg weekly and 3-14 mg daily, respectively. Despite its low oral bioavailability of only 0.4%-1% [367], semaglutide is one of fewer than 15 US Food and Drug Administration (FDA) approved peptide drugs to be administered orally [23]. The formulation of Rybelsus® targets gastric absorption of the semaglutide API and employs sodium N-[8-(2-hydroxybenzoyl)amino]-caprylate (SNAC) as a transcellular permeation enhancer [68]. However, SNAC is suggested to additionally function by stabilising the semaglutide from aggregation by increasing the local pH and shifting the proximal environment polarity thereby weakening hydrophobic interactions between monomer species [68].

Several techniques may be applied to analyse peptide physical stability, but there are drawbacks to each. Several methods are non-predicative, in that they measure the current size of the peptide therapeutic and do not provide an indication of the long-term stability. This includes techniques such as size exclusion chromatography (SEC) and dynamic (DLS) and static (SLS) light scattering. Therefore, unless accelerated stability testing methods, such as elevated temperatures are used, these techniques require measurements at several time points over the period of aggregation to establish the point of aggregation initiation. For example, this time period is roughly 20 days in the case of 20 μM semaglutide in pH 8.0 phosphate buffer [320]. Conversely, circular dichroism (CD) may provide more information concerning the peptide conformation and so its long-term stability. However, whilst CD, along with SLS and DLS, have low sample requirements in their microplate reader configurations, they typically have a minimum concentration threshold in the range of 0.1-1.0 mg/mL. This concentration is higher than what may be desirable, when studying drug release within the GI tract.

Therefore, there is potential need for a technique which can predict long-term stability and can operate at low sample concentrations. Self-interaction chromatography (SIC) fulfils this niche requirement, wherein measurements of the second osmotic virial coefficient, a direct measurement of the strength of the interparticle forces, allow for the prediction of the peptide's long-term propensity to aggregate. Furthermore, it provides the increased similarity to *in-vivo* conditions, since the sample measurement is made immediately upon injection into a mobile phase comprised of a simulated intestinal or gastric fluid.

As mentioned in the previous Chapter, the use of SIC to measure the second osmotic virial coefficient of a peptide has so far been limited to one study on enfuvirtide [218]. Therefore, an initial intention of this investigation was to establish the feasibility of SIC to function as a technique to study peptide physical stability under a range of simulated gastrointestinal conditions, before establishing an operating window within which to operate. In the previous Chapter, the effects of resin end-capping were established, with the end-capping and SIC experimental procedures being optimised. The aim of the work carried out in this Chapter was to build upon work previously performed using SIC, such as that by Hedberg et al. and Tessier et al. and establish whether similar SIC procedures were appropriate for peptides using semaglutide as a model species. The work would be corroborated and validated by comparison of DLS measurements of the peptide under similar solution conditions.

7.2. Materials & Methods

All experiments in this Chapter were carried out using a tresyl-functionalised methcrylate resin from Tosoh Bioscience as described in Section 5.1.7. Within the same section, descriptions of the Tris(hydroxymethyl)aminomethane and sodium phosphate used are described as well as the buffer preparation procedure. Additionally, the semaglutide used was that described in Section 5.1.2.2.

Peptide immobilisation and resin end-capping procedures were performed as outlined in Sections 5.2.3.1 and 5.2.3.2. Raman Spectroscopy of the resin surface was carried out as in 5.2.3.5. The column packing procedures are outlined in Section 5.2.4.1. Chromatographic experiments were carried out on a LC-20 high-performance liquid chromatography (HPLC) system from Shimadzu, the configuration of which is described in 5.2.1. The self-interaction chromatography experimental procedures were performed as outlined in Section 5.2.5.1, whilst the amount of peak-tailing was determined following the description in Section 5.2.4.3. The Dynamic Light Scattering procedure is also outlined in Section 5.2.5.2.

7.3. Results

7.3.1. Semaglutide Stability

7.3.1.1. Dynamic Light Scattering

Initial stability studies were conducted using 0.5 mg/mL semaglutide solutions as the sample peptide. Its propensity to aggregate was initially assessed across a range of simulated gastric and intestinal conditions using Dynamic Light Scattering. These results could be correlated to future SIC studies and would be used as an accurate indication of the solubility of the peptide and the size of any aggregate structures formed, allowing for suitable selection of conditions. Simulated gastric conditions were prepared using hydrochloric acid and 20 mM sodium citrate buffers to represent the fasted (pH <3.0) and fed states (pH 3.0-5.0) respectively. Simulated intestinal conditions (pH 6.0-8.0) were prepared using 20 mM sodium phosphate, whilst under both simulated environments the salinity of the conditions was varied using NaCl.

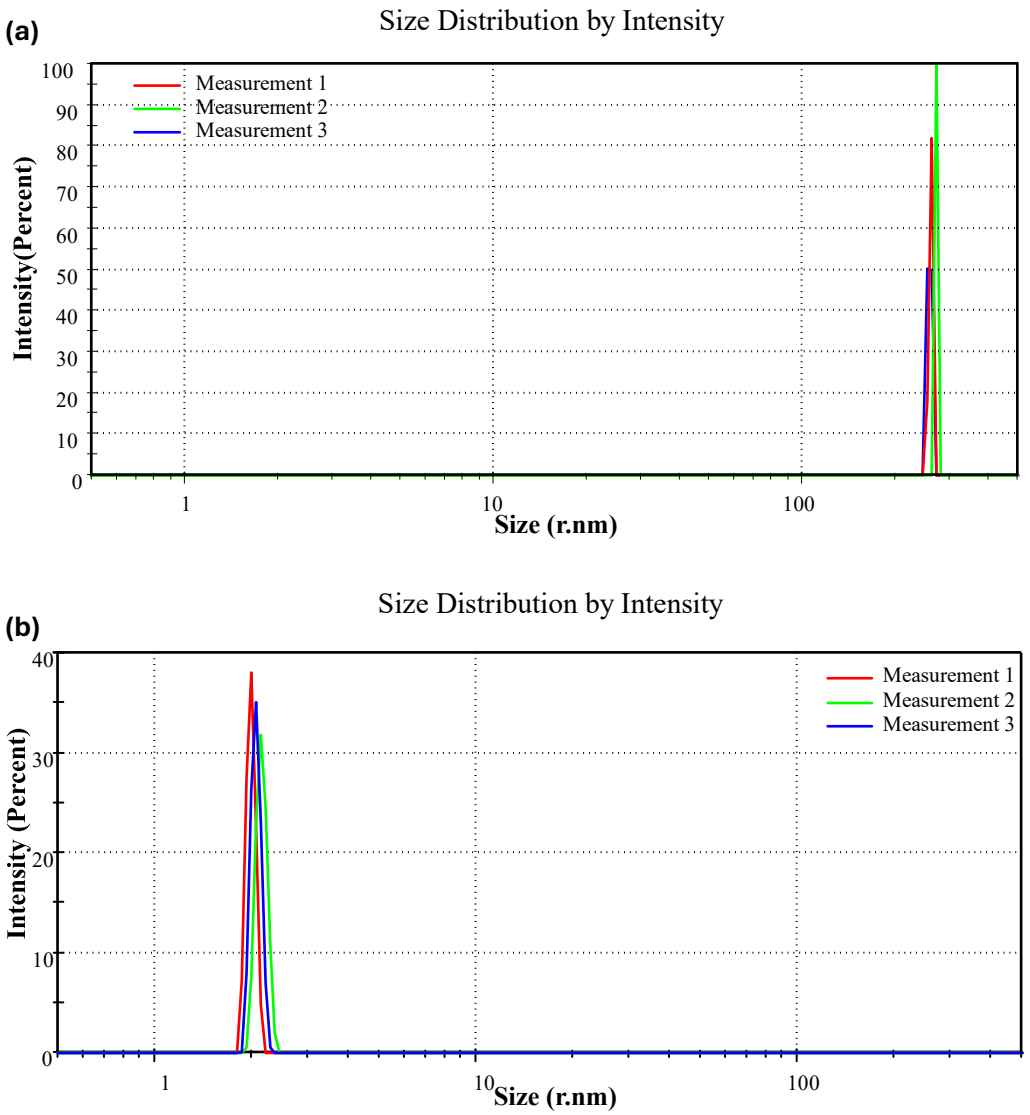


Figure 7.1: Dynamic Light Scattering radius size distributions for 0.5 mg/mL semaglutide *in 20 mM sodium buffer with 0 mM sodium chloride at (a) pH 4.0 citrate and (b) pH 8.0 phosphate (see Figure 12.6 in Supplementary Materials for correlograms)*

Under gastric conditions between pH 3.0 and pH 5.0, semaglutide initially exhibited poor solubility requiring significant gentle mixing and a settling time to ensure complete dissolution of the peptide sample as well as for any bubbles formed to dissipate (see Figure 7.1a for example DLS measurements at pH 4.0 juxtaposed with Figure 7.1b at pH 8.0 to highlight the approximate species size that indicates stability). Even then, there was a high degree of variation in the DLS measurements with an error of ±10%, due to the presence of large poorly soluble aggregates with R_h values of between 100 nm and 500 nm being observed (see Figure 7.2). Under these

conditions, the size of the aggregates showed little correlation with the salt concentration. However, as the pH increased and approached the isoelectric point (pI~5.4), the size of the aggregates increased, as would be expected due to the lower charge state of the peptide and so reduced electrostatic repulsions maintaining monomeric stability. Conversely, at pH <3.0, semaglutide showed much greater solubility and the radius of the species formed in solution correlated more strongly with the salt concentration of the environment. For example, at pH 2.0 and 0 mM sodium chloride, species with a radius of 2.5 nm were formed, whilst at 100 mM sodium chloride small aggregates with a radius of 50 nm were observed. These results indicated that gastric conditions would likely be unsuitable for self-interaction and size exclusion chromatography methods. This challenge is on account of the fact that the large size of the macro-aggregates formed would likely result in blocking of the columns.

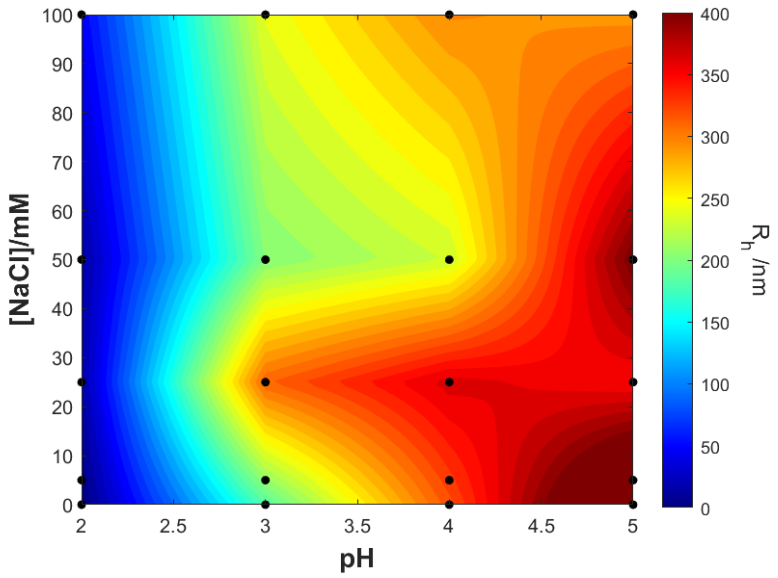


Figure 7.2: Plot of semaglutide hydrodynamic radius under various simulated gastric conditions *upon immediate dissolution (±10%)*

Conversely to gastric conditions, dynamic light scattering measurements of semaglutide under simulated intestinal conditions showed a significantly lower presence of aggregates and much higher solubility. However, at pH 6.0 the peptide did not instantly dissolve unlike at more basic conditions, requiring gentle agitation to ensure complete dissolution. The increased solubility

and reduced propensity to form aggregates instantaneously upon dissolution at pH ≥6.0 relative to pH 5.0 was assumed to be due to increased deprotonation of the imidazole group of the N-terminal His-7 group (pK_a~6.0), resulting in a structural change of the peptide conformation. However, it may also be due to complete dissociation of the stearic acid group (pK_a~4.50), increasing the localised charge of the hydrophobic side chain and so significantly disrupting and weakening hydrophobic interactions between peptide species. Nevertheless, DLS results showed that semaglutide predominantly adopted a structure with a hydrodynamic radius in the range of 2.3-2.7 nm and 2.4-3.0 nm (±0.15 nm) upon dissolution and after 24 hours incubation, respectively, in simulated intestinal conditions (see Figure 7.3). This size was larger than the semaglutide monomer hydrodynamic radius of 1.08-1.27 nm found by Venanzi et al. [320] at lower concentrations. However, it did approximately correlate with hexamer conformations observed by Li et al. [195], suggesting that a stable oligomeric structure was adopted upon dissolution.

Similar to Li et al. it was observed that the size of the oligomeric structure appears to increase with increasing salt concentration by an average of 0.2 nm from 0 mM to 100 mM sodium chloride across the pH range. Furthermore, there was a similar increase of 0.2 nm as the pH decreases from pH 8.0 to pH 6.0 also observed in Figure 7.3. It would be expected that the reduced surface charges at the pHs closer to the isoelectric point and the increased charge shielding at higher salt concentrations should both result in more dominant, attractive hydrophobic interactions, and thus more compact structures being adopted in solution for a given n-mer. However, here the opposite effect was observed. Li et al. suggest that increasing salt concentration promotes the formation of a larger n-mer structure being adopted via slow kinetics as suggested for liraglutide by Wang et al. [192, 195]. However, whilst an increase in the oligomer radius was observed over the 24 hour incubation period, this occurred for all simulated intestinal conditions tested and did not result in a clear stepwise size change which would signify two specific oligomer conformations. Therefore, it is suggested that the variance in average species size

across the range of conditions was indicative of a distribution of the number of species adopting one of two oligomeric species, which were too close in size for the Zetasizer to easily distinguish when mixed together as individual species.

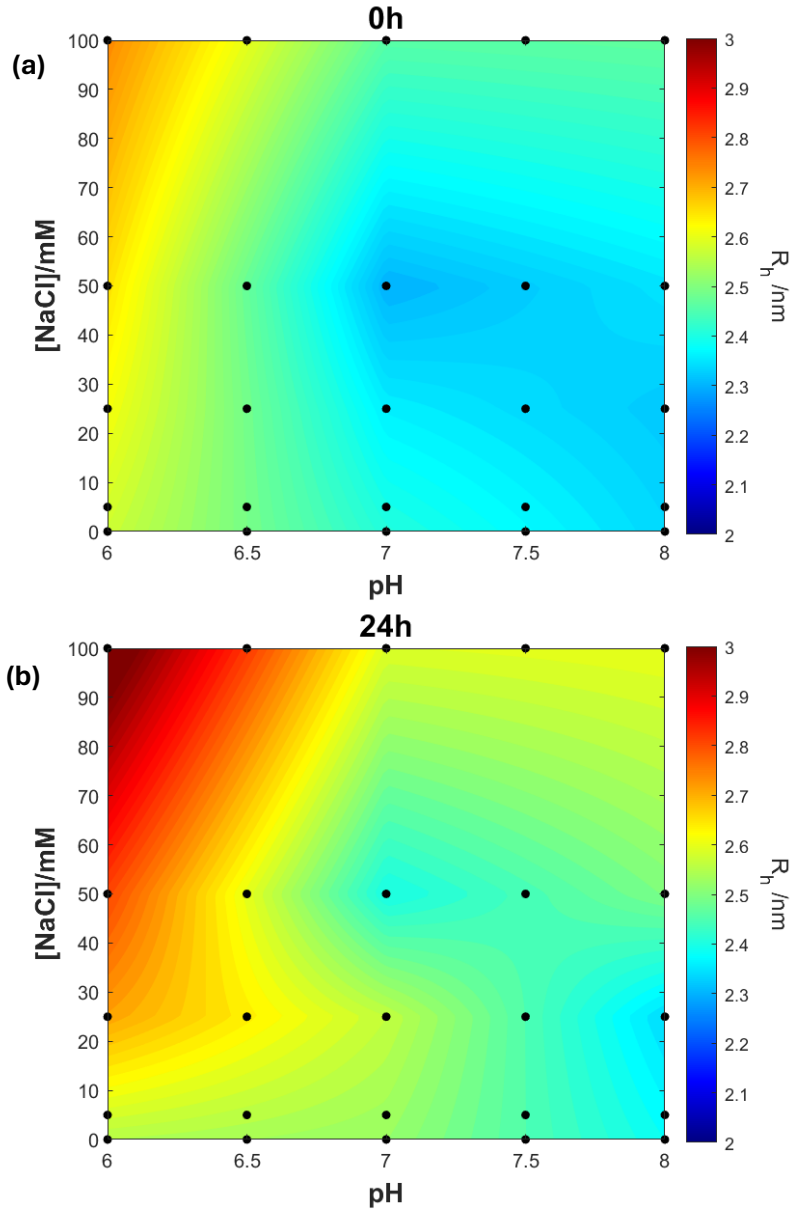


Figure 7.3: Plot of semaglutide hydrodynamic radius as measured via DLS *under different simulated intestinal conditions at 37 °C after (a) 0 hours and (b) 24 hours incubation (±0.15 nm)*

7.3.1.2. Self-Interaction Chromatography

The mass of semaglutide immobilised onto 100 mg Tosoh AF-tresyl-650M resin was determined to be approximately 0.727 ± 0.017 mg (i.e. 7.27 ± 0.17 mg/g.resin) by analysis of the concentration

of the initial and final reaction mixtures in a UV-Vis spectrophotometer. This resulted in a semaglutide molar surface coverage of only 2.21% of the available surface tresyl group sites (assuming a one peptide per tresyl site), which was significantly lower than for other proteins previously reported by Hedberg et al. [346]. This low estimate was assumed to be due to the lower concentration used in the reaction mixture as well as the reduced number of available surface amine groups on the shorter peptide chain relative to larger proteins. Peptide immobilisation was confirmed via Raman spectroscopy measurements of the functionalised-surface and by comparing the spectra with that of an end-capped resin. As may be observed in Figure 7.4, two additional peaks were observed representing carbonyl stretching of the amide bond at 1660 cm^{-1} and aromatic amino acid group ring breathing vibrations (likely phenylalanine) at approximately 1000 cm^{-1}.

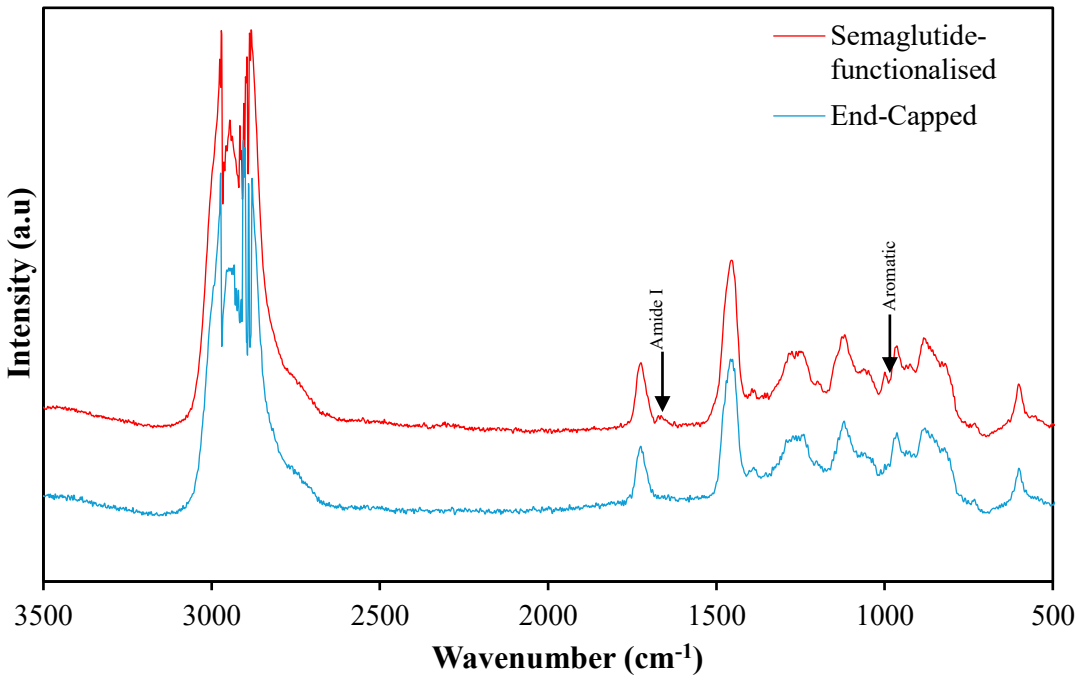


Figure 7.4: Averaged Raman spectra of Tris-end-capped AF-Tresyl-650M resins with (red) and without (blue) semaglutide immobilised onto the surface

The estimated or measured resin, peptide and column parameters required for calculation of the second osmotic virial coefficient via self-interaction chromatography are tabulated in Table 7.1.

The method of calculation of the surface density, ρ_s, volume immobilised, V_i, and the peptide volume of exclusion, B_{HS}, are outlined in Section 5.2.5.1.5, whilst the accessible surface areas were estimated from reverse size exclusion chromatography data reported by DePhillips et al. [214].

Table 7.1: Measured and calculated parameters used for calculation of semaglutide osmotic second virial coefficient (B_{22})
from self-interaction chromatography retention factors

Parameter	Value
Porosity (P)	0.811
Accessible surface area, ϕ_{1r} (m^2 /mL)	21.3
Effective phase ratio, ϕ_{3r} (m^2 /mL)	18.7
Column Volume, V (mL)	0.177
Peptide Molecular Weight, M_w (kDa)	4.11
Peptide Monomer Radius, r (nm)	1.04
Peptide Excluded Volume, B_{HS} (mL)	1.88×10^{-20}
Volume Immobilised, V_i (mL)	0.002
Surface Density, ρ_s (molecule/m^2)	3.43×10^{16}

After column preparation and flushing, initial experiments were conducted to establish the uniformity of the column packing by injecting 2% v/v acetone in water under various conditions. Gaussian peaks with an acceptable, very slight amount of fronting (tailing factor in the range of 0.95-0.98) were observed, an example of which is shown in Figure 7.6. The Gaussian peaks confirmed that the packed columns were uniform in structure. 0.1 mg/mL semaglutide samples in water were prepared and filtrated through 0.22 µm low-protein binding cellulose acetate filters for use as the injected sample. The aggregation state of the peptide was checked at hourly intervals via dynamic light scattering. Initially monomers with a hydrodynamic radius of 1.04 nm were observed as outlined in Figure 7.5a. However, after 24 hours, small peaks with a hydrodynamic radius of 3.51 nm were observed, representing the initial formation of a small number of oligomers (see Figure 7.5b). Therefore, in future experiments the semaglutide sample used in self-interaction chromatography was changed for a fresh sample after 24 hours to minimise sample waste whilst removing any changes in the observed elution behaviour that would be caused by the presence of the oligomer.

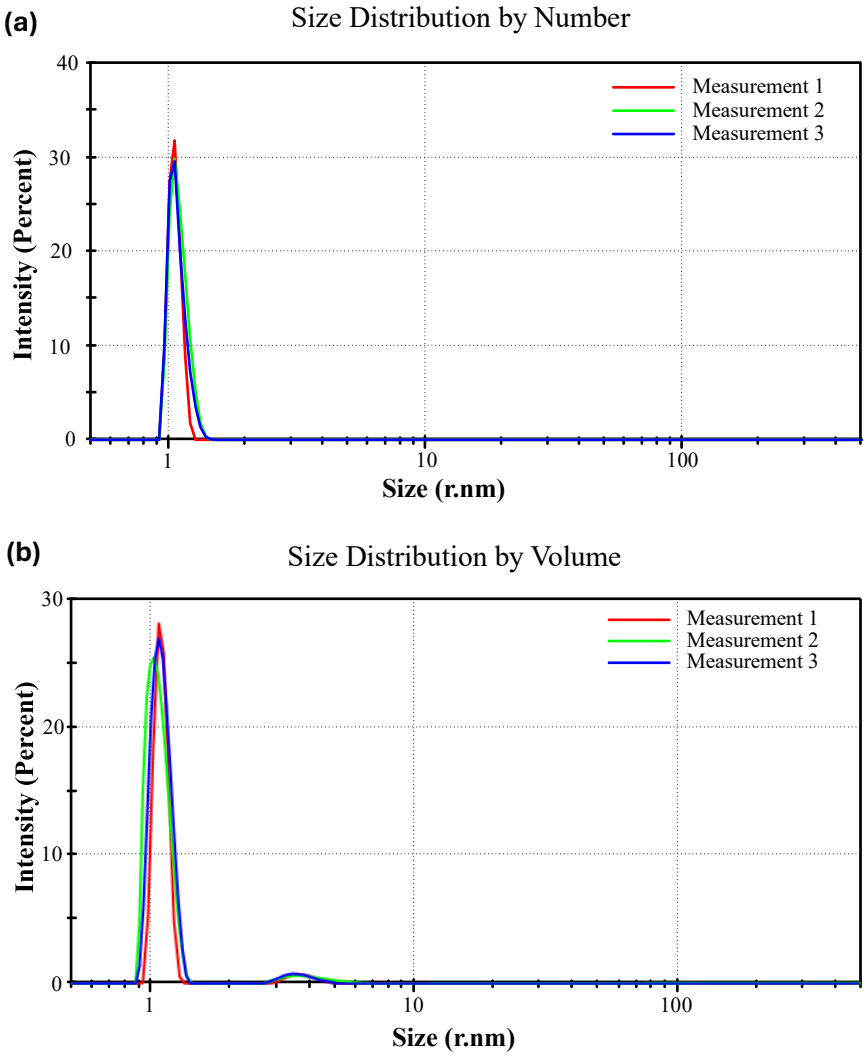


Figure 7.5: Example dynamic light scattering measurements of 0.1 mg/mL semaglutide in water *self-interaction chromatography sample (a) immediately after dissolution and (b) after 24 hours (see Figure 12.7 in Supplementary Materials for correlograms)*

Initial semaglutide experiments displayed similar behaviour to that observed in non-peptide immobilised columns (as discussed in Section 6.3.4). As shown Figure 7.6 for the semaglutide immobilised column, at pH 8.0 the peptide eluted with a lower retention volume than acetone exhibiting minimal/very weak interactions between the semaglutide molecules in the mobile phase with those immobilised to the resin surface. Whilst at pH 6.0, peak shouldering and tailing were observed with much larger retention volumes consistent with larger attractive intermolecular forces. This observation was in agreement with DLS findings suggesting the formation of a different, larger oligomeric aggregate structures, perhaps being accelerated by the high monomer concentration on the surface of the column, resulting in two elution behaviours

(the shouldering) being observed. Furthermore, under fed state gastric conditions (pH 3.0-5.0) no elution was observed. That is, the mobile phase semaglutide species irreversibly interacted with the resin-immobilised semaglutide species or simply formed very large aggregates that could not pass through the porous resin structure. These results were fully consistent with the DLS measurements (see Figure 7.2), and it could be concluded that the SIC technique was not appropriate for highly unstable conditions at which very large aggregates can rapidly form.

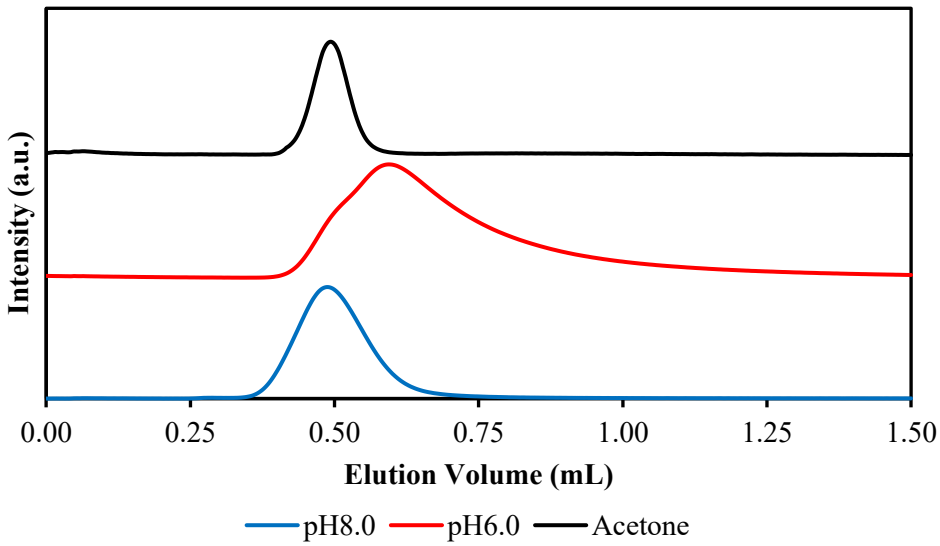


Figure 7.6: Elution chromatograms of 0.1 mg/mL semaglutide at pH 8.0 and pH 6.0 *(0 mM sodium chloride) and 2% acetone solutions through a 177 µL column packed with semaglutide-functionalised AF-Tresyl-650M resin at 0.05 mL/min*

The resultant estimated second osmotic virial coefficients (B_{22}) across the conditions investigated have been plotted in Figure 7.7. As may be observed in the similarity between the plots in Figure 7.3 and Figure 7.7, self-interaction chromatography captured the trend with relatively high accuracy, as smaller oligomers correlated to larger (more positive) B_{22} estimations. This was what would be expected as more positive B_{22} values indicate stronger, net repulsion forces between individual peptide species and so a lower propensity to aggregate. Figure 7.6 provides a clear example of the relative change in retention volume between the most

stable condition, and so fastest eluted resulting in lowest B_{22} values estimation, at pH 8.0 and a less stable condition at pH 6.0, emphasising the sensitivity of the technique to slight changes in interaction strength. Furthermore, errors in retention volume between replicates were at most 0.002 mL, being significantly reduced due to the use of a pre-column detector reducing sample injection error. This resulted in a <1% error between repeats for the final B_{22} values calculated. However, whilst the trend observed aligned qualitatively with that exhibited by DLS measurements, the overall values did not. The DLS results suggested that oligomers were formed instantaneously under all conditions, with larger less stable oligomers being formed at higher salt concentrations and acidities closer to the isoelectric point. Therefore, it would be expected that highly negative B_{22} values would be observed across the range of conditions investigated. Since, slightly positive and slightly negative B_{22} values were observed, it was suggested that the interactions being measured were those between immobilised monomers and mobile phase oligomers. Therefore, the findings suggested that oligomers under more basic conditions and lower salt concentrations were more stable, whilst at alternative conditions the oligomers interacted more strongly with monomers, showing a slightly greater propensity to grow to larger sizes. This explanation correlated with what was observed via DLS. Whilst a B_{23} estimation was not what was expected, nor was the final result accurate in that case, it did not detract from the ability of the technique to broadly predict peptide aggregational behaviour.

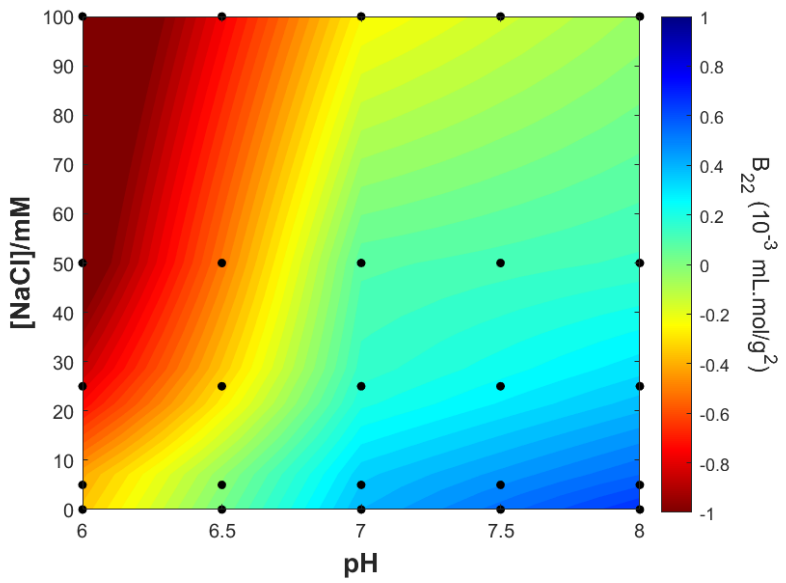


Figure 7.7: Plot of the Second Osmotic Virial Coefficient of semaglutide *at 37 °C in 20 mM sodium phosphate buffers simulating intestinal conditions as estimated via SIC (±2.5 x 10^{-5} mL.mol/g^2)*

7.3.2.Semaglutide Lipid Stability Analysis

Upon initial confirmation that the SIC was able to predict trends in the stability of oligomeric structures formed by semaglutide under simulated intestinal conditions, it was decided to establish whether the technique could be applied to more complex simulated intestinal fluids (SIFs), for which other techniques such as DLS and SEC are not appropriate. Therefore, SIFs, including low concentrations (up to 100 µM) of soybean lecithin to simulate dietary lipids and higher levels of sodium chloride, were prepared. However, only the most stable solution conditions previously investigated (pH 7.0-8.0) were evaluated to ensure poor dissolution of the sample was not a factor in the analysis that may be confused with micelle formation.

7.3.2.1. Dynamic Light Scattering

DLS measurements of semaglutide in soybean lecithin containing SIFs, predominantly showed two primary peaks: the first with a diameter in the range of 4.0-6.0 nm, which aligned with the oligomer peaks previously discussed in Section 7.3.1.1, and a second larger peak, the diameter of which varied between 100 nm and 200 nm depending on the solution conditions (see Figure 7.8a for an example distribution). This larger peak was observed in the presence of lecithin and

was observed in the absence of semaglutide. Furthermore, due to its diameter cut-off being approximately the size of the pore diameter of the 0.22 μm bottle top filters used in buffer preparation, it was assumed to be caused by the presence of lecithin micelles and not lecithin-induced aggregates. Furthermore, investigations by Mkam Tsengam et al. and Yamada et al. confirm that lecithin forms micelle or vesicular studies of 100-500 nm in diameter under similar aqueous conditions [368, 369]. At lecithin concentrations >50 μM, the smaller peak, assumed to be the semaglutide oligomer, was no longer observed, whilst the intensity of the larger peak was greatly increased (see Figure 7.8b). This increase in size further supported the hypothesis that the larger peaks were due to the presence of lecithin micelles. However, due to the limitations of DLS, it was unknown whether the increased micelle concentration resulted in increased scattering preventing oligomer peaks from being observed or whether the semaglutide dissolved into the micelle structure.

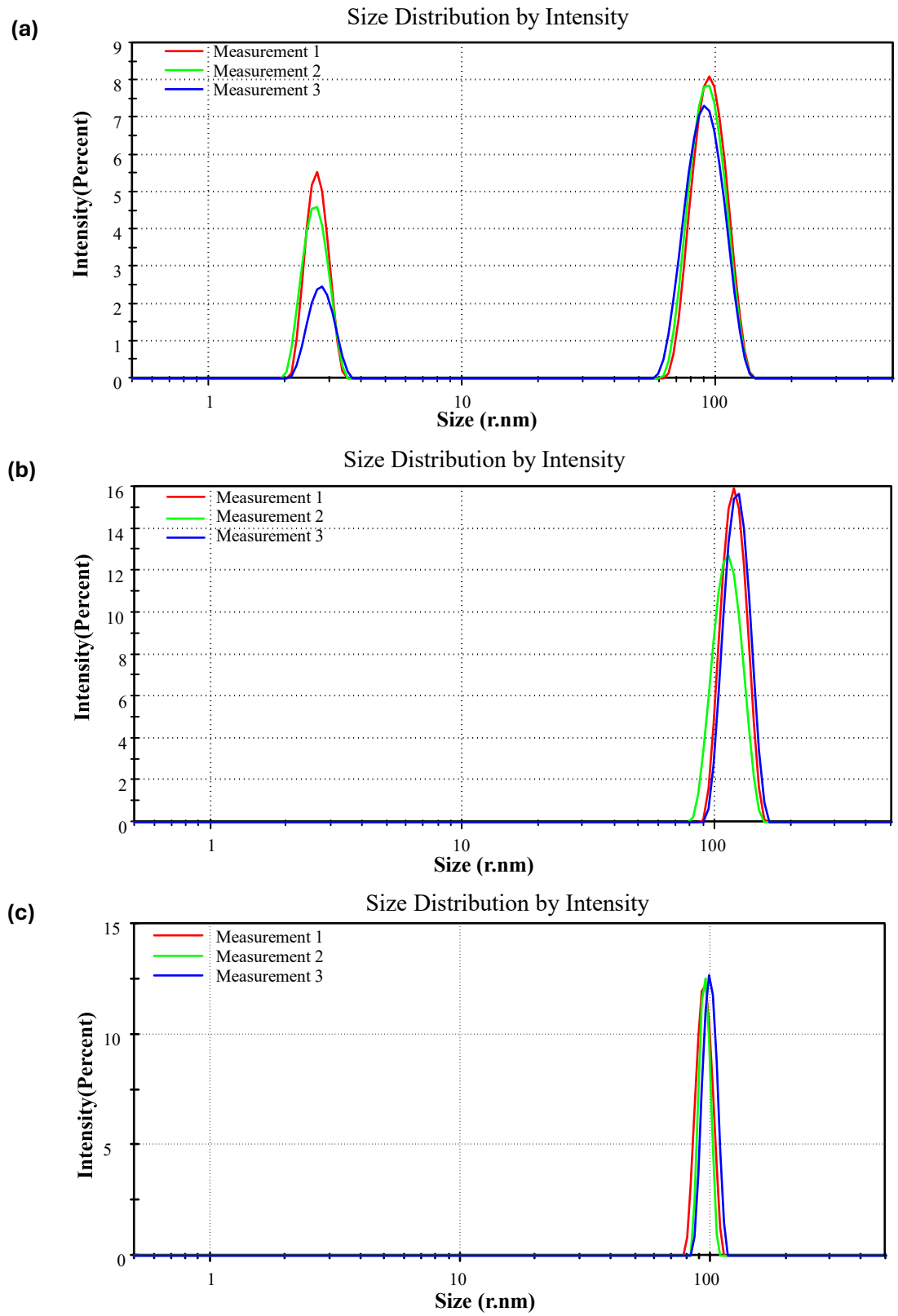


Figure 7.8: Exemplary Dynamic Light Scattering size distributions for 0.5 mg/mL semaglutide *in 20 mM pH 8.0 sodium phosphate with 500 mM sodium chloride and (a) 0.05 mM and (b) 0.1 mM soybean lecithin concentration, and (c) 0.05 mM soybean lecithin in the absence of semaglutide*

Only the sizes of the oligomers specifically were analysed in DLS results and so conditions of >50 μM lecithin were not plotted, since it could not be readily determined, which peaks were directly related to semaglutide aggregation under such conditions. As seen in Figure 7.9, similar results were observed to the previous DLS study: a species with a radius in the range of 2.0-3.0 nm, assumed to be semaglutide hexamers, with higher salt concentrations and lower pHs correlating to a larger size of this species. Three additional observations were made relative to the previous study. Firstly, the increased sodium chloride concentration range resulted in a continuation of the trend increasing species radius without significant plateauing, suggesting the formation of larger oligomers. Secondly, increasing lecithin concentrations resulted in smaller oligomer radii measurements, suggesting that the accompanied increase in hydrophobicity of the aqueous solution stabilises the species formed, thereby reducing their propensity to form larger oligomers. This aligned with the suggested mechanism of action of SNAC in the oral formulation of semaglutide (Rybelsus®). SNAC, due to its amphiphilicity, is suggested to weaken self-association interactions of semaglutide, stabilising the monomer and disrupting oligomer formation [68]. Herein, the lecithin is suggested to have a similar effect of weakening the hydrophobic interactions, albeit to a much lesser degree. The third and final observation is that in the presence of lecithin, the change in size of the species with time was minimal at high lecithin and low salt concentrations. Conversely, the increase in species radius was greater than 0.2 nm over the 24-hour period in the absence of lecithin and at sodium chloride concentrations greater than 100 mM. These trends were observed most clearly at pH 7.0 (Figure 7.9 A1-3), where, of the conditions tested, stability of the oligomer has been suggested to be lowest.

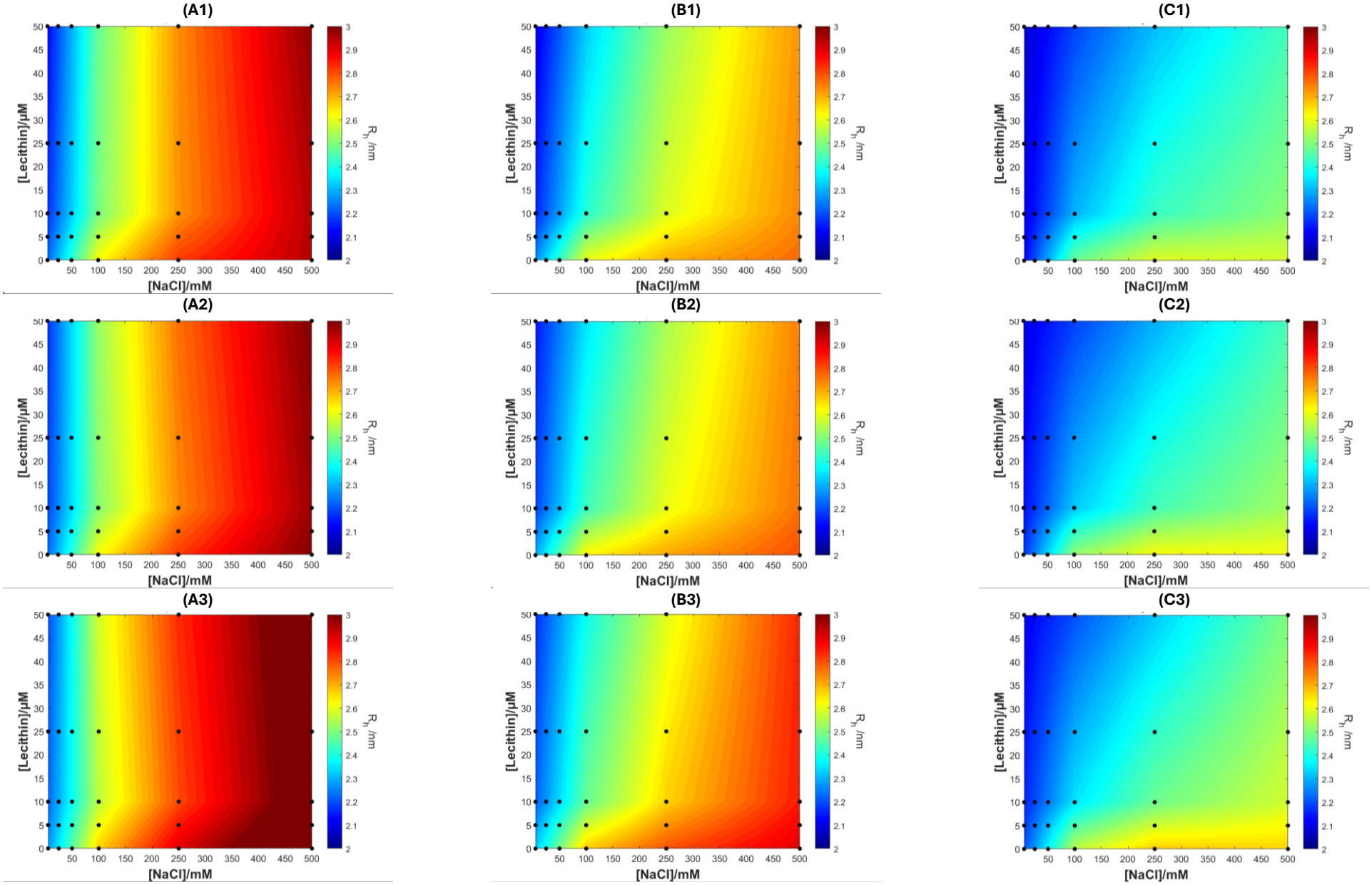


Figure 7.9: Plots of hydrodynamic radii of 0.5 mg/mL semaglutide species
at 37 °C in 20 mM sodium phosphate buffer at (A) pH 7.0, (B) pH 7.5 and (C) pH 8.0 at (1) 0 h, (2) 6 h and (3) 24 h timepoints (±0.12 nm)

7.3.2.2. Self-Interaction Chromatography

Semaglutide immobilisation was reperformed to prepare fresh self-interaction chromatography columns. However, due to the low initial reaction levels, the concentration of semaglutide in the immobilisation solution was doubled to increase coverage or the extent of the immobilisation reaction. UV-Vis analysis of the initial and final immobilisation solutions showed the mass of semaglutide to be approximately 3.23 ± 0.06 mg onto 100 mg (i.e. 32.3 ± 0.6 mg/g.resin) of the same resin as previously, which represented approximately 9.8% molar coverage of the available surface tresyl sites (assuming a one peptide per tresyl site). This was a significant improvement over the initial immobilisation coverage but meant that in order to achieve coverages of >20%, even higher initial concentrations of peptide would be required. The finalised parameters for calculation of the second osmotic virial coefficient (B_{22}) for the new columns were the same as those outlined in Table 7.1 with the updated values due to the difference in surface coverage shown in Table 7.2.

Table 7.2: Updated parameters used for the new self-interaction chromatography column

Parameter	Value
Volume Immobilised, V_i (mL)	0.009
Surface Density, ρ_s (molecule/m^2)	1.53×10^{17}

As for the previous experiments, similar packing quality analysis experiments were conducted via injections of 10 µL of 2% acetone solution which showed tailing factors of between 1.05 and 1.10. However, at lecithin concentrations of ≥50 µM, the 2% acetone injection gave tailing factors that increased up to 1.20 and inconsistent ghost peaks began to appear in the elution profiles. These peaks were assumed to be on account of lecithin micelles in the mobile phase being held up within the UV-Vis detector chambers causing a small temporary spike-increase in absorbance, since they were detected in both pre- and post-column detectors whether there was injection or not. The overall result was that the elution profile was significantly less smooth than in the absence of lecithin and as such smaller peaks had to be ignored, where the primary elution

peak was self-evident (see Figure 7.10 for example profiles). As a result, the accuracy of the retention volumes was slightly reduced due to the calculation of the peak centre of mass being impacted by the additional inconsistent peaks.

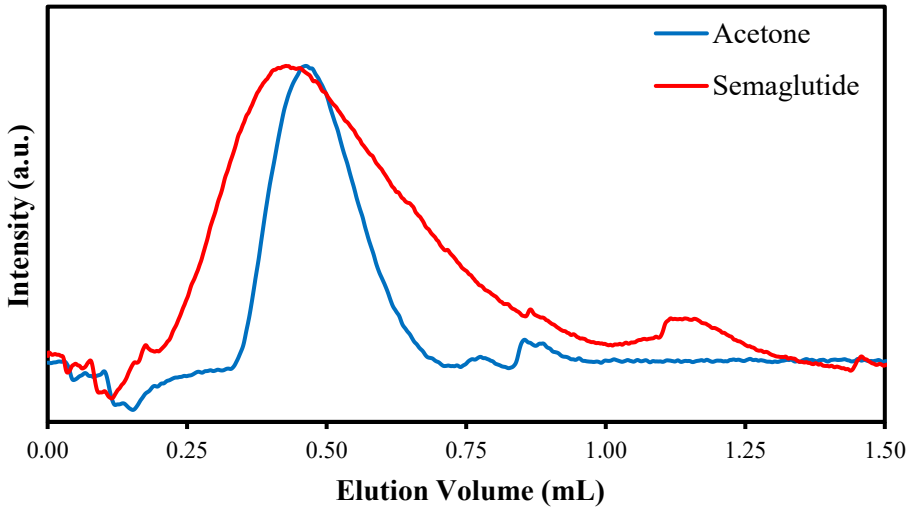


Figure 7.10: Elution profiles of 2% acetone and 0.1 mg/mL semaglutide through a semaglutide-fixated SIC column
(20 mM pH 8.0 sodium phosphate,100 mM sodium chloride and 100 µM soybean lecithin at 37 °C)

The calculated B_{22} values showed good correlation with the trends observed in the DLS results. Higher salt concentrations and lower pH and lecithin concentrations correlated with more negative B_{22} values, indicating an increasing propensity to aggregate (see Figure 7.11). However, where the results from the two techniques differed was that the effect of increasing lecithin concentration appeared to have a greater impact on increasing the second osmotic virial coefficients calculated via SIC than it did on the hydrodynamic radii measured via DLS. This difference was assumed to have been caused by the presence of micelles resulting in one of two phenomena. The first was increased scattering in the DLS measurements such that the oligomer size measured was greater than in actuality. The alternative effect would be a reduction in void volume in the self-interaction chromatography column due to lipid micelle species filling the porous network, resulting in reduced surface area for retention and faster elution and so an apparently more positive B_{22} value. The latter effect of micelles in solution should in fact be

accounted for in the estimations of the void volumes via dead column measurements and acetone injections. Therefore, it was concluded that it was more likely that the discrepancy in the trend was caused by DLS overestimating the size of the oligomeric species at higher lecithin concentrations due to increased scattering from micelles in solution, given the scattering intensity is proportional to the diameter of the macromolecule to the sixth power [199]. The additional benefit of self-interaction chromatography may also be observed in Figure 7.11, wherein estimations of B_{22} values were made for semaglutide at soybean lecithin concentrations up to 100 μM. These results showed that increasing the lipid concentration and so solution hydrophobicity acted further to stabilise the peptide structure without any indication of a plateau.

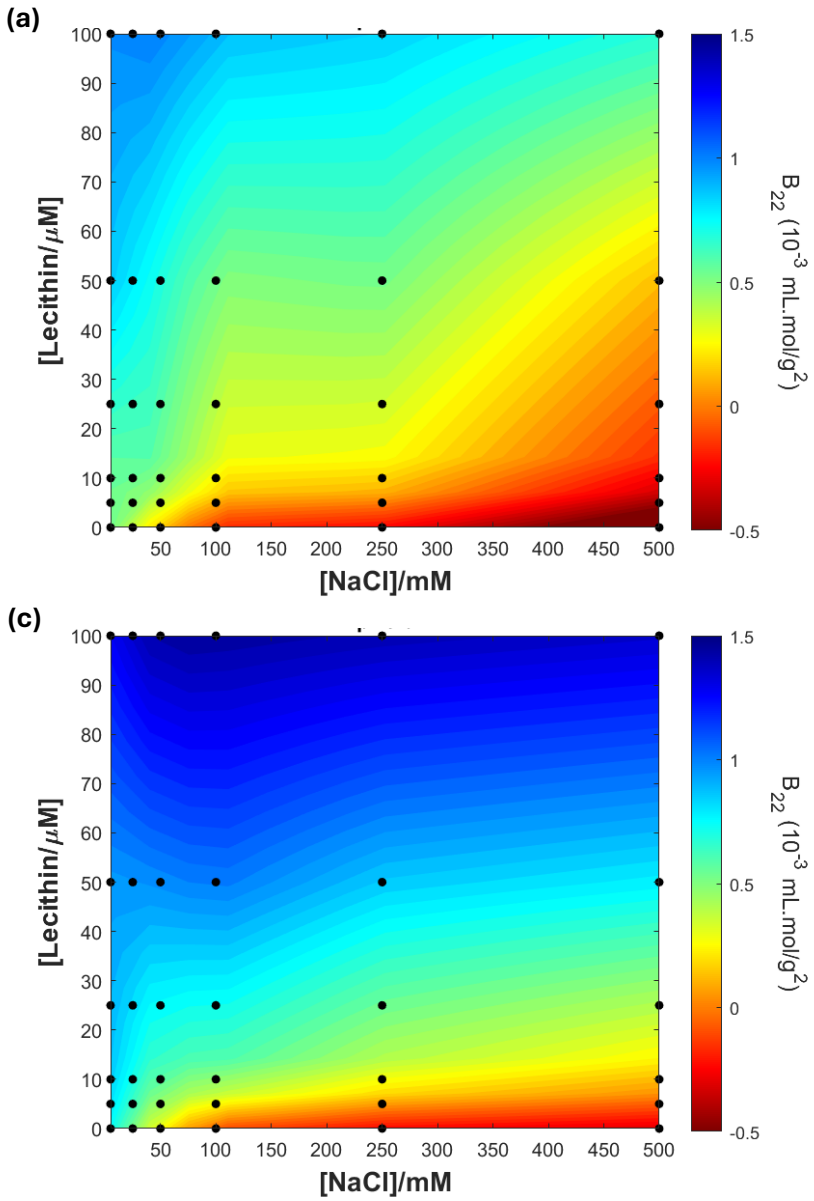


Figure 7.11: Plots of the second osmotic virial coefficient of semaglutide *at 37 °C in 20 mM sodium phosphate buffers with varying sodium chloride and soybean lecithin concentrations at (a) pH 7.0, (b) pH 7.5 and (c) pH 8.0 simulating intestinal conditions as estimated via SIC (± 8 x 10^{-5} mL.mol/g^2)*

7.4. Conclusions

Variations in the second osmotic virial coefficient (B_{22}) of semaglutide were measured across a range of simulated intestinal conditions from pH 6.0-8.0, with varying sodium chloride and soybean lecithin concentrations mimicking dietary salts and lipids. Gastric conditions were not included in the study as they were observed to result in instantaneous and complete aggregation of the peptide across almost all conditions. Of the parameters investigated, pH was observed to have the greatest impact on B_{22} values with more basic conditions providing the greatest stabilising effect on the peptide structure. Reducing salt and increasing lipid concentrations also resulted in greater B_{22} values and so predicted a more stable conformation aligning with what would be expected theoretically.

The overall trend in B_{22} results correlated with the trend in oligomer sizes observed via DLS, with larger oligomers being observed under conditions where more negative B_{22} measurements were made via SIC. However, since oligomers were observed via DLS under all conditions, the associated B_{22} values would be expected to be highly negative, indicating strongly attractive forces between monomers. Rather, a range of slightly positive and slightly negative B_{22} were measured via SIC indicating mild stability. This discrepancy was assumed to be due to SIC measuring the interaction strength of mobile phase oligomers with immobilised monomers, thereby showing that the oligomers were somewhat stable, which DLS observations over a 24-hour period showed them to be. Additional studies showed that soybean lecithin was measured to have a greater stabilising effect via SIC than observed via DLS, this was deemed to be due to increased scattering caused by micelles resulting in increased estimation of the oligomer size, thereby appearing to weaken the correlation.

Therefore, SIC showed potential as a relatively high-throughput technique to realise trends in behaviour of semaglutide oligomers in simulated intestinal environments. However, the method did not accurately predict the exact behaviour of the peptide monomer and understanding of the

results required understanding of the solution assembly behaviour adopted by the peptide upon dissolution. For the method to accurately predict monomer aggregation propensity in the long-term, semaglutide would likely need to be studied using formulations allowing it to maintain monomeric conformations at concentrations sufficient to be observed via other methods such as DLS.

CHAPTER 8

8. SELF-INTERACTION CHROMATOGRAPHY OF INSULIN AND LIRAGLUTIDE

8.1. Introduction

Following on from the previous Chapter, in which the applicability of self-interaction chromatography to study semaglutide solution stability under intestinal conditions was studied, it was considered necessary to expand the study to other peptides to establish the consistency and adaptability of the technique to species with different aggregation pathways. The peptides selected were liraglutide, another glucagon-like peptide-1 (GLP-1) analogue with a structure highly similar to semaglutide, and insulin. The latter of the two being a two-chain peptide with a very different structure and aggregation pathway, but the physical stability of which has previously been study extensively and so is well understood [370, 371].

Despite the similarities in their structures and unlike semaglutide, an oral formulation of liraglutide is yet to be approved by the U.S. Food and Drug Administration (FDA) or the European Medicines Agency (EMA). The sodium N-[8-(2-hydroxybenzoyl)amino]-caprylate (SNAC) employed as a transcellular permeation enhancer in the oral formulation of semaglutide, does not exhibit the same effectiveness at stabilising liraglutide against oligomerisation, thereby hindering its absorption [68]. The suggested cause of liraglutide's increased propensity to form oligomers is the relatively higher hydrophobicity of the palmitoyl acylation region relative to that of the stearic acid group on semaglutide. This difference emphasises how minor modifications to a peptide structure can fully alter their aggregation pathway and so their formulation design.

Despite research into developing oral insulin delivery strategies or formulations being extensive, few have made it past stage II clinical trials and none are yet to be approved by the FDA or EMA [372, 373]. In accordance with Lipinski's rule of five, despite its application being predominantly for small molecule drugs, one of the factors impeding oral insulin from permeation of the intestinal membrane and achieving a high bioavailability, is the high molecular weights of these peptides [13]. Therefore, its propensity to dimerise or hexamerise in neutral conditions [370, 371] or even aggregate in Lys- or Arg-containing simulated intestinal fluids [374] further impedes its ability to be absorbed through the intestinal epithelium.

The work in this Chapter aims to further develop the self-interaction chromatography investigations in the last Chapter to establish the sensitivity of the setup and procedure previously optimised for semaglutide to a change in peptide. An additional goal was to confirm whether the suggestion of measuring the second osmotic virial coefficient between monomer and oligomer for semaglutide was accurate by determining whether similar observations were made for other peptides. Aside from this, if the method showed good alignment with size exclusion chromatography and dynamic light scattering results, it was hoped that results would further the understanding of the aggregational propensity of liraglutide and insulin under a range of simulated intestinal conditions.

8.2. Materials & Methods

All experiments in this Chapter were carried out using a tresyl-functionalised methcrylate resin from Tosoh Bioscience as described in Section 5.1.7. Within the same section, descriptions of the tris(hydroxymethyl)aminomethane and sodium phosphate used are described as well as the buffer preparation procedure. Additionally, the liraglutide and insulin used was that described in Sections 5.1.2.1 and 5.1.1, respectively.

Peptide immobilisation and resin end-capping procedures were performed as outlined in Sections 5.2.3.1 and 5.2.3.2. Raman Spectroscopy of the resin surface was carried out as in

5.2.3.5. The column packing procedures are outlined in Section 5.2.4.1. Self-interaction and Size exclusion chromatographic experiments were carried out on LC-20 and Nexera high-performance liquid chromatography (HPLC) systems, respectively, from Shimadzu the configuration of which is described in 5.2.1. The Self-interaction chromatography experimental procedures were performed as outlined in Section 5.2.5.1, whilst the amount of peak-tailing was determined following the description in Section 5.2.4.3. Dynamic Light Scattering and Size Exclusion Chromatography procedures are outlined in Sections 5.2.5.2 and 5.2.5.3, respectively.

8.3. Results

8.3.1. Liraglutide Stability

8.3.1.1. Dynamic Light Scattering

Initial studies of liraglutide stability were performed using Dynamic Light Scattering (DLS) measurements under simulated intestinal conditions, varying the acidity from pH 6.0 to pH 8.0 and the salinity from 0 mM to 50 mM NaCl. These measurements allowed for an accurate measurement of the size of the liraglutide species in solution, which would provide an indication of an appropriate choice of size exclusion chromatography column. Then, these results could be correlated and compared to final self-interaction chromatography second virial coefficient values to establish the accuracy of the technique under such conditions.

Upon dissolution of the liraglutide, almost the entirety of the species detected in a single DLS measurement had a hydrodynamic radius in the range of 3.0-3.5 nm. However, some measurements detected a small number of smaller species. An example sample triplicate measurement, each comprised of at least 12 runs, is shown in Figure 8.1. This example illustrates the consistency in size of the species formed and the inconsistent detection of smaller species highlighted by the small peak at R_h~0.65 nm in the second replicate.

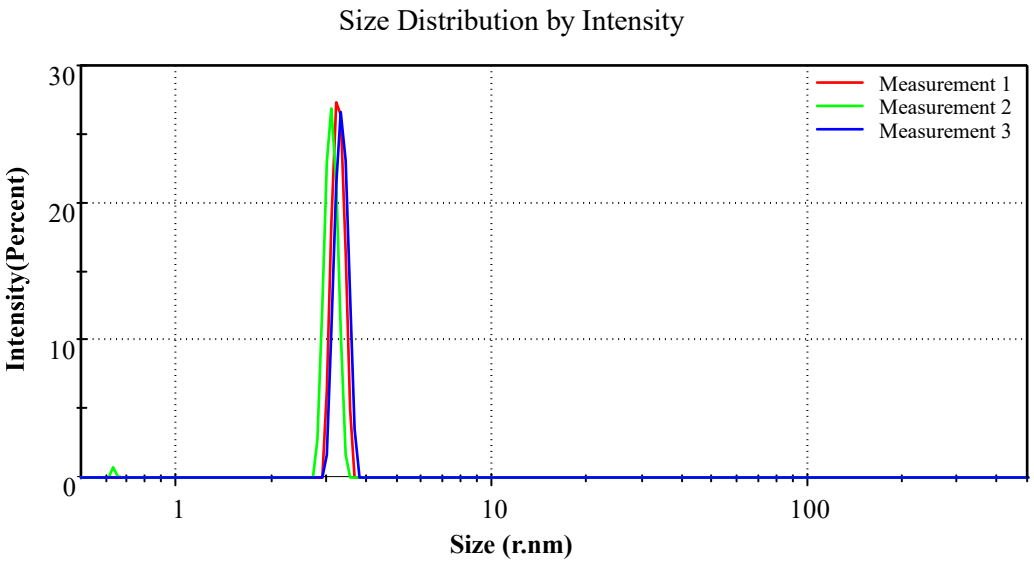


Figure 8.1: Example dynamic light scattering triplicate measurement of 0.5 mg/mL liraglutide *in pH 8.0 20 mM sodium phosphate buffer with 0 mM sodium chloride (see Figure 12.9 in Supplementary Materials for correlograms)*

The initial timepoint DLS measurements of liraglutide across the range of conditions tested showed the same species with a 3.0-3.5 nm hydrodynamic radius (see Figure 8.2a). Each measurement showed strong consistency with an error of ±0.15 nm. However, this is a far larger size that would be expected of a peptide of the same molecular weight (3751.2 Da). The minimum hydrodynamic radius (R_{min} in nm) for a spherical protein or peptide of a given molecular weight (M in Da) may be estimated by the following equation [375]:

$$R_{min} = 0.066M^{\frac{1}{3}} \qquad \textbf{Equation 8.1}$$

Therefore, assuming an ideal spherical monomer, liraglutide would be expected to have a hydrodynamic radius of 1.02 nm. Given that the shape factor for even the most elongated proteins such as fibrinogen is less than 2.5 [375] and the fact that DLS measurements show sharp, narrow peaks (as seen in Figure 8.1), atypical of highly non-spherical species. Therefore, it could be quickly concluded that liraglutide formed an aggregate or oligomer species instantaneously upon dissolution under the conditions tested. Species of approximately 3.4 nm hydrodynamic radius, according to Bothe et al. [193] are indicative of12-or 13-mer micelle

structures previously reported by Wang et al. [192]. However, in previously published studies, this 12-mer was observed primarily at pH < 7.0. It was hypothesised to be due to the protonation of N-terminal Histidine causing a structural change as discussed in Section 5.1.2.1, whilst a 7- or 8-mer oligomer species (R_h ~2.6 nm) was observed under more basic conditions [192, 193]. This difference from literature values is assumed to be due to the higher, physiological temperature of 37 °C in this study, relative to those previous room temperature, 25 °C studies. The higher temperature is hypothesised to result in increased thermodynamic stress destabilising the octamer species and shifting the equilibrium to favour the larger 12- or 13-mer oligomers following initial sample dissolution.

The deviation in R_h measurements recorded from the two oligomer species observed in literature may also be as a result of DLS measurements detecting a mixture of the two species and not differentiating between them. This would account for the non-uniformity associated with a DLS size distribution peak for a non-spherical species. Assuming this, a mixture of the two similarly-sized oligomer species being detected would result in peak overlap and a single, unresolved, broad measurement peak with an average between R_h value between 2.6 nm (octamer) and 3.4 nm (dodecamer). Therefore, this may provide an alternative explanation for the range of hydrodynamic radii being 3.0-3.5 nm and not either 2.6 nm or 3.4 nm as described by Wang et al. [192] and Bothe et al. [193].

The oligomer species detected appeared to be highly favourable, since, as observed in Figure 8.2a, there was little variation in their hydrodynamic size with increasing NaCl concentration at all pHs. This observation suggests that the structure is primarily held together through hydrophobic interactions of the lipid side chain group in an optimal micelle configuration, since increasing NaCl concentration would be expected to result in an increase in hydrodynamic radius. This is on account of the fact that higher NaCl concentrations result in reduced electrostatic repulsions between the peptide chains allowing for more dense packing. However,

since this was not observed, it can be assumed that electrostatic forces between individual peptides do not dictate oligomer conformation under these conditions, but rather the hydrophobic interactions of the lipid sidechain.

After incubating the samples for 6 hours and 24 hours, a different size trend began to be observed. For peptide species stored at pH 7.0 and above, the oligomeric structures observed remained approximately the same size, if not decreasing very slightly by approximately 0.2 nm on average after 24 hours (see Figure 8.2c). This is hypothesised to be due to the 12-mer transitioning to the 8-mer, which as previously mentioned was observed to be the favoured oligomeric structure of liraglutide at pH >7 and room temperature. Wang et al. state that the kinetics of the transformation is slow with several days being required for equilibrium to be reached [192]. Furthermore, due to the close similarity in size of the two oligomers, it is likely that DLS and the Zetasizer data processing software would be unable to distinguish between them very easily, given a difference in size of approximately a factor of 5 is required for clear resolution [203]. Therefore, the peak observed is merely an average radius of the 12-mer and 8-mer species. Therefore, had the sample been incubated for longer, perhaps the hydrodynamic radius may have approached 2.6 nm, indicating the octamer species would have been in the majority.

Conversely at pH < 7.0, an increase in the hydrodynamic radius of the oligomer was observed by an average of 0.2 nm after 6 hours and 0.4 nm after 24 hours, finally to be in the range 3.6-3.8 nm. This could have been due to one of three following reasons (from most likely to least):

1. Any 7- or 8-mer conformations present in the original 0 h timepoint measurements transitioning to 12- or 13-mer micelles as previously observed by Bothe et al. [193]. In this case, the discrepancy in hydrodynamic radius of the oligomer species compared to literature values may have been caused by the method through which the Zetasizer software estimates R_h. It fits the data to a discrete size distribution, the discretisation of

which is dependent on user settings using the CONTIN algorithm, and so may have resulted in a slightly higher average value than provided in literature [376]

2. The 12- or 13-mer micelles beginning to coalesce together showing the initiation of peptide aggregation or fibrillation
3. The 12- or 13-mer peptide conformation transitioning to another larger oligomeric structure, which is more stable at higher temperatures and pH <7.

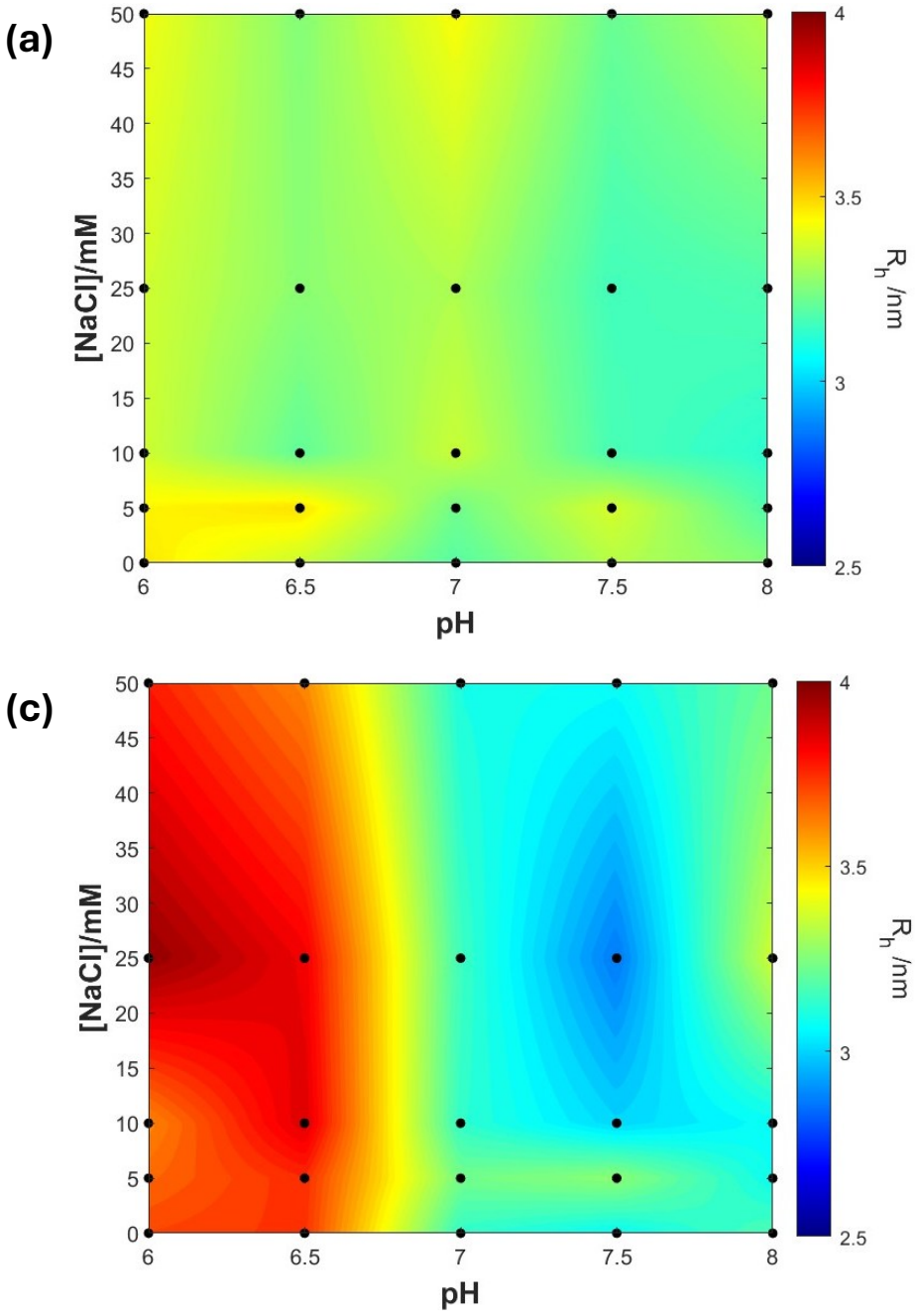


Figure 8.2: Plot of liraglutide hydrodynamic radius as measured via DLS *under various simulated intestinal conditions (a) upon immediate dissolution and after (b) 6 hours and (c) 24 hours incubation (±0.13 nm)*

The exact mechanism that occurs is outside the scope of this work, as primarily the concern is whether the peptide aggregates, how is that effected by solution conditions, and whether self-interaction chromatography may predict the same behaviour. Nevertheless, it was assumed that the most likely reason was the first (of the three listed), since there have not been reports of larger oligomeric structures being observed in the literature. Additionally, some large amorphous aggregates were observed at pH 6.0 at both 6 hours and 24 hours timepoints, thereby reducing the likelihood of the initiation of fibril formation being observed. However, these aggregates were not always observed, perhaps due to sedimentation during sample equilibration prior to sample measurement. Additionally, they were not observed at all at pH >7.0. Example measurements showing large, >150 nm hydrodynamic radius, peptide fibrils at pH 6.0 and their absence at pH 8.0 are represented in Figure 8.3. Wang et al. previously observed aggregation of 4 mg/mL liraglutide within 10 minutes at pH <6.3 [192]. Since it is known that for pH <7.0 Liraglutide fibrillates more rapidly at lower peptide concentrations [193], the observation of fibril formation at pH <7.0 as shown in Figure 8.3a, appeared to be in agreement with those results.

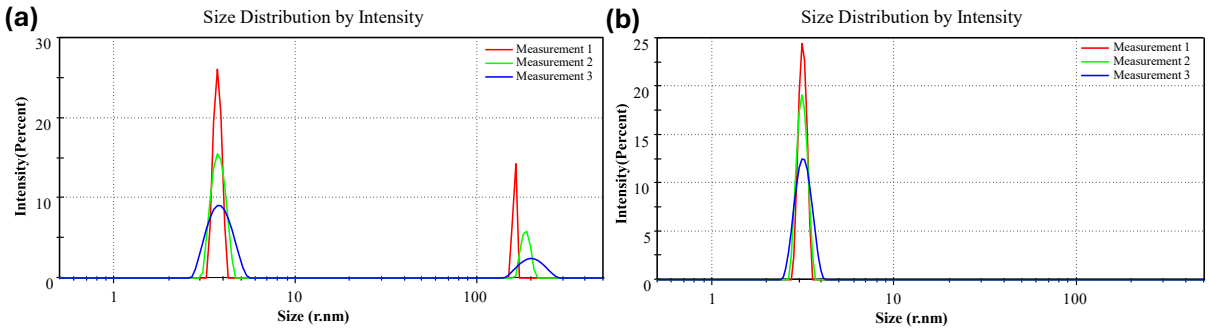


Figure 8.3: Example DLS measurements of 0.5 mg/mL liraglutide after 24 hours incubation *in (a) pH 6.0 and (b) pH 8.0 20 mM sodium phosphate (see Figure 12.10 in Supplementary Materials for correlograms)*

8.3.1.1. Size Exclusion Chromatography

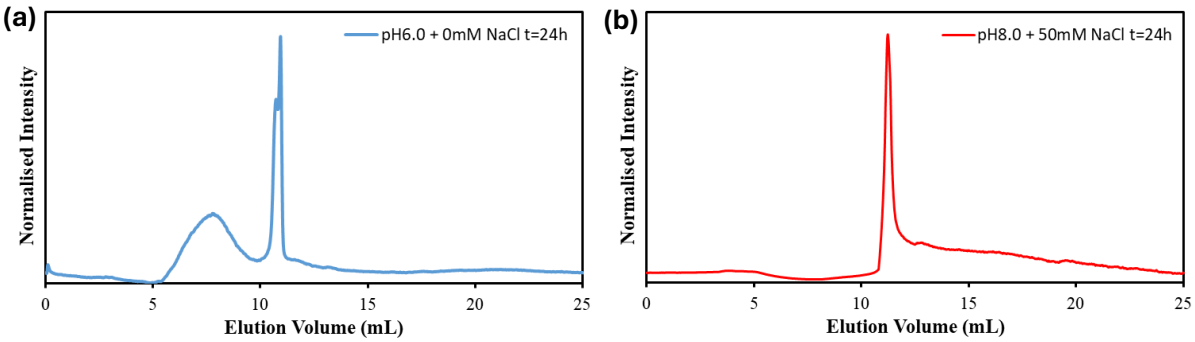


Figure 8.4: Example SEC elution curves of 0.5 mg/mL liraglutide samples *incubated in (a) pH 6.0 and (b) pH 8.0 20 mM sodium phosphate at 37 °C for 24 h (1.0 mL/min 50 mM pH 7.0 sodium phosphate + 150 mM sodium chloride*

Since DLS results showed the rapid formation of an oligomer conformation upon dissolution, a TSKgel G3000SW$_{XL}$ size exclusion column, typically used for the analysis of monoclonal antibodies and large protein aggregates, was selected for sample analysis. Whilst this column would exhibit relatively poorer separation and resolution between monomer and oligomer peaks, given its calibration range of 10-500 kDa, it would be much less likely to become blocked by sample aggregates, enabling safer operation for the samples at pH <6.3. Two example elution curves have been displayed in Figure 8.4 for liraglutide samples incubated at 37 °C for 24 hours, one at pH 6.0 and at pH 8.0. They exhibit similar trends to the corresponding example DLS measurements in Figure 8.3. A primary peak with a slightly larger retention volume at the higher pH, representing a slightly smaller oligomer, and a broad early eluting aggregate peak at the lower pH, representing a broad distribution of precursor aggregate species.

The two chromatograms in Figure 8.4 together typify four key trends that were observed throughout the size exclusion chromatography (SEC) experiments. Firstly, the initial broad peak in the pH 6.0 elution curve was only observed under conditions at pH <6.5 and, due to its correlated molecular weight being >100 kDa, it was assumed to be indicative of large aggregates and their sheared fragments. Second, the peak splitting observed in the pH 6.0 but not the pH 8.0

chromatogram (see Figure 8.4) was a consistent feature within the data but was assumed to be due to physical hinderance in the frit at the column inlet due to the presence of the larger liraglutide aggregates. The difference in retention volume between the maxima within the split peak was insufficient to be the difference between the two different oligomer species, although it may still have been indicative of a mixture of the two oligomer species. Thirdly, the retention volumes of the primary peaks of different samples were all within a range of <1 mL, with an average error of ± 0.25 mL. Therefore, trends observed between different conditions were not much greater than the margin of error. Finally, as more clearly observed in the liraglutide sample incubated at pH 8.0, a reasonably large amount of peak tailing was often observed. This could be indicative of one of two factors: shear-induced dissociation of the sample oligomer, which has been previously observed for polystyrene although not peptides [377], resulting in several smaller sized species eluting, or sample interaction with the stationary phase, resulting in retention in the column. This latter phenomenon has previously been demonstrated to result in a slight increase in insulin retention time and a great increase in peak broadening and tailing with a silica-based SEC stationary phase [378]. These four properties of the data were required to be factored into any analysis of the final data results.

The SEC data, as shown in Figure 8.5, exhibited a very similar overall trend to the DLS results: a reasonably uniform size distribution of species across the tested conditions upon initial sample dissolution, with species incubated at pH <7.0 increasing in size over the 24 hour period and those at pH ≥7.0 reducing in size. However, the size of the species, as estimated from the column calibration, was in disagreement with the DLS results. For example, the measurements after initial dissolution as observed in Figure 8.5a were in the 20-25 kDa range, indicative of a 6- or 7-mer species or a hydrodynamic radius of between 2.3-2.4 nm, significantly smaller than the 3.1-3.4 nm range measured via DLS for the same timepoint. This was true across all measurements with the largest species estimated at approximately 34 kDa or a hydrodynamic radius of 2.8 nm at pH 6.0 and 0 mM sodium chloride after 24 hours incubation (see Figure 8.5c). This species

would only be equivalent to approximately a 9-mer, which is significantly smaller than those collated via Dynamic Light Scattering. Furthermore, the smallest estimated data was estimated at 17.5 kDa equivalent to a 5-mer or a hydrodynamic radius of approximately 2.1 nm. This species is significantly smaller than those suggested by Bothe et al. or Wang et al. [192, 193] (although pentamers were observed for liraglutide by Marassi et al. [379]). This furthered the trend of the SEC consistently underestimating the species sizes relative to DLS. Therefore, it could be concluded that as discussed regarding the tailing in Figure 8.4, either there was sample interaction with the column stationary phase resulting in a higher retention volume, or shear-dissociation of the oligomer was occurring resulting in smaller species with longer mean-free elution paths. However, due to the narrow, sharp peak and the presence of tailing, it was assumed that the sample was interacting with the stationary phase via a consistent, linear mechanism that resulted in higher retention volumes relative to the calibration curve, as previously discussed by Javeri et al. for the size exclusion analysis of insulin [378]. Therefore, it was not unexpected, due to the great difference in structure of the peptide micelle relative to the calibration species, that the oligomers did not elute according to the same linear calibration. Nevertheless, since the overall trend observed was the same as for DLS, with the presence of instantaneous oligomer formation upon dissolution, the qualitative information that may be correlated to a self-interaction chromatography method was the same.

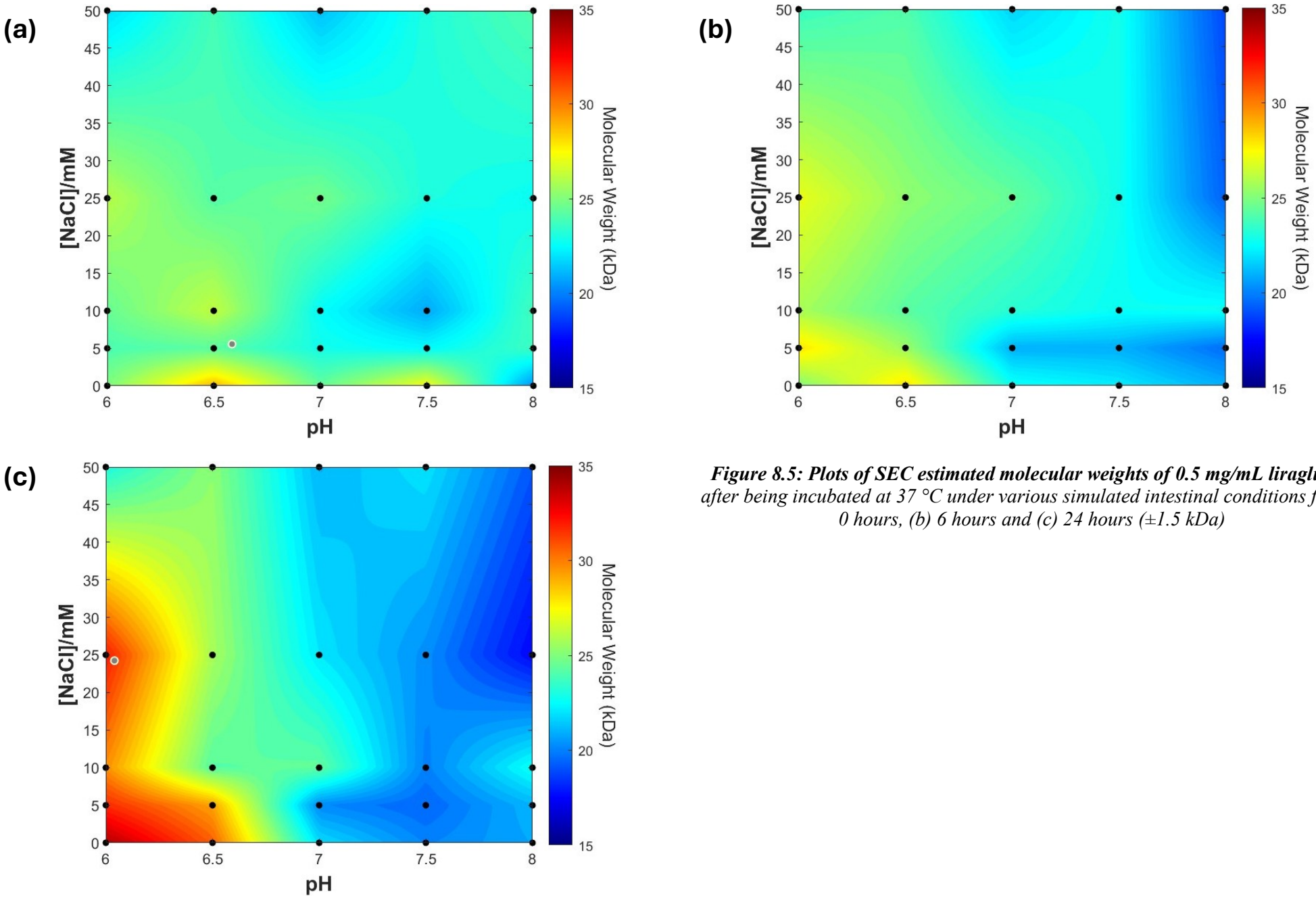


Figure 8.5: Plots of SEC estimated molecular weights of 0.5 mg/mL liraglutide *after being incubated at 37 °C under various simulated intestinal conditions for (a) 0 hours, (b) 6 hours and (c) 24 hours (±1.5 kDa)*

8.3.1.2. Self-Interaction Chromatography

Liraglutide was first immobilised onto an AF-Tresyl-650M resin from Toyopearl. The initial and final concentrations of liraglutide in the immobilisation solution were measured to be approximately 1.83 mg/mL and 0.76 mg/mL by UV spectroscopy, resulting in an estimated 5.37 ± 0.08 mg liraglutide immobilised onto 100 mg of resin (i.e. 53.7 ± 0.8 mg/g.resin). This represented a 17.9% molar coverage of the available surface tresyl sites (assuming a one peptide per tresyl site). Liraglutide immobilisation was further confirmed using Raman-spectroscopy of the peptide-immobilised and dead column resins via the presence of three small additional absorption peaks relative to the end-capped resin: the amide I band, at approximately 1660 cm^{-1}, represents carbonyl stretching of the peptide bond; the C=C aromatic stretching at approximately 1600 cm^{-1}, indicates the addition of the aromatic groups Phe-12, Tyr-19, Phe-28 and Trp-31 to the resin surface, whilst the peak at approximately 1000 cm^{-1} is indicative of aromatic amino acid ring breathing vibrations.

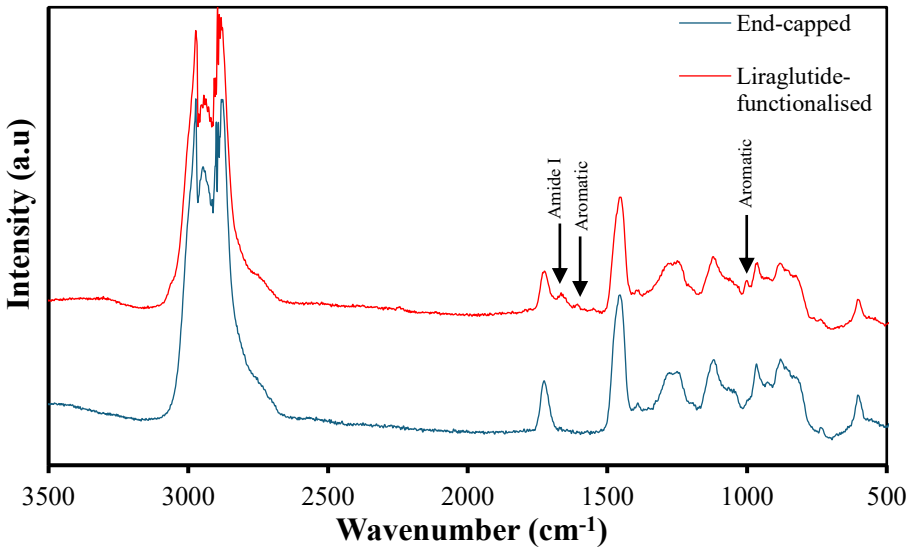


Figure 8.6: Averaged Raman spectra of Tris-end-capped AF-Tresyl-650M resins with (red) and without (blue) liraglutide immobilised onto the surface

The calculated and measured peptide, resin and column values required for the estimation of the second osmotic virial coefficient from the self-interaction chromatography (SIC) retention volumes are tabulated in Table 8.1. The accessible surface area, ϕ_{1r}, and the effective phase ratio, ϕ_{3r}, were calculated through interpolation of reverse size exclusion data of Dextran molecules through the same resin performed by DePhillips et al. [214]. The surface density, ρ_s, and the peptide excluded volume, B_{HS}, were calculated as outlined in Section 5.2.5.1.5.

Table 8.1: Parameters used for calculation of liraglutide osmotic second virial coefficient (B_{22}) *from SIC retention factors*

Parameter	Value
Porosity (P)	0.811
Accessible surface area, ϕ_{1r} (m^2 /mL)	21.3
Effective phase ratio, ϕ_{3r} (m^2 /mL)	18.7
Column Volume, V (mL)	0.177
Peptide Molecular Weight, M_w (kDa)	3.75
Peptide Monomer Radius, r (nm)	1.04
Peptide Excluded Volume, B_{HS} (mL)	1.88×10^{-20}
Volume Immobilised, V_i (mL)	0.0162
Surface Density, ρ_s (molecule/m^2)	2.78×10^{17}

Once the column had been packed with liraglutide fixated beads and thoroughly flushed, initial tests were performed to analyse the packing quality and establish the retention behaviour of the sample. Injections of 2% Acetone showed that the column was well-packed with a Gaussian peak with minimal tailing and a tailing factor of between 1.00 and 1.05 (see Figure 8.8 for an example elution profile). 0.1 mg/mL liraglutide samples in water were prepared and filtrated through 0.22 µm low-protein binding cellulose acetate filters for use as the injected sample. The aggregation state of the peptide was checked at hourly intervals via dynamic light scattering. Initially monomers with a hydrodynamic radius of 1.04 nm were observed as outlined in Figure 8.7a. However, after just 6 hours, noticeable peaks with a hydrodynamic radius of 2.94 nm were observed, representing the initial formation of oligomers (see Figure 8.7Figure 7.5b). Therefore, in future experiments the liraglutide sample used in self-interaction chromatography was changed for a fresh sample after 6 hours minimise any observed elution behaviour changes that

would be caused by the presence of oligomers. This was reasonably wasteful so careful preparation of 0.5 mL of liraglutide sample had to be carried out to minimise wasting expensive sample.

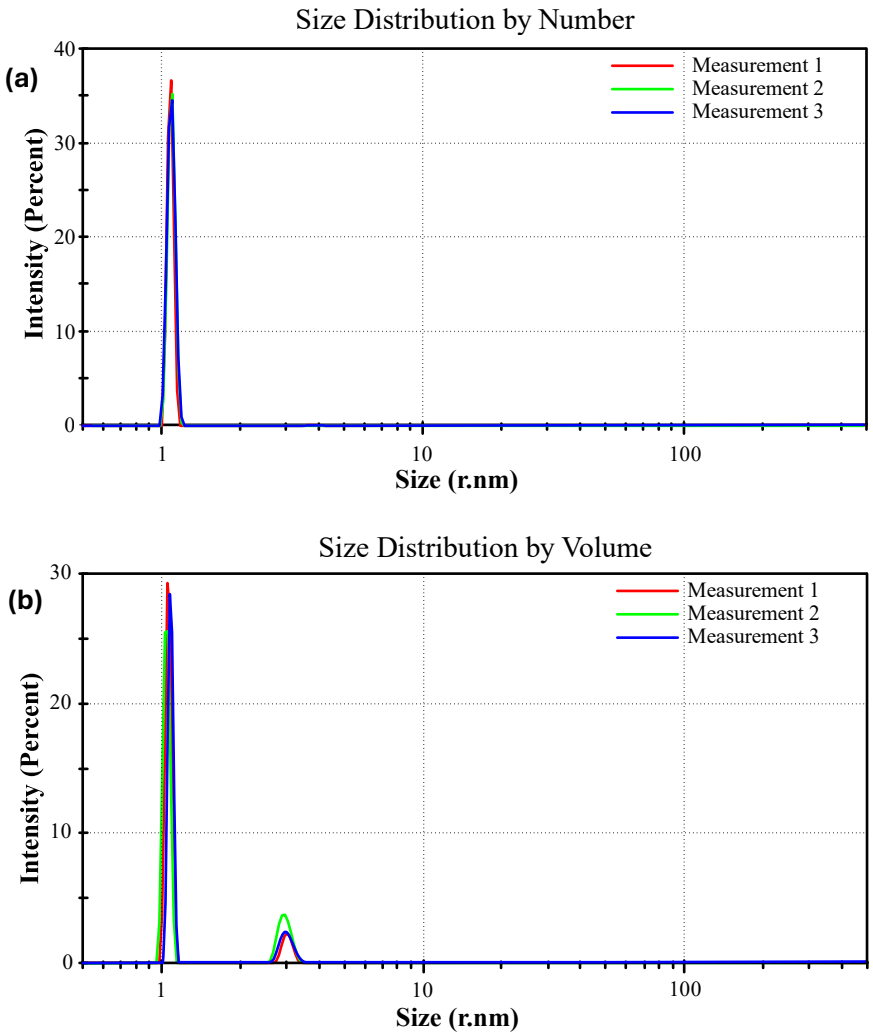


Figure 8.7: Example dynamic light scattering measurements of 0.1 mg/mL liraglutide in water *self-interaction chromatography sample (a) immediately after dissolution and (b) after 6 hours (see Figure 12.11 in Supplementary Materials for correlograms)*

Conversely, liraglutide samples displayed higher tailing, although still within the acceptable range, with a tailing factor between 1.35 and 1.40. Furthermore, as may be observed in Figure 8.8, liraglutide showed a lower or similar retention volume than the acetone void volume marker species for both the dead column and the liraglutide fixated column. This was to be expected based on the DLS and SEC results showing that the liraglutide oligomerises in solution. As such,

the oligomers would be size excluded from the stationary phase resulting in a lower retention volume than acetone, with the larger oligomers, previously observed at pH 6 eluting before the smaller ones at pH 8. However, this may also be indicative of weak peptide-peptide interactions and even perhaps repulsion between the mobile phase and immobilised peptides, indicating peptide stability.

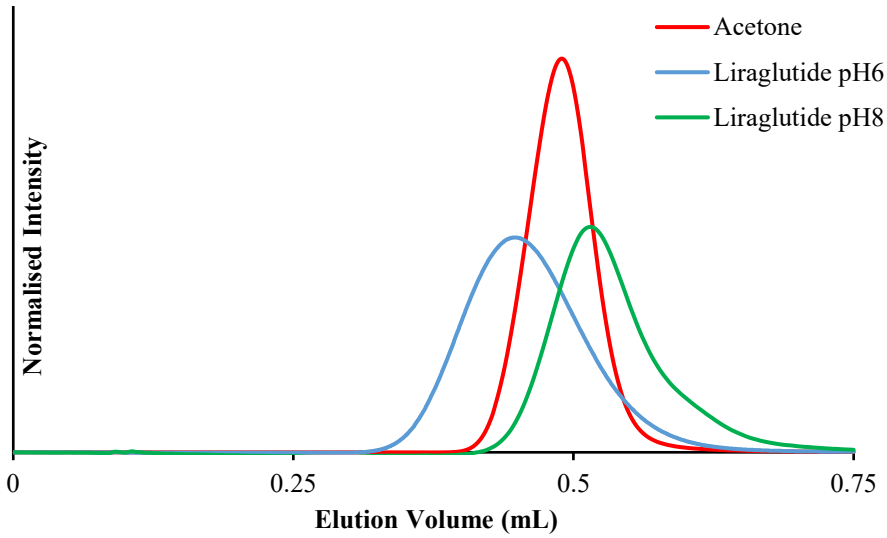


Figure 8.8: Example elution profiles of 2% Acetone solution (red) and 0.1 mg/mL liraglutide through a liraglutide fixated stationary phase at pH 6.0 (blue) and pH 8.0 (Green)
(20 mM sodium phosphate mobile phase, 0.05 mL/min flowrate and 37 °C Temperature)

The second osmotic virial coefficients (B_{22}), as estimated via self-interaction chromatography, across a range of simulated intestinal conditions are plotted in Figure 8.9. Since more positive B_{22} values (highlighted in blue in Figure 8.9) are indicative of stronger repulsive forces between species, and so reduced propensity to aggregate, the trend observed appeared to directly contradict that displayed by the SEC and DLS results. Since oligomers were formed upon dissolution under all conditions, it would be expected that under all conditions that there would be a highly negative B_{22} value estimated. Furthermore, it would also be predicted that at pH <7.0, where larger oligomers were formed, and at pH <6.5, where there was a propensity to aggregate, according to DLS and SEC data, the values of B_{22} estimated would be more negative than those at pH ≥7.0. However, as observed in Figure 8.9, the opposite was observed. This was assumed to

be primarily because the stationary phase surface was functionalised with liraglutide monomer species, whilst the species in the mobile phase were oligomer species. Therefore, the interactions that were being measured were not monomer-monomer, as typically measured by SIC, but were between monomer and oligomer species. Therefore, the parameter being estimated was not exactly B_{22}, but more similar to a second osmotic virial coefficient between two different species (B_{23}) as discussed by Hedberg et al. [358].

Two primary contributing factors led to the estimated B_{22} values. The first explanation is that since the oligomer species present at pH ≥7.0, as measured via DLS and SEC, were smaller oligomers than those at pH <7.0, they had a greater relative surface area for interaction with the immobilised liraglutide on the stationary phase. Furthermore, the smaller oligomers were also not size excluded from the resin volume to the same degree and so were able to access more of the stationary phase surface. The same effective phase ratio was used for all calculations. It was estimated assuming the surface accessible for three coeluted species, since this is the typical value applied in SIC. This was the approximate size of the individual oligomers according to DLS measurements. If a different phase ratio were applied for each condition, based on DLS measurements of the oligomer species hydrodynamic size, it would prevent the method from being truly predicative. Therefore, the method would be rendered somewhat obsolete. A second explanation as to the resultant B_{22} values appearing to contradict the DLS and SEC results is hypothesised based on the findings of Bothe et al.. They suggest that the two different oligomer species are in equilibrium with one another [193]. Therefore, despite the smaller (7- or 8-mer) oligomer being favoured at pH ≥7.0, in the presence of increased monomer concentration, (i.e. that immobilised on the surface of the resin which is incapable of self-associating into an oligomer structure), the equilibrium would be expected to shift to the larger (12- or 13-mer) oligomer. Conversely, for the larger oligomer species, the interaction strength with the monomer would be expected to be relatively weaker as it is not expected to readily associate with the monomer. However, this explanation does not account for the aggregation that was observed at

pH <6.5, unless the aggregation occurred from the self-association of the oligomer species and not from interaction with monomer species. If the alternative is the case (i.e. aggregation at pH <6.5 occurred from monomers associating with the larger oligomer), then strong retention and highly negative B_{22} values would have been expected under these conditions.

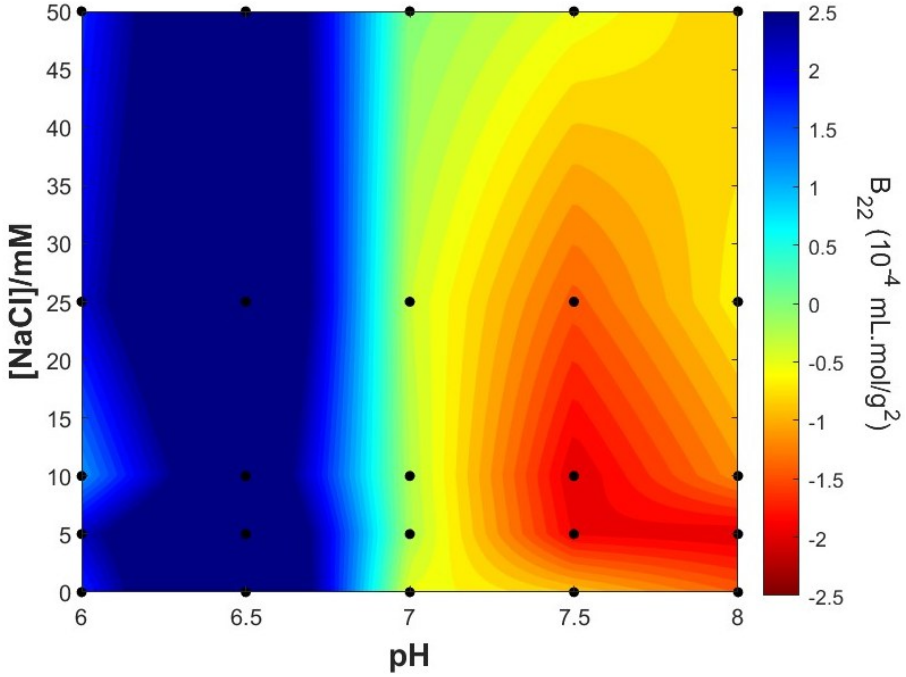


Figure 8.9: Plot of the Second Osmotic Virial Coefficient of liraglutide
at 37 °C under various simulated intestinal conditions as estimated via SIC (±1.1 x 10^{-5} mL.mol/g^2)

8.3.2. Insulin Stability

8.3.2.1. Dynamic Light Scattering

Dynamic light scattering (DLS) measurements of insulin in 20 mM sodium phosphate buffers simulating intestinal conditions showed that despite gently mixing until a transparent solution remained, the peptide did not instantaneously dissolve. Instead, as well as monomer and oligomer peaks, an apparent large aggregate peak was also observed (see Figure 8.10a for an example measurement exhibiting the various peaks). This observation was increasingly apparent at lower pH values approaching the isoelectric point of insulin (pI ~ 5.4), with a decreasing fraction of the species measured in solution having a radius of <5 nm (an arbitrary point selected, with no species close to it in size but representing species smaller than a 100-mer and so those

which may be classified as oligomers rather than larger aggregates), as may be seen in Figure 8.11. (Due to the presence of large species in the measured samples that were dependent on the sample preparation and had the potential to sediment, there was a relatively high error of ±10% for the data in Figure 8.11a).

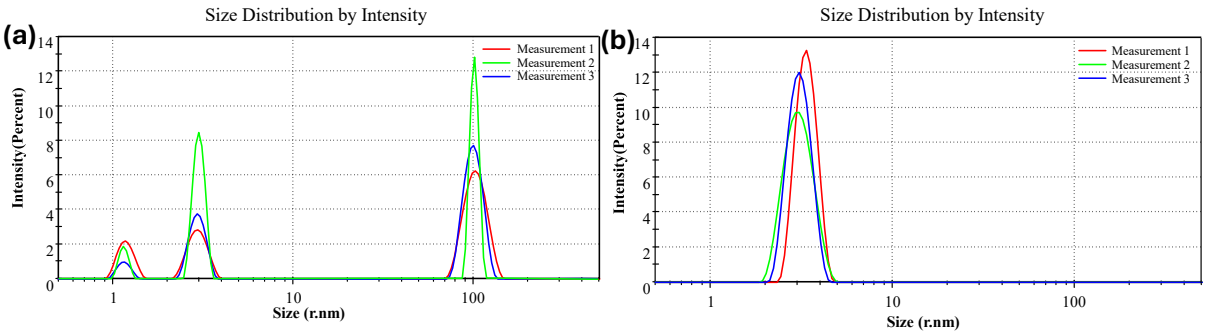


Figure 8.10: Example DLS measurements of 0.5 mg/mL insulin
in pH 7.0 20 mM sodium phosphate with 0 mM sodium chloride (a) immediately and (b) 24 hours after dissolution (see Figure 12.12 in Supplementary Materials for correlograms)

Furthermore, increasing salt concentrations up to approximately 25 mM NaCl increased the rate of dissolution of insulin and so the fraction of oligomer species present. However, at concentrations of sodium chloride higher than 25 mM, this trend reversed. This was likely due to the initial increasing salt concentration "salting-in" the insulin species, enabling their dissolution by reducing the strength of attractive electrostatic forces between neighbouring peptides. This phenomenon and its effects has previously been discussed in detail by Baldwin [173]. However, at even higher concentrations, the increased concentrations of ions may shield all electrostatic interactions and enable attractive hydrophobic interactions to become the more dominant forces between the dissolving sample peptides, thereby slowing the rate of their dissociation. However, as may be observed in Figure 8.11b and c (and supported by the example 24 hour measurement in Figure 8.10b), at pH ≥6.5 the insulin, after sufficient time, was able to fully dissolve, dissociating into smaller oligomer species, which remained stable over the 24-hour incubation. These findings were in agreement with those from Dimitrov et al. who showed the slow dissolution kinetics of insulin crystals under stagnant conditions [380]. Below pH 6.5

though, the large aggregates remained present throughout the course of the incubation period, showing that insulin is not readily soluble under these conditions at 0.5 mg/mL. The reduced charge under conditions approaching the isoelectric point resulted in poorer observed solubility, due to reduced electrostatic repulsive forces resulting in reduced propensity to separate into monomer or smaller oligomer species. This resulted in the individual species maintaining more open structural conformations with more increased hydrophobic site availability.

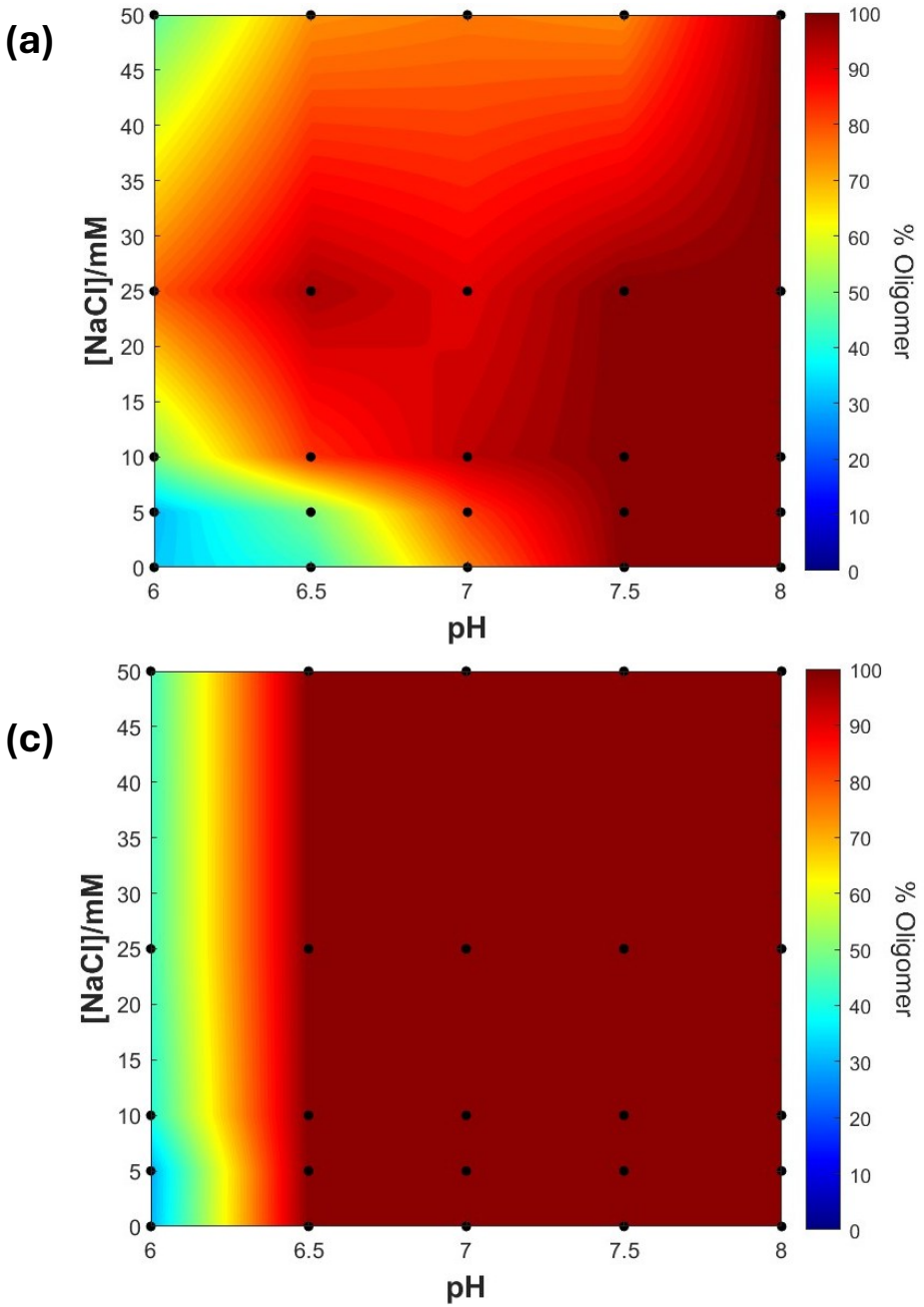


Figure 8.11: Plots of fraction of insulin species with a hydrodynamic radius <5 nm

at 0.5 mg/mL concentration and 37 °C in various 20 mM sodium phosphate buffers (±10%) after (a) 0 hours, (b) 6 hours and (c) 24 hours as measured via Dynamic Light Scattering

Therefore, attractive hydrophobic forces became the more dominant stabilising interactions between neighbouring insulin species resulting in the formation of larger aggregates.

The species formed upon full dissolution predominantly had an average hydrodynamic radius of between 2.5 nm and 3.0 nm (see Figure 8.12), indicative of an insulin hexamer according to literature values [370, 371, 381, 382]. The hexamer structure is comprised of three dimers, typically stabilised by two divalent metal cations (most prevalently Zn^{2+}). In fact, the propensity to form hexamers relative to dimers being 400 times greater in the presence of the Zn^{2+} ions than in their absence [370]. Figure 8.12 does not show measurements recorded immediately after dissolution, due to the high variance in accordance with the incomplete dissolution as represented in Figure 8.11a. Furthermore, the species recorded for pH <6.5 are only those <5.0 nm with the larger aggregates excluded on account of the high degree of variance in their size and to focus on the relative difference in the oligomeric structures. Across the pH range, it can be observed that the largest species on average were detected in the pH 6.5–7.0 range. However, aside from at pH 6.0, where the concentration was lower due to the incomplete dissolution, the average sizes reported (~3.0 nm) appear to be approximately the size of (or slightly larger than) that of the hexamer. However, Das et al. also discuss the existence off-pathway oligomer intermediates, which may be of a similar size to the hexamer species [310]. This suggests that the reduction in average radius as the pH was increased to pH 8.0 was due to the polydispersity of the results, with an increasing fraction of the sample population taking a smaller oligomer structure reducing the resultant average size reported. This was assumed to be the case since larger stable oligomeric structures of insulin have not commonly been reported. Additionally, there appeared to be little variation in the size of the species formed with sodium chloride concentration, implying that the relative stability of the apparent hexamer formed was not due to monovalent sodium ions but rather was a function of the concentration, charge state and the temperature. In fact, the primary difference in the conditions carried out in this experiment, relative to size exclusion data reported by Hvidt et al who recorded the presence of

predominantly dimers in a zinc-free environment [370], is the temperature being at 37 °C rather than at 21 °C. The results of this study contradict this since, whilst smaller species (R_h ~1.5-2.0 nm) were present in some measurements (see Figure 8.10a), the average species size recorded of the polydisperse peak in the Zn^{2+} free environment was approximately that of the hexamer. Therefore, one of the following must be assumed to have caused the difference in result:

1. The sample preparation method and experimental temperature resulted in a change of conformation that favoured the formation of a hexamer or a similarly sized off-pathway oligomer, as proposed by Das et al. [310], and not the dimer nor larger aggregates
2. The sample insulin contained divalent metal cationic species such as zinc or calcium such that upon dissolution the concentration was sufficient to stabilise the formation of the hexamer species
3. There was consistent error in the Zetasizer data processing software such that there was an overestimation of the species size by approximately 1 nm. (However, this seems highly unlikely given that DLS data aligned well with SEC data for liraglutide).

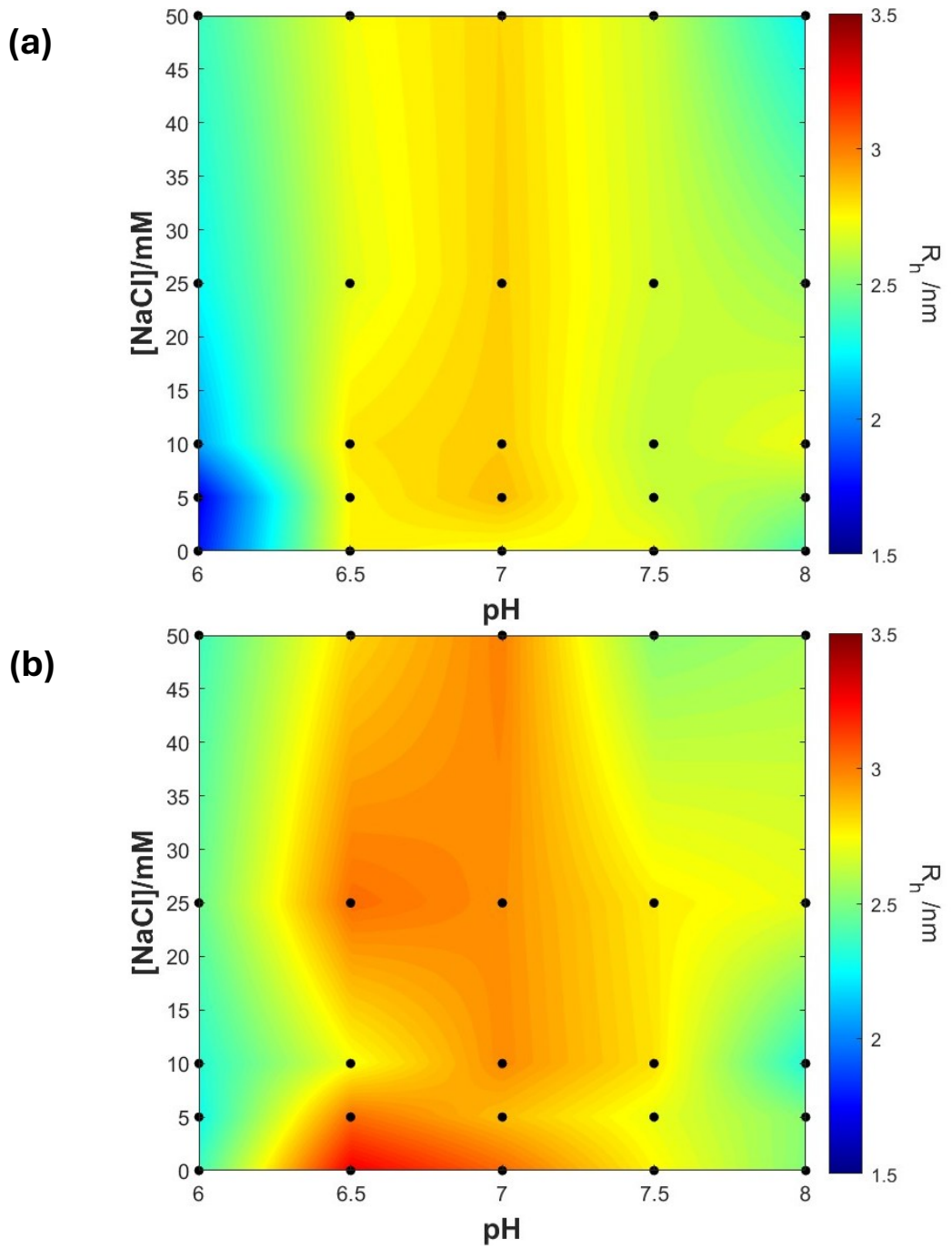


Figure 8.12: Plots of radius of insulin species (<5 nm) at 0.5 mg/mL concentration *and 37 °C in various 20 mM sodium phosphate buffers simulating intestinal conditions (±10%) after (a) 6 hours and (b) 24 hours (±0.15 nm)*

8.3.2.2. Size Exclusion Chromatography

Due to DLS measurements showing the presence of large aggregate samples and large oligomers, a TSKgel G3000SW$_{XL}$ size exclusion column, typically used for the analysis of monoclonal antibodies and large protein aggregates, was selected for analysis of insulin

aggregation. Furthermore, measurements were not taken at pH <6.5, due to the recurring presence of aggregates in DLS measurements and a resultant increase in column pressure during initial sample testing, suggesting that aggregates present may be blocking the column packing. Initial analysis of the size exclusion chromatography elution profiles of insulin samples incubated in various 20 mM sodium phosphate buffers at 37 °C, indicated similar results to DLS with a sharp main peak eluted after approximately 9.5 to 10 mL, followed by a much smaller peak at approximately 11.0-11.5 mL (see Figure 8.13 for an example elution profile). These peak ranges correlated to species with molecular weights in the range of 28-43 kDa and 8.2 kDa to 12.4 kDa, respectively. This was in good agreement with the DLS data, with the higher molecular weight species correlating to an oligomer in the 5-mer to 7-mer range and the smaller species accurately describing the dimer molecular weight (based on the molecular weight of a human insulin monomer being 5808 Da).

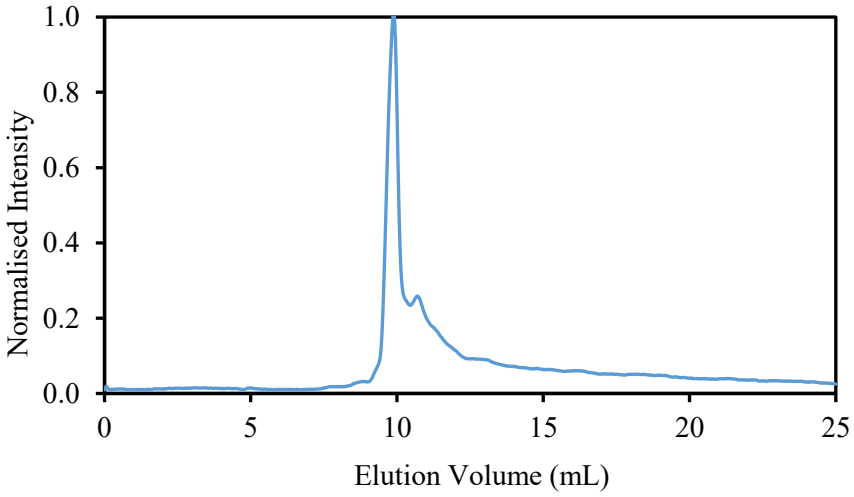
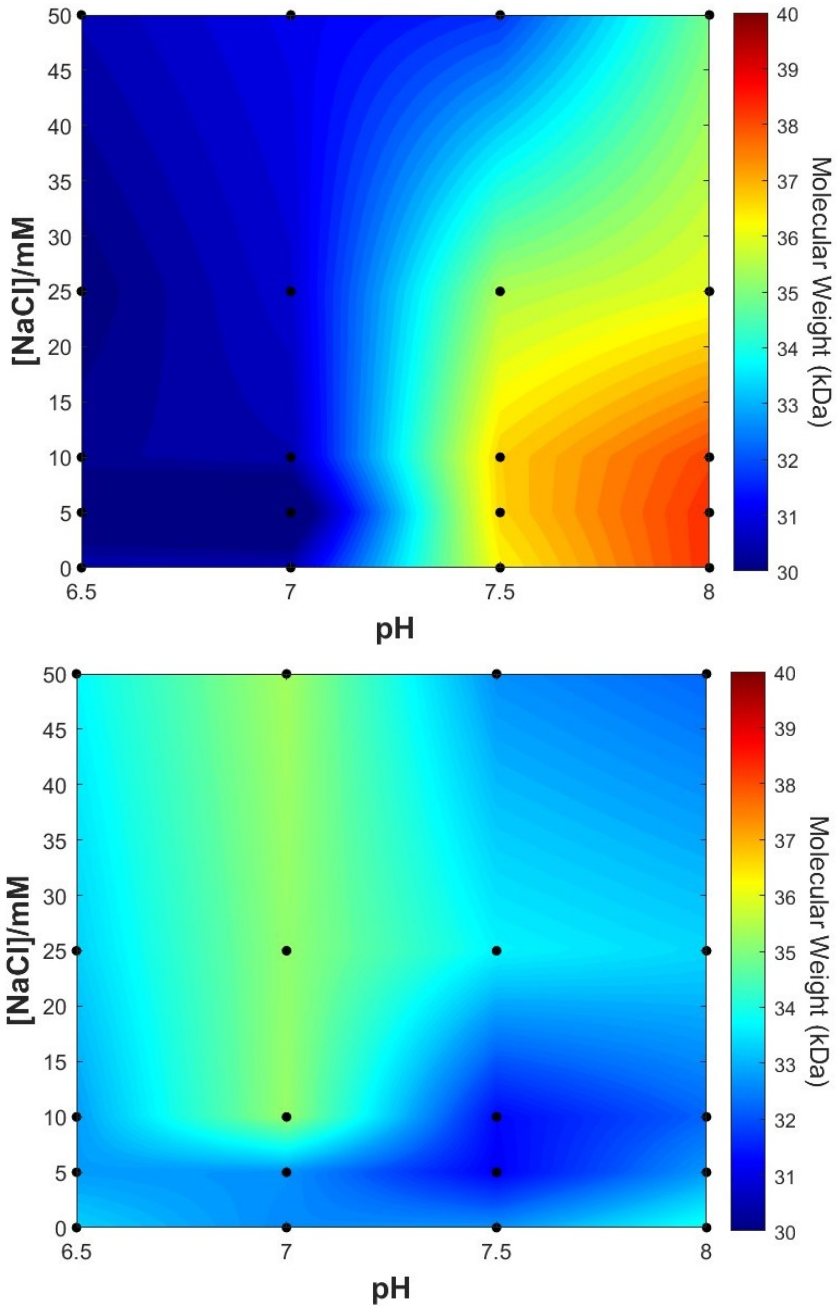


Figure 8.13: Example SEC elution profile of 0.5 mg/mL insulin
incubated in pH 7.0 20 mM sodium phosphate at 37 °C for 24 h (1.0 mL/min 50 mM pH 7.0 sodium phosphate + 150 mM sodium chloride)

Size exclusion chromatography enabled a more accurate analysis of the non-aggregate species upon initial dissolution (see Figure 8.14a) due to the absence of the larger aggregate species either fragmenting due to the high-pressure shearing forces within the SEC column or not being

eluted at all. It was observed that upon initial dilution the hexamer species formed at higher pHs and lower sodium chloride conditions were larger in size, likely due to the stronger electrostatic repulsion forces between their dimer constituents, the theory of which was discussed in Section 2.3.6.1. Although, the structure formed may be an off-pathway oligomer as proposed by Das et al., the structure of which may shift with time or varying conditions [310]. However, after 24 hours (see Figure 8.14c) incubation time the trend was in closer agreement with what was observed during DLS experiments with the largest species being observed at approximately pH 7.0. At higher pHs, the relative fraction of hexamers was smaller with an increased number of smaller species. This resulted in increased peak tailing and a shifted overall apparent peak time, and thus a smaller estimated species size. Conversely, at lower pHs the hexamer would likely be more stable due to the reduced repulsion forces of the peptides, resulting in a more compact structure and a smaller apparent hydrodynamic radius and hence a reduced species size. However, it should be noted that the variation may also be in part as a result of the standard measurement error which was approximately ±0.1 mL for the triplicate retention volume measurements or a molecular weight error of approximately ±2 kDa.



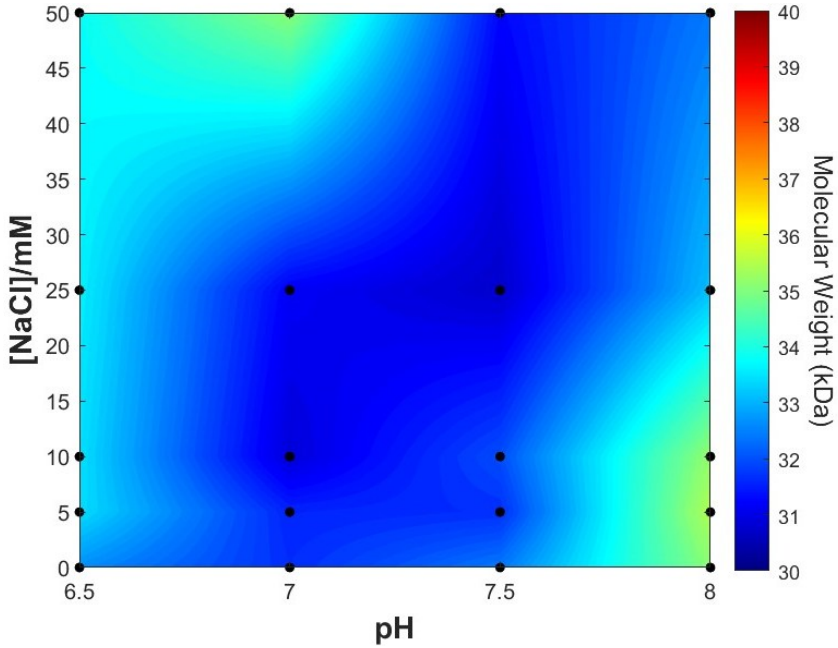


Figure 8.14: Plots of SEC estimated molecular weights of 0.5 mg/mL insulin *after being incubated at 37 °C under various simulated intestinal conditions for (a) 0 hours, (b) 6 hours and (c) 24 hours (±0.8 kDa)*

8.3.2.3. Self-Interaction Chromatography

Insulin immobilisation was performed as in Section 5.2.3.1 with initial and final concentration of the peptide in the immobilisation solution being measured to be approximately 2.03 mg/mL and 0.53 mg/mL by UV spectroscopy. This resulted in an estimated 7.50 ± 0.16 mg insulin immobilised onto the surface of 100 mg of resin (i.e. 75.0 ± 1.6 mg/g.resin), which represented a 16.1% molar surface coverage of the surface tresyl-conjugated sites (assuming a one peptide per tresyl site). Raman spectroscopy confirmed that insulin had been immobilised onto the resin surface, with the same three small additional peaks being observed as for liraglutide. Since more insulin was immobilised onto the surface, it had been expected that the peaks would be clearer than for liraglutide, but, as may be seen in Figure 8.15, the additional peaks were perhaps even less distinct. Furthermore, an additional peak in the region of 425-550 cm^{-1} had been expected to be observed indicating the presence of the disulfide bond of which insulin has three. Its absence and the reduced peak sizes relative to liraglutide emphasises that peptide detection may be highly measurement site-dependent, especially given a measurement consisted of a grid of 16 sites and were repeated in triplicate across three separate beads. This was to be expected to a degree, due to a high percentage of the resin surface area being within the porous network of the resin rather than the external surface, meaning that the majority of the peptide would not have been on the surface readily available to be detected.

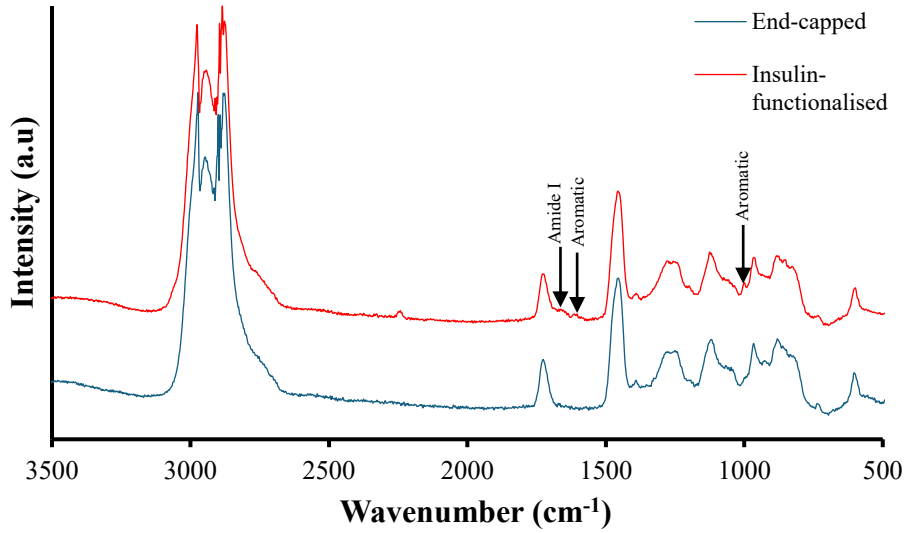


Figure 8.15: Averaged Raman spectra of tris-end-capped AF-Tresyl-650M resins with (red) and without (blue) insulin immobilised onto the surface

The calculated and measured peptide, resin and column values required for the estimation of the second osmotic virial coefficient from the self-interaction chromatography retention volumes are tabulated in Table 8.1. The accessible surface area, ϕ_{1r}, and the effective phase ratio, ϕ_{3r}, were calculated through interpolation of reverse size exclusion data of Dextran molecules through the same resin as proposed by DePhillips et al. [214]. The surface density, ρ_s, and the peptide excluded volume, B_{HS}, were calculated as outlined in Section 5.2.5.1.5. The insulin monomer hydrodynamic radius was an average of measurements reported in literature by Bromberg et al. [383] and Wang et al. [384].

Table 8.2: Calculated parameters applied in the calculation of the osmotic second virial coefficient (B_{22}) of insulin

Parameter	Value
Porosity (P)	0.811
Accessible surface area, ϕ_{1r} (m^2 /mL)	21.2
Effective phase ratio, ϕ_{3r} (m^2 /mL)	17.7
Column Volume, V (mL)	0.177
Peptide Molecular Weight, M_w (kDa)	5.81
Peptide Monomer Radius, r (nm)	1.34
Peptide Excluded Volume, B_{HS} (mL)	3.99×10^{-20}
Volume Immobilised, V_i (mL)	0.0310
Surface Density, ρ_s (molecule/m^2)	2.56×10^{17}

0.1 mg/mL insulin samples in water were prepared and filtrated through 0.22 µm low-protein binding cellulose acetate filters for use as the injected sample. The aggregation state of the sample was checked at hourly intervals via dynamic light scattering. Initially monomers with a hydrodynamic radius of 1.26 nm were observed as outlined in Figure 8.16a. However, after approximately 10 hours, noticeable peaks with a hydrodynamic radius of 2.62 nm were observed, representing the initial formation of oligomers (see Figure 8.16b). Therefore, in future experiments the insulin sample was changed for a fresh sample after 10 hours minimise any observed elution behaviour changes that would be caused by the presence of oligomers.

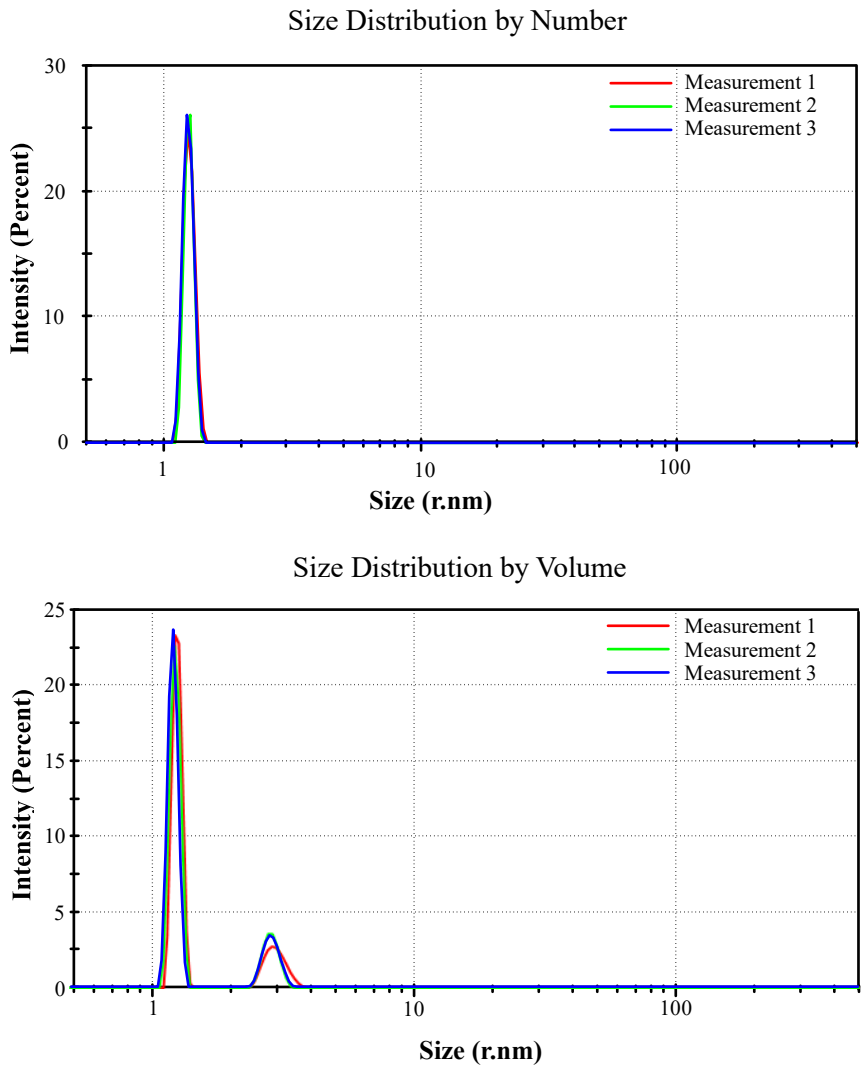


Figure 8.16: Example dynamic light scattering measurements of 0.1 mg/mL insulin in water *self-interaction chromatography sample (a) immediately after dissolution and (b) after 6 hours (see Figure 12.13 in Supplementary Materials for correlograms)*

Initial insulin dead column experiments were conducted to establish the interaction strength of the peptide with the end-capped stationary phase. As observed in the example elution profiles in Figure 8.17b, across the range of conditions the insulin eluted with an acceptable gaussian distribution. However, increasing retention was observed as the acidity of the mobile phase increased, with more peak tailing suggesting that it was a result of weak non-linear stationary phase interactions rather than the species being smaller and therefore size excluded to a lesser extent. This was supported by the fact that at pH <7.0 the insulin had a greater retention volume than acetone injections. Furthermore, as the acidity increased, the injection profile detected by the pre-column spectrophotometer showed increasing peak splitting of the injected sample (see Figure 8.17a) suggesting the presence of poorly-diffusing, large aggregates. This hypothesis was supported by the decrease in the fraction eluted, estimated by comparing the peak area measured by the pre- and post-column detectors, at the more acidic pHs. At pH 8.0, the peak area ratio was 103% that of acetone, whilst at pH 7.0, where peak splitting began, it dropped to 96.7%, and at pH 6.0 to 82.3%, suggesting a large fraction of the injected insulin was in the form of poorly eluting aggregates or a conformation that formed strong interactions with the end-capped resin surface.

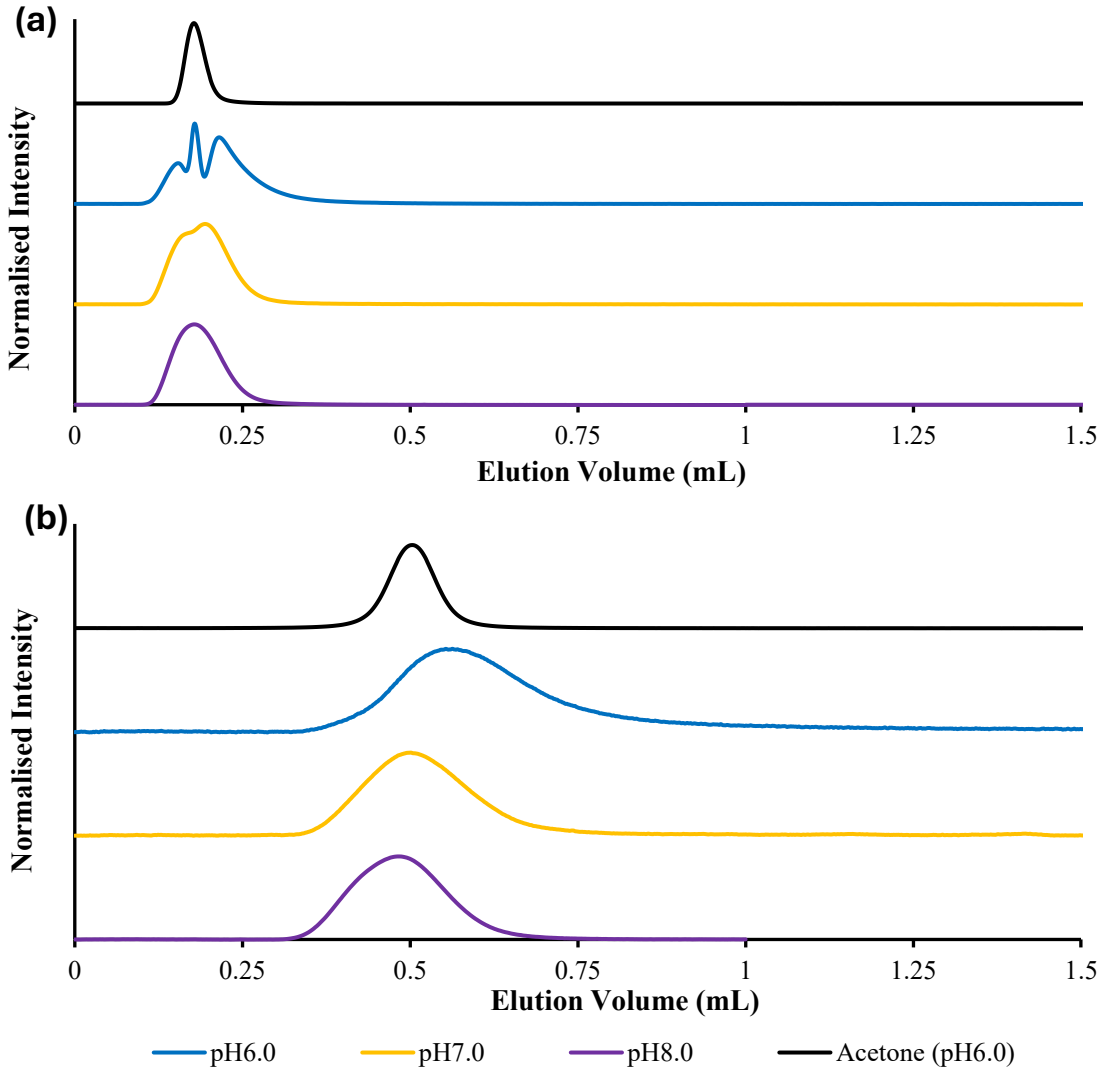


Figure 8.17: Dead column (a) injection and (b) elution profiles of 0.1 mg/mL insulin (unlabelled) or 2% v/v Acetone
with 20 mM sodium phosphate mobile phases at example pHs to simulate the intestinal environment at 37 °C and a flowrate of 0.05 mL/min

Upon completion of column packing with insulin-fixated AF-tresyl-650M beads and thorough flushing, initial packing quality analysis was performed. Injections of 2% Acetone showed that the column was well-packed with a Gaussian peak with minimal tailing and a tailing factor of between 0.95 and 1.05. However, by contrast insulin elution varied significantly with mobile phase acidity as observed in Figure 8.18. At pH <7.0, there was an absence of an elution profile peak, indicating that the insulin injected, observed in the pre-column detector, was fully retained in the column, suggesting very strong interactions with the stationary phase functionalised insulin. This retention behaviour was confirmed when large elution peaks were observed when

the column was flushed with pH 8.0 mobile phase prior to the subsequent study condition, releasing the strongly retained species. The strength of retention interactions decreased significantly as the mobile phase acidity decreased from pH 7.0 to pH 8.0, indicated by narrowing peaks with lower retention volumes in Figure 8.18. This was likely on account of a change in charge state of His-B10 and His-B5 (pK_a~6.0), resulting in a change in conformation of the stationary phase monomers or mobile phase species, resulting in weaker peptide-peptide interactions. Protonation of the His-B10 residue has previously been suggested to result in electrostatic repulsions between adjacent monomers in insulin aggregate species [315]. Furthermore, peak-splitting was observed between pH 7.0 and pH 7.5, despite the injection profiles being similar to those observed during the dead column experiments, suggesting that the high concentration of insulin immobilised on the resin surface assisted in the formation of larger oligomer or aggregate structures. These species were either eluted first due to being size-excluded from the porous resin structure upon their formation, having less available surface immobilised insulin with which to interact or, due to having a slow formation time, were eluted more slowly than species retaining the initial injected conformation. However, at pH 8.0, due to the increased charge of the insulin species at the higher basicity (further from its isoelectric point), the stronger electrostatic repulsion forces resulted in weaker peptide-peptide interactions and thus a shorter retention time with an acceptable gaussian peak, as would be expected for a poorly-retained sample. Due to the absence of peak-splitting at pH 8.0, it can be assumed that its cause was not due to column overloading or column voids, as it would have been observed at all conditions. Therefore, it was assumed to be due to the presence of multiple conformational forms of the mobile phase peptide. Nevertheless, there was a high tailing factor in the range of 1.6-1.8, indicating that the analyte peptide interactions with the stationary phase insulin were non-linear.

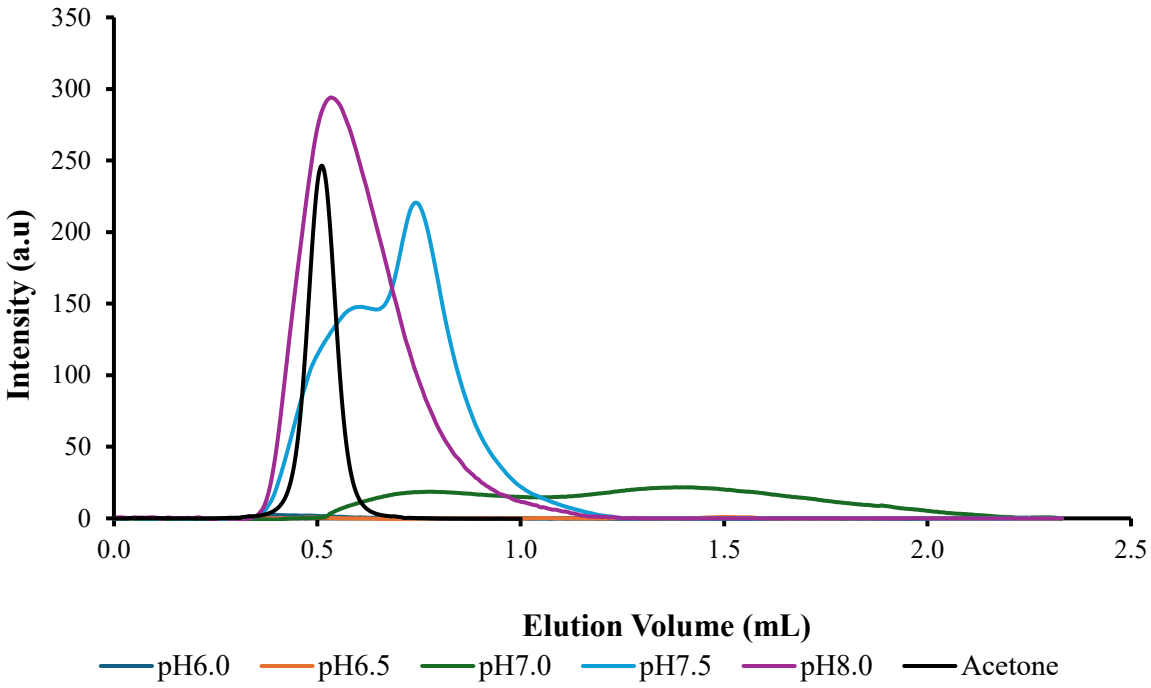


Figure 8.18: Example elution profiles of 0.1 mg/mL insulin through an insulin fixated AF-Tresyl-650M packed column
with 20 mM sodium phosphate mobile phases at various pHs to simulate the intestinal environment at 37 °C and a flowrate of 0.05 mL/min

Due to the absence of an elution peak for conditions pH <7.0, the second osmotic virial coefficient (B_{22}) could not be reasonably estimated using self-interaction chromatography under these conditions as the value produced was negative infinity. Furthermore, due to the large difference in retention volumes between pH 7.0 and pH 8.0, the resultant B_{22} values differed by orders of magnitude and therefore were not all plotted on one chart, since the trends in variance with changing salt concentration would be lost. Furthermore, due to the peak-splitting observed between pH 7.0 and pH 7.5, the second osmotic virial coefficient was calculated for both peaks to analyse if these peaks represented two different peptide conformations and the different trends in their interaction strength with the immobilised monomer.

As may be observed in Figure 8.19a and b, the two peaks observed at pH 7.0 and pH 7.5, displayed very similar trends in their second osmotic virial coefficients with changing salt concentration: increasing salt concentration showed stronger peptide-peptide interactions (more negative B_{22} values) due to increased charge shielding and so reduced electrostatic

repulsion between peptides. The second peak's value was more negative due to its higher retention volume, inherent with it being the second peak, but this difference in value remained relatively constant with changing sodium chloride concentration. This suggests that the ionic strength of the mobile phase affected the interaction strength of the two different mobile phase peptide conformations with the immobilised peptide by not the same relative amount but the same amount. Therefore, the two peaks shift overall as one peak with changing salt concentration, with little change in the relative splitting of the peak, implying that the mobile phase insulin had one conformation injected, some of which formed a second conformation upon interaction with the stationary phase insulin. This species was size excluded upon formation, hence changes in salinity have the same effect on interaction strength, as it was the same initial conformational species. Nevertheless, despite this the aggregational propensity (scale of the negativity of the B_{22} value), as observed in the scales in Figure 8.14a-c, appeared to be predominantly pH-dependent, with the salt concentration accounting for only minor differences. However, at pH 8.0, the effects of salt concentration were in complete contradiction to those at more acidic conditions, with increasing concentrations of sodium chloride resulting in more positive B_{22} values. This indicated that the immobilised or mobile phase insulin conformation changed as the pH increased from pH 7.5 to pH 8.0, likely caused by a deprotonation of the B1-Phe N-terminal amino group. The pH is too high for the structural change to be cause by protonation of Histidine groups as discussed by Noormagi et al. [315]. The resultant structure was perhaps such that the increase in Cl^- ions between 0 mM and 50 mM more effectively shielded the remaining local positively charged A1-Gly, B22-Arg and B29-Lys groups than the Na^+ groups shielded local negatively charge groups. Therefore, the result was that there was an overall net increase in electrostatic repulsion between the immobilised and mobile phase peptide groups reducing retention volume and so estimated B_{22} value.

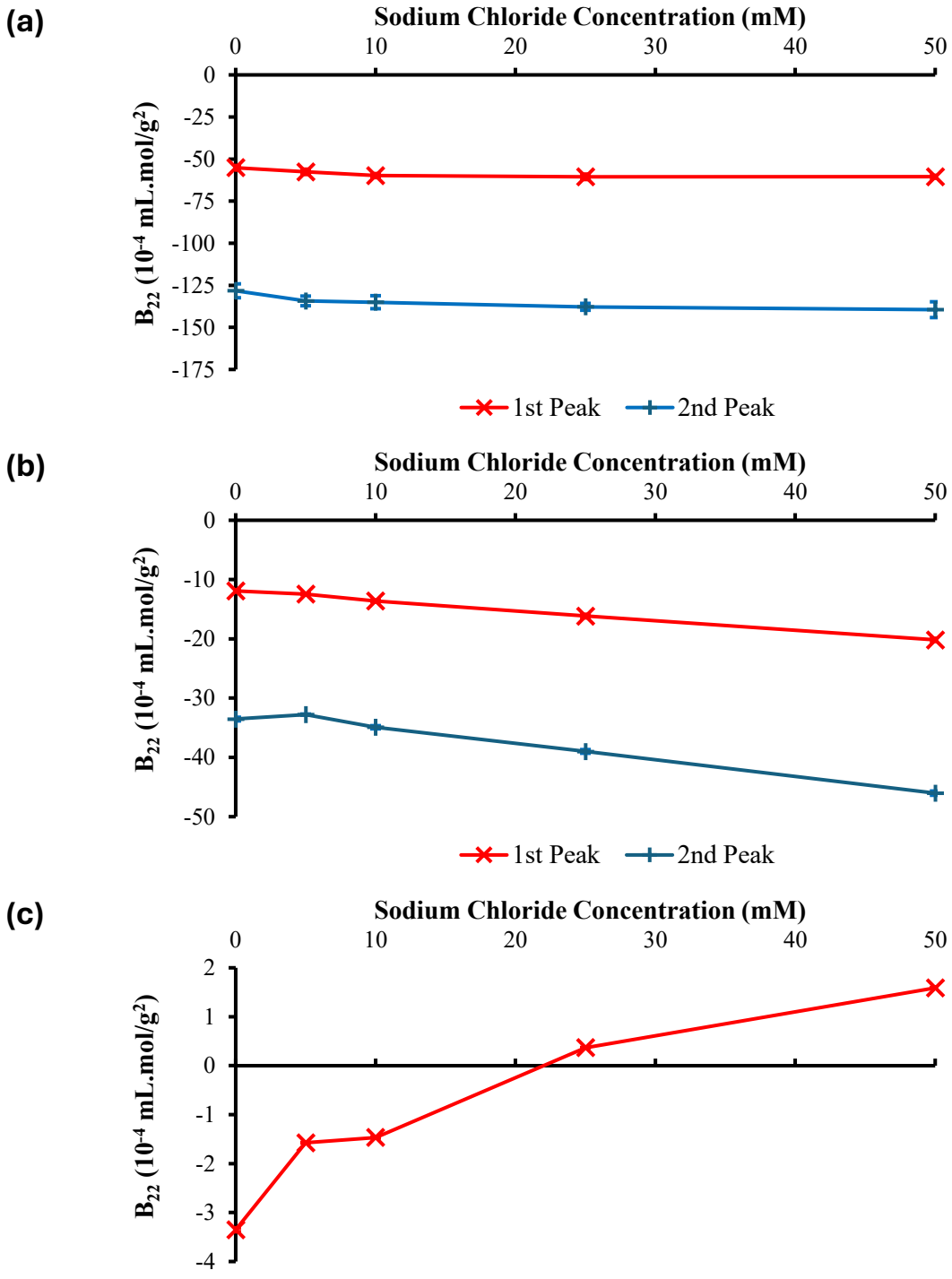


Figure 8.19: Plots of the change in the second virial coefficient of insulin with NaCl concentration

at (a) pH 7.0, (b) pH 7.5 and (c) pH 8.0 as measured via self-interaction chromatography (177 µL column, 0.05 mL/min flowrate, 20 mM sodium phosphate mobile phase, Temperature 37 °C)

8.4. Conclusions

The physical stability of liraglutide and insulin were measured under a range of simulated intestinal conditions from pH 6.0-8.0 and varying sodium chloride concentrations from 0-50 mM. DLS was applied to measure the radius of the species, which were correlated to molecular weights using SEC. The results were then compared to SIC measurements of the peptides' second osmotic virial coefficient (B_{22}).

Liraglutide was observed to form 12- or 13-mers with a radius of 3.0-3.5 nm instantaneously upon dissolution. However, repeated measurements made after incubation times of 6 hours and 24 hours found that at pH ≥7.0, the radius of the species began to decrease suggesting the oligomers began to transition to form 8-mers. Conversely, B_{22} values measured via SIC showed the opposite trend, with positive values at pH <7.0 indicating stability, and negative values at pH ≥7.0. This was suggested to be due to the B_{22} measurements indicating the interaction strengths between the monomer and the two different oligomers. The effect of sodium chloride concentration by contrast seemed to be much less than that of pH.

Insulin was observed to be poorly soluble forming large aggregates at pH <6.5, whilst almost entirely forming oligomers at all other conditions, predominantly hexamers with a radius between 2.5 nm and 3.0 nm. These hexamers were observed to be most stable under more basic conditions with sodium chloride concentration not having a major effect on their size. SIC measurements were only obtainable at pH ≥7.0 due to poor elution quality and were highly negative indicating a high propensity to aggregate, which was observed by the formation of oligomers at all conditions. However, the trends did not align well with what was observed via DLS and SEC.

The overall results indicated that SIC can be a useful tool for outlining the trends in physical behaviour of the peptides under simulated intestinal conditions. However, in this current state it requires prior detailed understanding of the aggregation behaviour of the peptide in order to

interpret the results. Therefore, unless reduced concentrations or formulations which maintain the monomeric structure of the peptide are employed, the method is likely not the most well-suited for studying readily aggregating or oligomerising peptides under a wide range of conditions.

CHAPTER 9

9. GASTRIC PROTEOLYTIC DIGESTION OF PEPTIDES VIA IMMOBILISED PEPSIN REACTORS

9.1. Introduction

Gastric delivery is not the typical route targeted by oral drug formulations primarily for three reasons: the wall of the gastric mucosa is thick thereby limiting the rate of drug absorption; the stomach itself has a low absorption area to volume ratio with a surface area of approximately 0.1 m^2 [226], resulting in a low concentration gradient for absorption; the highly acidic gastric conditions averaging pH <2 [229] under fasted conditions are not ideal for long term stability nor ensuring high hydrophobicity, and so drug permeability. Therefore, to bypass these constraints and target intestinal delivery, formulation strategies aim to protect the active pharmaceutical ingredient (API) from the gastrointestinal (GI) environment. Such strategies include enteric coatings which are typically composed of pH-dependent polymers, which maintain their structure at low pH but upon exposure to increased pH, dissolve to release the API payload [50]. Similarly, in oral peptide drug delivery research, a sodium alginate and calcium chloride hydrogel was shown by Ren et al to effectively shield insulin from the harshly acidic gastric conditions [385]. Nevertheless, since of the approximately 10 orally delivered peptide therapeutics available, only the oral formulation of semaglutide, Rybelsus, perhaps the most prevalent within media currently, targets a gastric absorption pathway. Understanding the barriers that the gastric conditions may provide to oral peptide therapeutic absorption remains of relevant research importance [68]. Included within these barriers to the gastric delivery of oral peptide drugs are the proteolytic enzymes present within the stomach, which act to degrade proteins and peptides into smaller peptides and amino acids. The primary of these enzymes is pepsin which is secreted

by gastric chief cells and mainly functions as an endopeptidase cleaving at the carboxyl side of hydrophobic residues as discussed in Section 5.1.5 [40].

Therefore, developing methods to measure the rate of peptide proteolytic degradation under simulated gastric conditions for therapeutic peptides targeting gastric peptide delivery remains of importance within research. Typically, the rate of proteolytic digestion of sample proteins or peptides is measured in solution via exposing the specific substrate to pepsin under gastric conditions followed by timepoint sampling and immediate deactivation of the enzyme via precipitation, prior to centrifugation and further analysis [259]. One of the drawbacks of this technique is the fact that pepsin, being a protein, has a similar structure to the substrate peptides and proteins, and so will likely precipitate under similar conditions, thereby resulting in either incomplete separation or product loss prior to analysis at insufficient or excessive precipitant concentrations, respectively. The use of immobilised enzymes in a packed bed reactor set-up removes the necessity for post reaction separation or time-controlled sampling, since, provided there is minimal sample resin interaction. The flowrate dictates the residence time and automatically separates the substrate from the enzyme. Additionally, immobilisation of the enzymes allows their repeated reusage under identical conditions whilst maintaining their activity [386] potentially reducing experimental error. Moreover, they may also show greatly improved stability with a storage time of 3-6 months [250]. The major drawbacks, however, are that immobilisation may result in a reduction in activity or change in selectivity of the enzyme, thereby rendering the immobilised enzyme as an inferior model to solution reactions [250]. However, previously developed immobilised pepsin reactors have been shown to maintain reasonable activity and maintain similar selectivity [348, 386-388].

Immobilised pepsin reactors have been used sparingly in the field for studying the stability of oral peptide therapeutics targeting gastric delivery. Nevertheless, workflows have previously been described to develop successful immobilised enzyme reactors, specifically focussing on

matching enzymes and resins based on their properties [389]. Therefore, three primary aims were established for the work carried out in this Chapter: to establish a method of pepsin immobilisation that enable high pepsin loadings onto a resin suitable for prolonged exposure to the harsh simulated gastric environments; to determine qualitatively and quantitatively how immobilisation affected, most importantly, the selectivity, but also the activity and stability of the enzyme; and finally, to examine the peptide-dependency of the proteolytic activity of pepsin in a mock study. This current work aims to build upon previous work carried out using immobilised pepsin reactors, but attempts to develop an improved immobilisation procedure and methodology.

9.2. Materials and methods

All experiments in this Chapter were carried out using a propylamino silica resin as described in Section 5.1.7. Within the same section, descriptions of the monoethanolamine, glutaraldehyde, potassium phosphate, sodium sulfate, sodium citrate, citric acid, hydrochloric acid and formic acid used are described as well as the buffer preparation procedure. Additionally, the porcine pepsin used as the proteolytic species throughout is that which is described in Section 5.1.5, whilst all the substrate peptides are described in Sections 5.1.1 - 5.1.4.

Pepsin immobilisation and resin end-capping procedures were performed as outlined in Sections 5.2.3.3 and 5.2.3.4. Raman Spectroscopy of the resin surface was carried out as in 5.2.3.5. The column packing procedures are outlined in Section 5.2.4.2. Chromatographic experiments were carried out on a Nexera high-performance liquid chromatography (HPLC) system from Shimadzu, the configuration of which is described in 5.2.1. The immobilised and solution pepsin reactions were performed as outlined in Sections 5.2.6.1 and 5.2.6.2, respectively. Liquid chromatography-mass spectrometry (LC-MS) analysis of the degradation products was performed as outlined in Section 5.2.6.3.

9.3. Resin selection and pepsin immobilisation

Covalent immobilisation via the surface amine groups onto the resin particles was selected as the preferred strategy for the immobilisation of porcine pepsin primarily due to the ease with which orientation of the enzyme may be controlled to ensure accessibility of the active site. As may be observed in Section 5.1.5, the only surface arginine, lysine and N-terminal residues of porcine pepsin are all located on the opposite side of its 3-dimensional conformation to the active site. Therefore, immobilisation via these groups ensures the active sites are orientated towards the mobile phase and available for substrate interaction. Moreover, a Schiff's base mechanism of immobilisation allows for these surface amine groups to be targeted during immobilisation without operating at conditions that would result in denaturation of pepsin. The additional reasons for choosing a covalent immobilisation strategy are the increased stability it provides relative to other immobilisation strategies [270] and the prior expertise that the Surfaces and Particle Engineering Laboratory (SPEL) at Imperial College has in protein immobilisation.

There are three primary desired properties of a resin for use as the support phase of an immobilised pepsin reactor: first and most importantly, stability under the highly acidic simulated gastric conditions to which it will be consistently exposed; second, a surface that will interact minimally with substrate peptides under the aforementioned conditions; third, high porosity and surface coverage to allow for high loading of sterically available and active enzyme species. Typical polymer resins used as substrates for enzyme immobilisation are comprised of polymeric matrices, such as the styrene-divinylbenzene copolymer from Sunresin or the methacrylate resins from Purolite. These resins, whilst being highly non-interactive, typically have a stability range from pH 3-10, which is not sufficient for gastric condition testing. Therefore, a propylamino-silica resin was selected as it fulfilled the previous desired criteria: first, silica is highly stable under a wide range of conditions including resistant to acidic conditions; second, given the lowest pK_a values of chromatographic silanol groups is estimated to be in the 3.5-4.6 range, since the operating pH is lower, the groups will be protonated and provide reduced relative

interactions; third, the resin selected had a high amine loading of 1 mmol/g with a surface area of 550 m^2/g and a microporous structure with a pore size of 60 Å. In addition to this, the silanol groups that may remain available during immobilisation, may provide additional sites for physisorption, as is the primary route of immobilisation for enzymes onto mesoporous silicates, and thus increase potential pepsin loading [389].

Initial immobilisation experiments were conducted in triplicate and resulted in an average pepsin coverage of 185 mg/g.resin ± 5.4 mg/g.resin according to the change in concentration of the reaction solution as measured via UV-Vis spectroscopy. This constituted a molar surface coverage of only 0.54% of the available surface amino groups. Raman spectroscopic analysis of the resin surface was subsequently carried out, after the resin was sufficiently washed and dried to remove any potential physisorbed enzyme, to confirm the competition of a covalent immobilisation reaction. The results comparing a Glutaraldehyde-activated (without subsequent pepsin immobilisation) and pepsin-functionalised propylamino silica resins are outlined in Figure 9.1. There were two primary differences between the two spectra: first, the presence of a small peak at approximately 1650 cm^{-1} in the glutaraldehyde-fixated spectrum, representing the C=N of the imine bond formed between glutaraldehyde and the amine surface group; second, the presence of the large peak at approximately 990 cm^{-1}, due to aromatic breathing vibrations, indicating the presence of Phe, Tyr or Trp residues. The imine bond peak was absent from the pepsin-functionalised resin, which is expected due to its degradation to an amine via $NaCNBH_3$. However, it would have been expected that there would be a small amide bond peak in a similar location. Nevertheless, the spectra provided clear confirmation of completion of the immobilisation reaction.

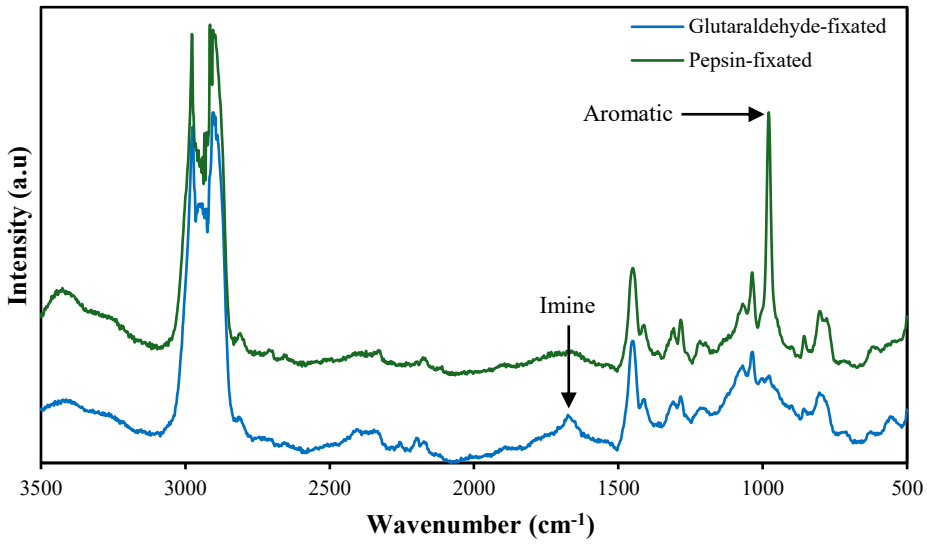
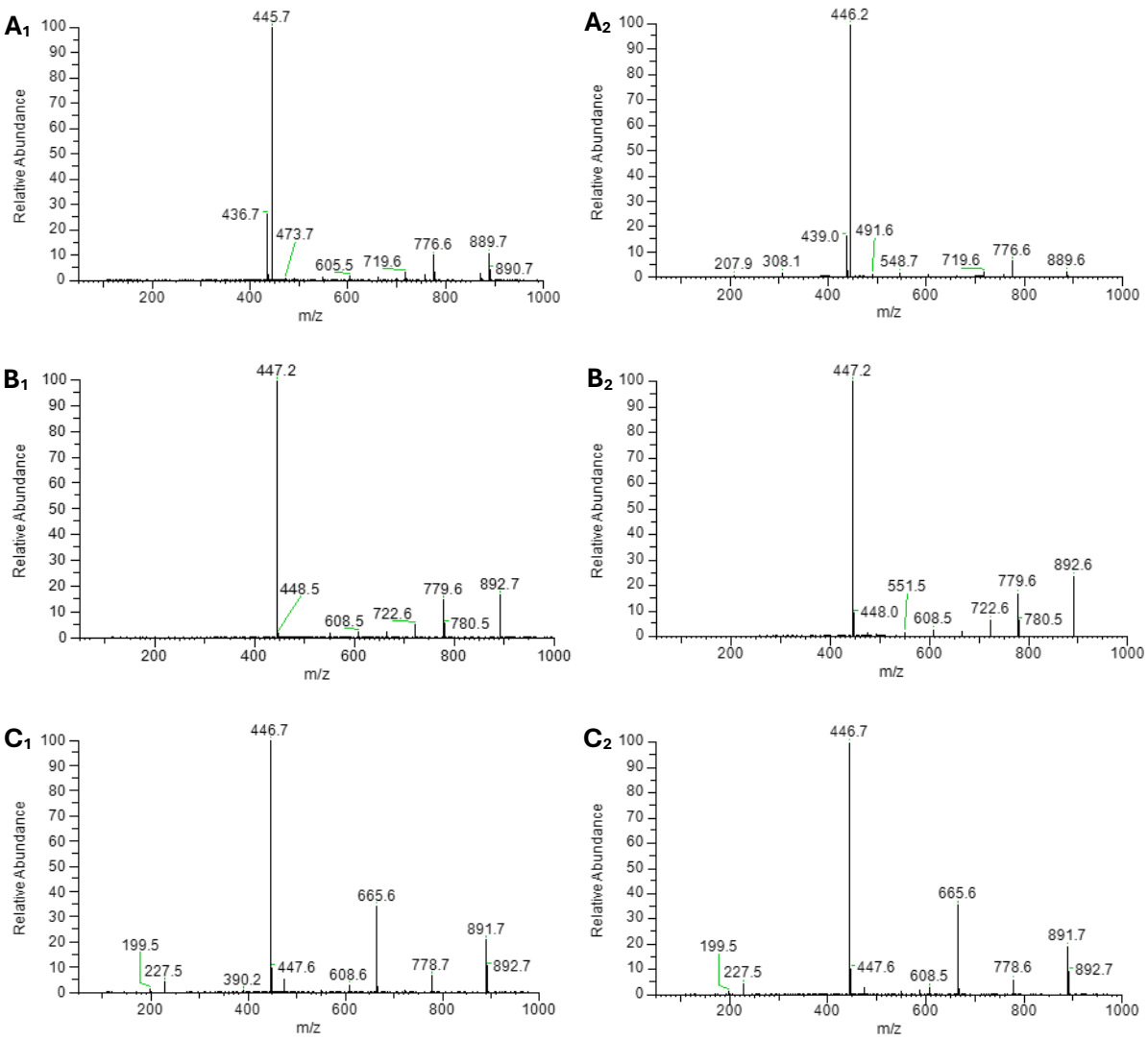


Figure 9.1: Raman Spectra of Glutaraldehyde-activated and pepsin-functionalised propylamino silica resin
with labelled peaks to establish completion of reaction

9.4. Initial qualitative analysis of immobilised pepsin activity

After packing approximately 0.45 g (±0.01 g) resin into the column (approximately 0.341 mL empty volume), initial experiments aimed to first establish that the immobilised pepsin had maintained activity and that the activity observed was qualitatively similar to that in solution reactions. Therefore, the reversed-phase high-performance liquid chromatography (RP-HPLC) peptide retention standards (see Section 5.1.3) were selected as qualitative reference standards for the proteolytic activity of the immobilised pepsin. This selection was made due to the fact that pepsin, as discussed in Section 5.1.5, is an endopeptidase, the action of which is specifically promoted via the presence of aromatic or hydrophobic residues, including Leucine. Therefore, with the increasing Leucine content within the peptide structure from P1 to P6, an increasing rate of proteolytic degradation would be expected, due to the increasing number of proteolytic sites. However, due to the decreasing length of the peptide chain from P1 to P5 and pepsin's preference to bind to seven-residue sequence sites, the sequences may allow for investigation of this aspect. Finally, P2-P6 each also contain a Phenylalanine group at the penultimate residue, since pepsin is an endopeptidase and the C-terminal is an arginine group which due to its positive

charge may inhibit the proteolytic rate, it would not be expected for this group to react. However, P1 does not contain this Phenylalanine group and therefore may be used as a comparison to P2 in this event. Based on these reasons, this hydrophobic peptide library was deemed a suitable selection for analysing the qualitative activity of the immobilised pepsin relative to in solution. Each were first eluted at 0.1 mL/min through a column packed with propylamino-silica end-capped with ethanolamine (hereafter referred to as a dead column) but without pepsin functionalisation to establish whether the peptides were eluted through the resin without any reaction or shearing, which may be conflated with proteolytic degradation. Mass spectra of the collected fractions were then compared to that of the pure diluted peptides.



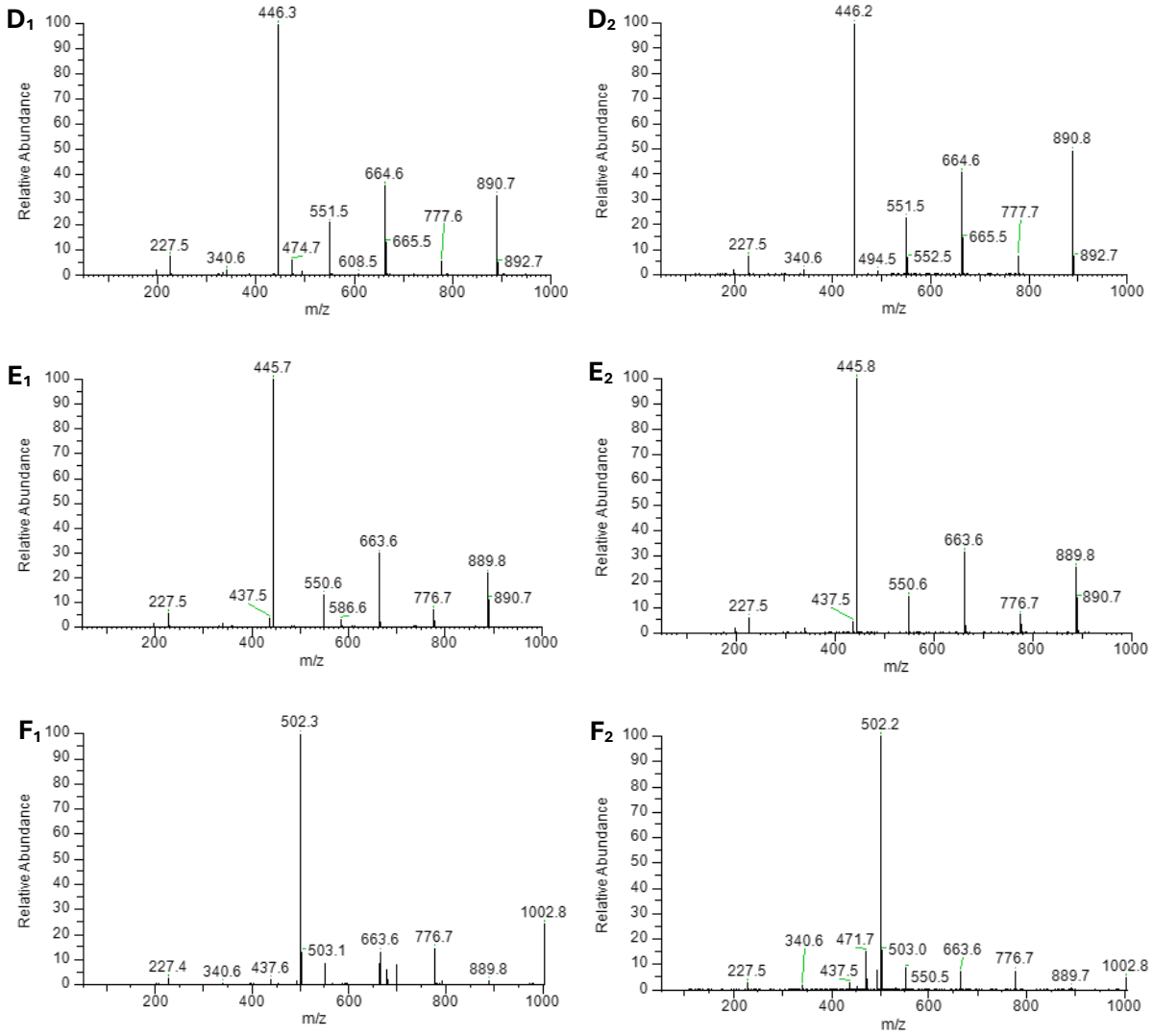


Figure 9.2: Mass spectra of hydrophobicity standard peptides P1-P6 (A-F)
with subscript 1 representing the pure initial sample dissolved in pH 2.0 formic acid and subscript 2 the sample after elution through a 30 mm L x 4.6 mm ID column packed with end-capped propylamino-silica resin at 0.1 mL/min in pH 2.0 HCl

For each of the hydrophobicity standard peptides, a single peak was observed for both the initial sample and after elution through a dead column. The mass spectra corresponding to these peaks are illustrated in Figure 9.2. The initial observation from each of these spectra is that the peptides fragmented during the nanoelectrospray ionisation process. This was confirmed by the fact that each of the peaks corresponded to a fragment of the peptide sequences, but were coeluted with the peptide through the PepMap C18 column, confirming that they were not present in the original sample. The primary peak in each mass spectrum was the $[M+2H]^{2+}$ peak clearly

indicating the peptide sample, whilst the fragment peaks were each present in the MS^2 spectra. Nevertheless, this may have complicated future spectral analysis of the proteolysis of peptides in the event that a degradation product coeluted with the original peptide substrate. However, when comparing the samples eluted through the dead column to the initial samples mass spectra, the fragmentation patterns remained relatively consistent. Therefore, due to this consistency, a significant change in these fragmentation patterns may be reasonably considered as the result of the coelution of proteolytic degradation products. Furthermore, the lack of difference between the mass spectra and the observation of only a single elution peak illustrates that the dead column resin alone did not react or catalyse any reaction involving the peptide substrates which resulted in the changes in molecular weight.

As previously mentioned in Section 9.3, an initial pepsin immobilisation procedure was carried out resulting in 0.5 g of resin with a surface coverage of approximately 185 ± 5.4 mg/g.resin pepsin as determined via UV-Vis spectroscopic concentration measurements of the reaction solution. This resulted in a column with approximately 83.4 mg (± 1.5 mg) pepsin at a concentration of 245 mg/mL (± 8.1 mg/mL). Initial injections of the peptide substrates were carried out to establish whether any reaction was observed, and any qualitative trends were captured by the immobilised pepsin column. For each of P1-P3 peptides only one clear peak was eluted, the mass spectra of which are outlined in Figure 9.3. As may be noted from the similarity of the results in Figure 9.2, it was clear that the immobilised pepsin caused no observable proteolytic degradation of the mobile phase P1-P3 peptide substrates. This aligned well with experiments in solution which similarly showed no reaction of the same peptide substrates even after 30 minutes of reaction, as also illustrated in Figure 9.3. This was expected proteolytic behaviour as P1, P2 and P3, the most hydrophilic of the hydrophobicity standard library, have the fewest leucine residues which are susceptible to degradation from pepsin and emphasised that the immobilisation of pepsin via this method did not result in a drastic decrease in specificity. The only primary difference in the spectra being the reduced sample to noise ratio of the post-

solution reaction P1 spectra. This was due to the high hydrophilicity of P1 and the increased hydrophobicity of the injection solution (80% ethanol for pepsin precipitation) into the LC-MS resulting in a lower fraction of the sample being concentrated in the PepMap trap column. This change in injection solution also explains the difference in the relative fragmentation patterns and charge distribution of the peptide within the mobile phase of the solution-reaction mass spectra relative to those previously in Figure 9.2 and the column-reaction spectra.

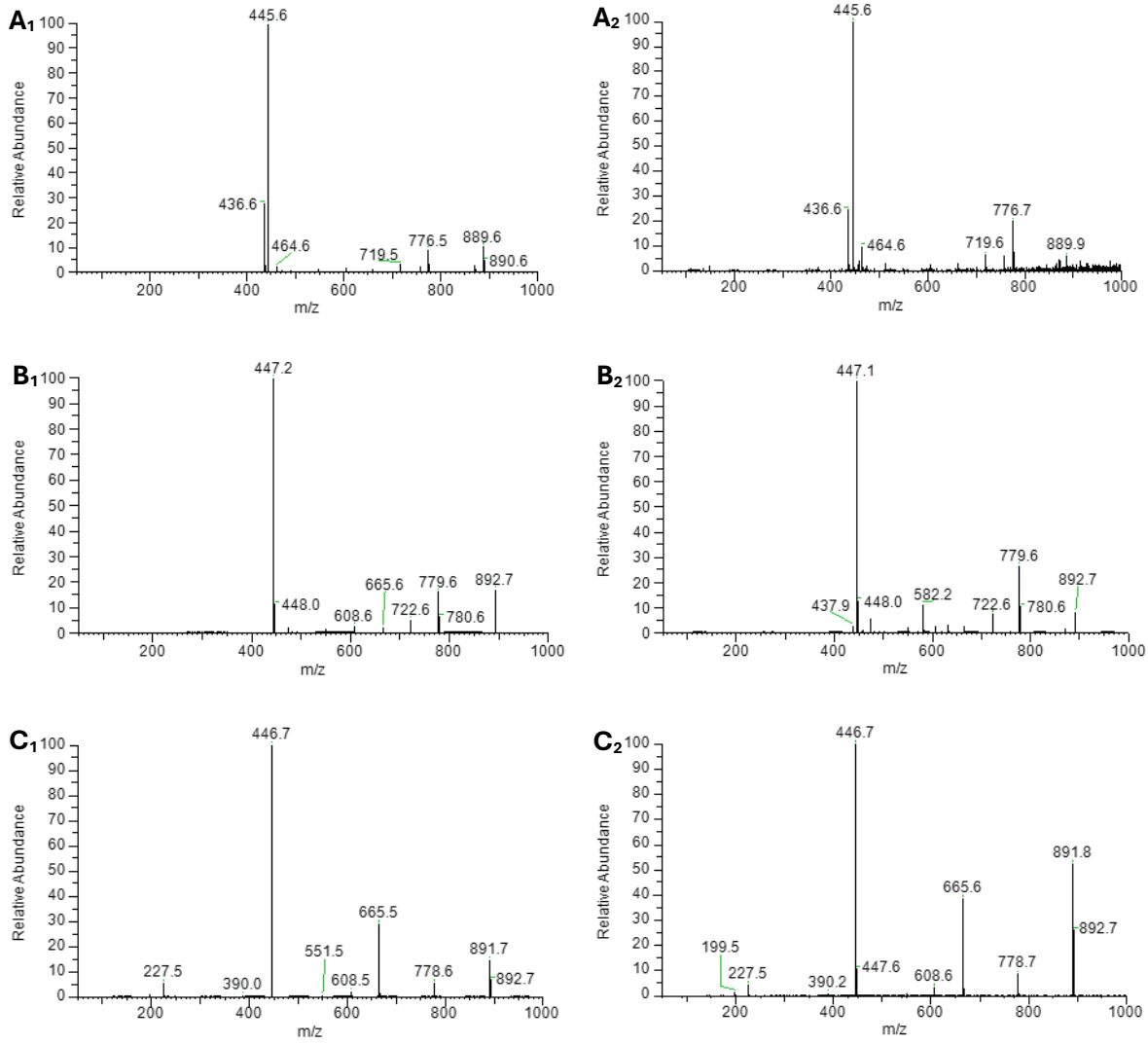


Figure 9.3: Mass spectra of P1-P3 (A-C)
after elution at 0.1 mL/min in pH 2.0 HCl through a 30 mm L x 4.6 mm internal diameter (ID) column packed with silica resin functionalised with 185 mg/g.resin pepsin (subscript 1) and after 30 minutes of reaction at 0.1 mg/mL in 0.5 mg/mL pepsin solution in pH 2.0 HCl (subscript 2)

By contrast P5 and P6 showed complete reaction after 30 minutes incubation solution and after elution at 0.1 mL/min through the immobilised pepsin reactor with insufficient remaining substrate to readily discern from spectral noise. The labelled base peaks for P5 and P6 for each condition are outlined in Figure 9.4. (Henceforth, due to the number of mass spectra, base peaks with labelled peaks shall primarily be presented). It should be noted that whilst the LDFR and LLDFR groups in Figure 9.4A_1 and Figure 9.4B_2 respectively appear to be mislabelled, since both m/z values refer to the sequences with one fewer Leucine residues, in fact both short peptides also fragmented during the nanoelectrospray ionisation process. This resulted in the fragment without one Leucine (approx. -113.1 m/z) being the major peak in both cases. The mass spectra of the two peaks are outlined in Figure 9.5. For both peptides, as would be expected, both the immobilised and free pepsin preferentially cleaved with Leucine at the P_1 position (see Section 5.1.5 for reference) and more specifically at an internal residue site, since pepsin is an endopeptidase. This was exemplified by the primary peaks for the P5 proteolytic degradation products being that of the tri- and tetra-Leucine peptides (see Figure 9.4A). However, the corresponding DFR and (in the case of the free enzyme reaction and both P6 degradation sets) the LDFR peptides were not observed as readily. For the DFR tripeptide, this was suggested to be due to the fact that it is relatively hydrophilic and therefore was less readily concentrated in the PepMap C_{18} trap column. However, for the LDFR peptide, this may have also been the case for the solution reactions but due to the increased hydrophobicity of the injection solvent the sample may have less readily been retained in the column. Additionally, the tetrapeptide may have been lost during the pepsin precipitation procedure due to its increased hydrophilicity. The limitations of the Velos Pro setup are that it was poorer at identifying peaks with an m/z value significantly less than 200, despite having a range of 50 to 2000, due to the increased noise. However, despite Leucine monomer peaks being indeterminable, due to the absence of their C-terminal partner fragments from the mass spectra, it could safely be assumed that neither the immobilised nor free pepsin readily degraded P5 or P6 at that position or any other unobserved fragment position.

Therefore, despite the relative abundance of the various fragments differing between the immobilised and solution degradation products, due to the presence of the same proteolytic degradation products within each sample, it could be concluded that there was not a significant change in the selectivity of pepsin upon immobilisation via this method. The difference in relative abundances may have been in part due to the more complex procedure required for solution reactions.

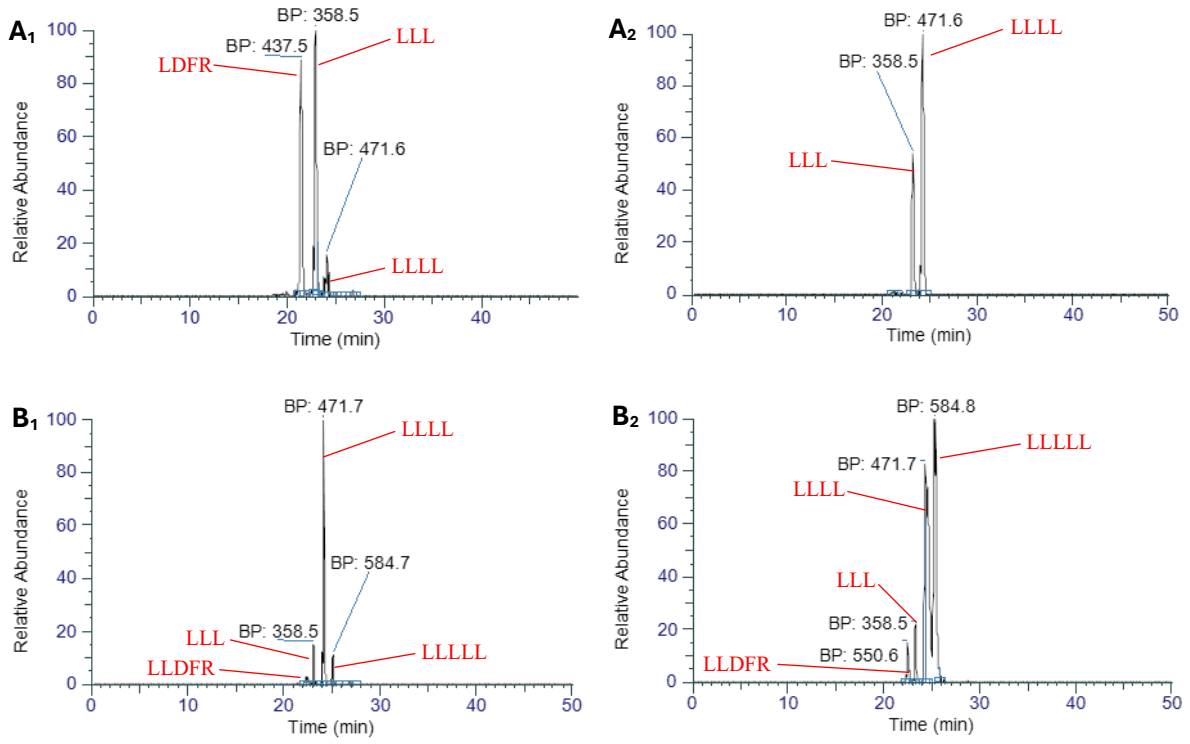


Figure 9.4: Mass spectra of P5-P6 (A-B)
after elution at 0.1 mL/min in pH 2.0 HCl through a 30 mm L x 4.6 mm ID column packed with silica resin functionalised with 185 mg/g.resin pepsin (subscript 1) and after 30 minutes of reaction at 0.1 mg/mL in 0.5 mg/mL pepsin solution in pH 2.0 HCl (subscript 2)

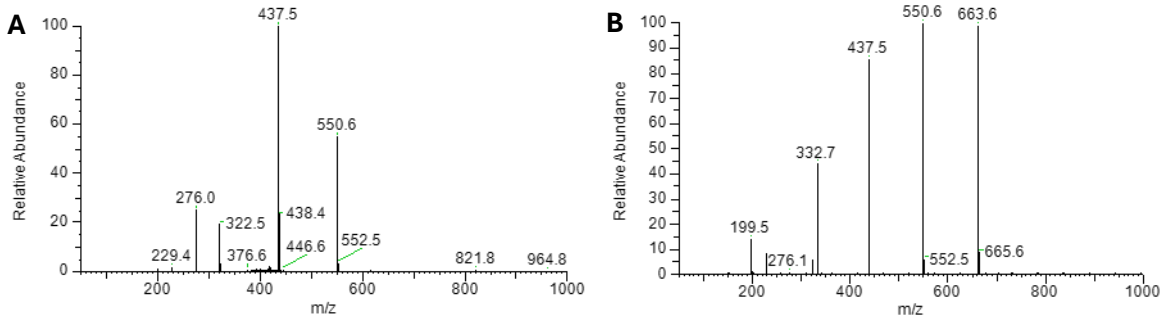


Figure 9.5: Mass spectra to illustrate the fragmentation patterns of proteolytic degradation products

(A) LDFR from P5 and (B) LLDFR from P6 highlighting that the primary peak is not the same as the $[M+H]^+$peak

P4 by contrast showed very minimal proteolytic degradation at pH 2.0 after both elution at 0.1 mL/min (approx. 3.5 minute reaction time) through the immobilised enzyme column and after exposure at 0.1 mg/mL to 0.5 mg/mL pepsin solution for 30 minutes. Under both conditions, a split peak was observed which predominantly consisted of the original peptide in both parts of the peak, as represented in Figure 9.6. However, in the second peak in both cases, there was an additional fragment with an m/z of 587.6, which corresponded to the first 6 residues of the octamer as highlighted in Figure 9.6A_2 and Figure 9.6B_2. This difference may be noted by comparison to the spectra displayed in Figure 9.2D_1 and D_2, in which this additional mass fragment was absent. From this, two qualitative conclusions may be drawn. Firstly, Glycine at the P_1' position inhibited the proteolytic cleavage at the Leu-3 residue relative to Asp at the P_1' in P5 and P6, resulting in no tri-Leucine product being observed. Secondly, whilst the relative amount of reaction in the immobilised column was lower, the same observation of very limited reaction was still observed, highlighting that the immobilisation procedure resulted in no significant reduction in selectivity of the enzyme. Since there was significantly more pepsin within the immobilised column relative to free solution but the reaction time was significantly shorter, initial observations were purely qualitative hence there is no mention of the relative activity of the enzyme.

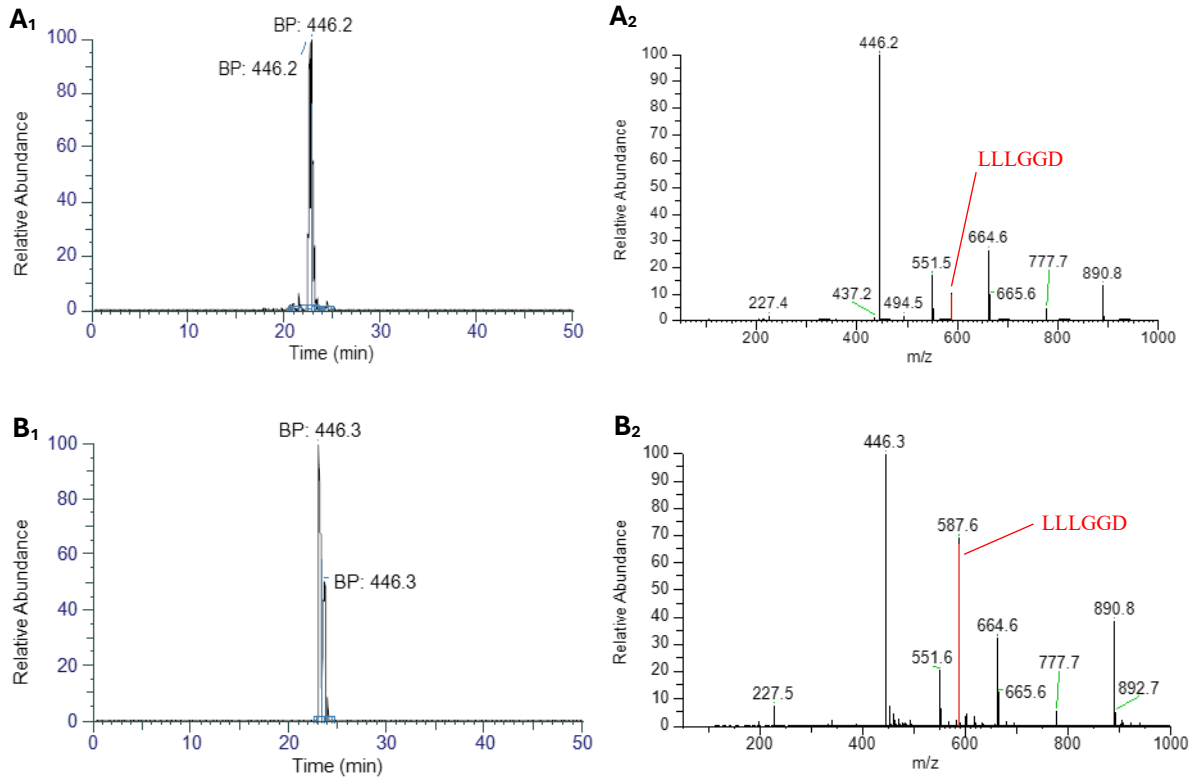


Figure 9.6: The LC-MS base peak profiles of the pepsin proteolytic degradation profiles of P4 *(1) after elution at 0.1 mL/min in pH 2.0 HCl through a 30 mm L x 4.6 mm ID column packed with silica resin functionalised with 185 mg/g.resin pepsin (A) and after 30 minutes of reaction at 0.1 mg/mL in 0.5 mg/mL pepsin solution in pH 2.0 HCl (B) with the mass spectrum of the second half of the split peak (2) highlighting the different m/z value*

9.5. Expanding to a complex peptide structure - Semaglutide

After establishing that the proteolytic degradation selectivity of the immobilised pepsin qualitatively exhibited similar behaviour to that of the free enzyme for a library of simple peptides with primarily Leucine target sites, it was necessary to briefly assess the immobilised pepsin performance with a more complex structure. Semaglutide was selected due to its prevalence within research interest at the current time and due to the glucagon-like peptide-1 (GLP-1) analogue sequence containing an expected seven residue sites susceptible to proteolytic cleavage, which may be increased or decreased by the presence of the stearic acid side chain. Similar to P1-P6, semaglutide was first eluted through a dead column to ensure that there was no structural degradation during the experimental process other than proteolytic cleavage. The mass spectra of the collected dead column fractions and of the diluted original peptide solution

are illustrated in Figure 9.7. Unlike P1-P6, semaglutide did not fractionate under the LC-MS operating conditions, with two peaks representing the $[M+3H]^{3+}$ and $[M+4H]^{4+}$ primarily being observed, the former being the more prominent.

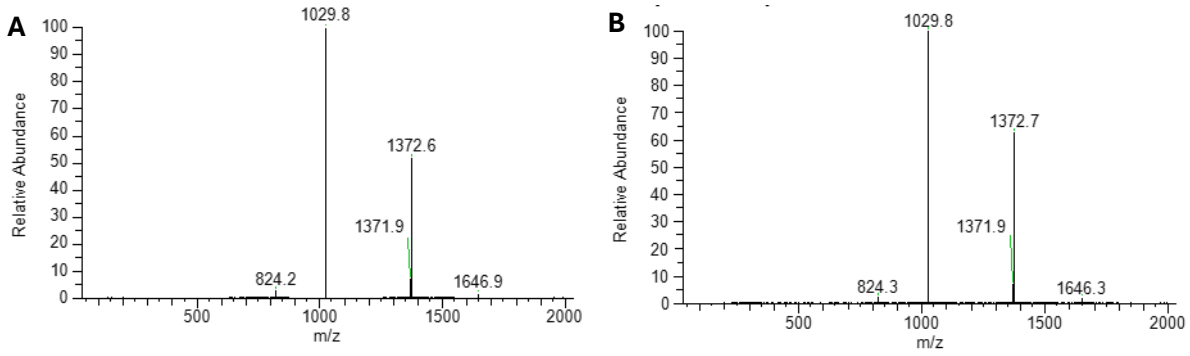


Figure 9.7: Mass Spectra of semaglutide
(A) diluted in 0.6 M pH 2.0 formic acid and (B) after elution 0.1 mg/mL through an end-capped propylamino silica packed column at 37 °C and 0.1 mL/min in 0.6 M pH 2.0 formic acid

Unlike for P1-P6, the comparison between the activity of the free and immobilised enzyme experiments required more detailed analysis. This was for two primary reasons: firstly, due to the increased size and complexity of the peptide structure, differences in the selectivity of the enzymes would result in drastically different degradation product distributions, whilst secondly, due to the difference in product separation required for free enzyme reactions, some products may have been lost to a greater extent than others, thereby affecting the relative product distributions. The first of these issues may compromise the ability of the setup to be used as a strategy to replace free enzyme analysis of peptide degradation stability, whilst the latter may actually represent why the setup is an improvement. However, there may be difficulty in differentiating the two.

Example mass spectrometry (MS) base peaks of the free and immobilised pepsin proteolytic degradation products of semaglutide are shown in Figure 9.8. What was immediately observed was the difference between the two elution profiles, with the immobilised enzyme degradation products displaying more identifiable peaks. In particular the profile of the 1062.9/532.4 peak in

the elution profile of the free enzyme degradation products was significantly more noisy. (The difference in the charge of this peak was associated with the higher acidity of the immobilised enzyme sample, since it did not require base neutralisation nor enzyme precipitation with ethanol). This peak in fact correlated to an Ala-25 - Glu-27 fragment, the smallest species containing the lipid side chain, hence it being the latest eluted from the reversed-phase liquid chromatography (RPLC) column. Additionally, this was suggested to be the cause for the staccato peak, as the species was being poorly nebulised on account of its lower volatility relative to the solvent. However, upon closer viewing of the spectra, what was noticed was that both contained the 626.8 (His-7 – Ser-18), 708.3 (His-7– Tyr-19) and 536.5 (Phe-28 – Trp-31) m/z peaks in addition to the aforementioned Ala-25 - Glu-27 fragment, with each exhibiting similar relative peak heights proportional to each other. In fact, three primary peaks were observed to be missing from Figure 9.8A which were present in Figure 9.8B, namely those pertaining to m/z values of 675.6 (His-7 – Phe-12), 458.2 (Ala-24 – Gly-31) and 453.5 (Phe-12 – Tyr-19). (Other fragments were also eluted but were coeluted with others such that they were never the dominant peak, hence the presence of broad overlapping peaks in each spectrum). Given the presence of the His-7 - Tyr-19 fragment in both spectra, it seemed unlikely that both the His-7 – Phe-12 and the Phe-12 – Tyr-19 fragments were both lost during the free enzyme precipitation, centrifugation and filtration processes. Therefore, appeared that the immobilised enzyme showed higher proclivity to cleave semaglutide at the Thr-Phe and Phe-Thr sites relative to the free enzyme. However, this did not explain why the relative abundance of the Phe-12 – Tyr-19 fragments relative to the other fragments present in both spectra remained relatively unchanged, unless equivalent amounts of each other peak were lost during the separation process too.

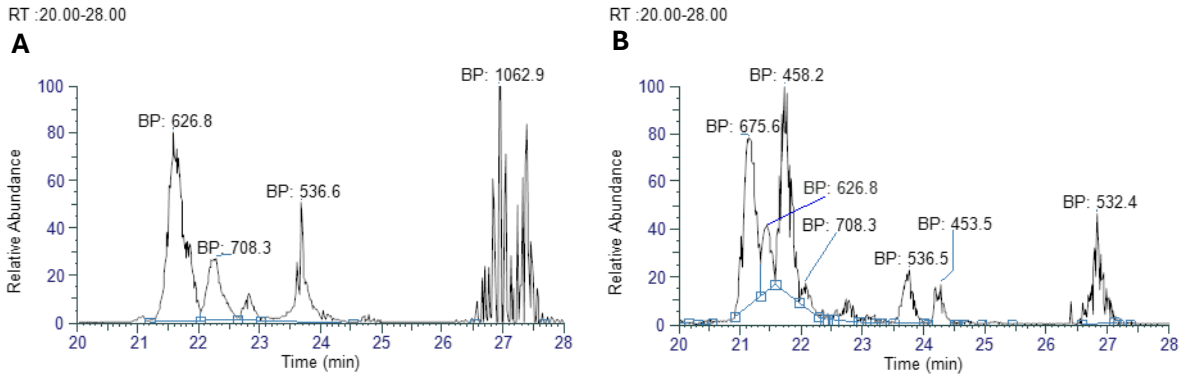


Figure 9.8: Example LC-MS base peaks of proteolytic degradation products of 0.1 mg/mL semaglutide

by (A) 0.5 mg/mL free pepsin for 5 minutes and (B) immobilised pepsin at 37 °C and in pH 2.0 HCl at 0.1 mL/min (highlighted only 20-28 min elution time for peak clarity)

Timepoint analysis of the solution reaction and varied flowrate analysis for the immobilised pepsin reactions for the various peaks observed for both conditions was carried out to establish whether the reaction profiles for each product and the initial sample were similar. From those results, it may be possible to determine whether any unobserved products from the free enzyme reactions were being lost or were not being produced. Therefore, three fragment m/z values were selected in addition to the unreacted semaglutide to analyse their relative formation and degradation kinetics. The three selected fragments were Ala-25 – Glu-27, as initial analysis of the various peaks present had established that it was a peak present in both samples that was not further degraded after initial reaction; His-7 – Ser-18, as previously discussed, it was present in both spectra but also contained the His-7 – Phe-12 fragment absent from the free pepsin product profiles, and finally, His-7 – Ala-24, as similarly it contained both peaks absent from the free pepsin product profiles, but also was more prevalent in the free enzyme samples.

As may be observed in Figure 9.9, the profiles of the substrate and products previously specified followed similar but not identical trends. Two general observations were immediately apparent. Firstly, product fragment fractions did not reach as high a relative abundance for the immobilised pepsin proteolysis (Figure 9.9B) than the free enzyme (Figure 9.9A). This was on account of other peaks that had higher relative abundances which were observed in the immobilised pepsin

product samples but not the free enzyme sample, as discussed previously. These may be observed in Figure 9.8. Secondly, the trends for the 7-18 and 7-24 product profiles appeared to be relatively produced and degraded at a much higher rate for the immobilised pepsin than the free enzyme solution. However, the relative abundance of the 25-27 fragment increased at a much slower rate by comparison. By contrast, the degradation of semaglutide itself followed a highly similar trajectory. These factors are all explained again by high and increasing relative abundances of fragments which were not present in the free enzyme spectra (see Figure 9.10 for example production profiles of two of these species). This caused lower relative amounts of all samples, resulting in more rapid reduction in 7-18 and 7-24 fragments and slower relative increase in the 25-27 fragment. However, this may have also been coupled with greater rates of reaction for the immobilised pepsin for the 7-18 and 7-24 fragments, since the 7-24 fragment was no longer observed after a space time of 80 s. However, the relative rate of reduction of the abundance of the m/z relating to semaglutide would have been expected to be higher too if this were solely the case.

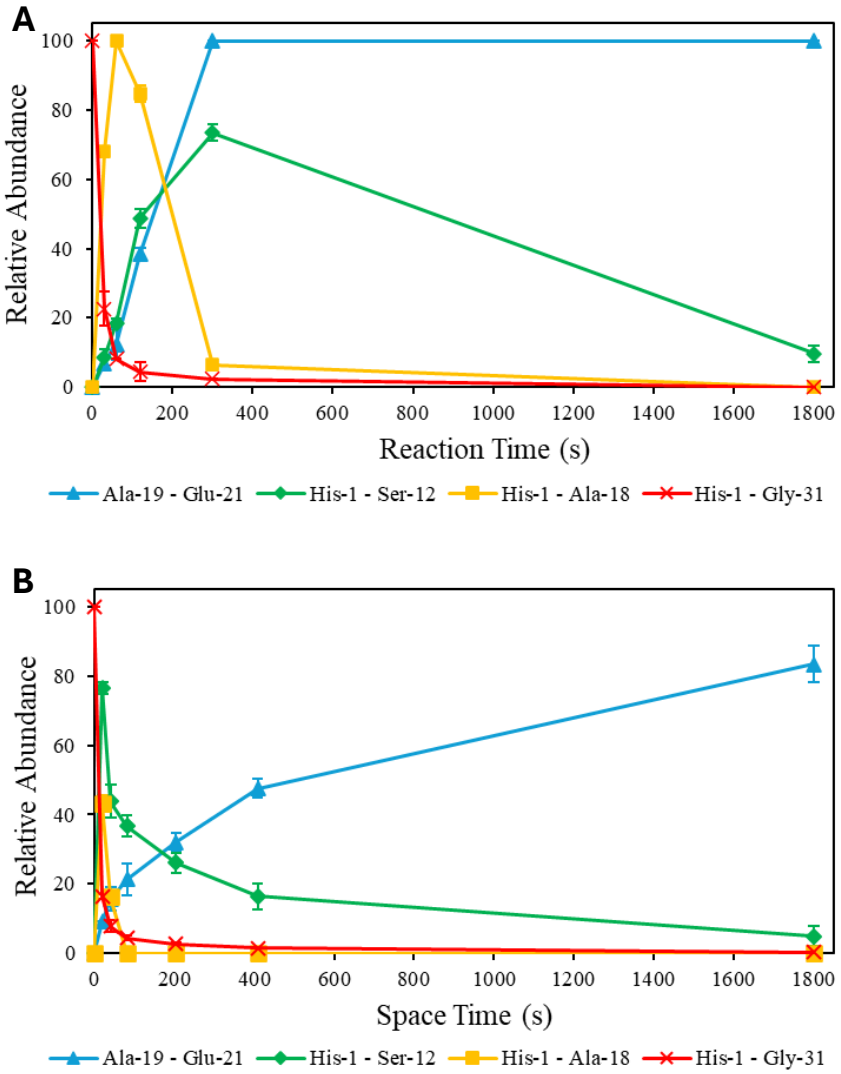


Figure 9.9: Reaction profiles of semaglutide and three of its proteolytic degradation fragments *upon exposure of 0.1 mg/mL semaglutide at 37 °C and in pH 2.0 HCl to (A) 0.5 mg/mL pepsin and (B) elution through a 0.34 mL immobilised pepsin reactor at various flowrates*

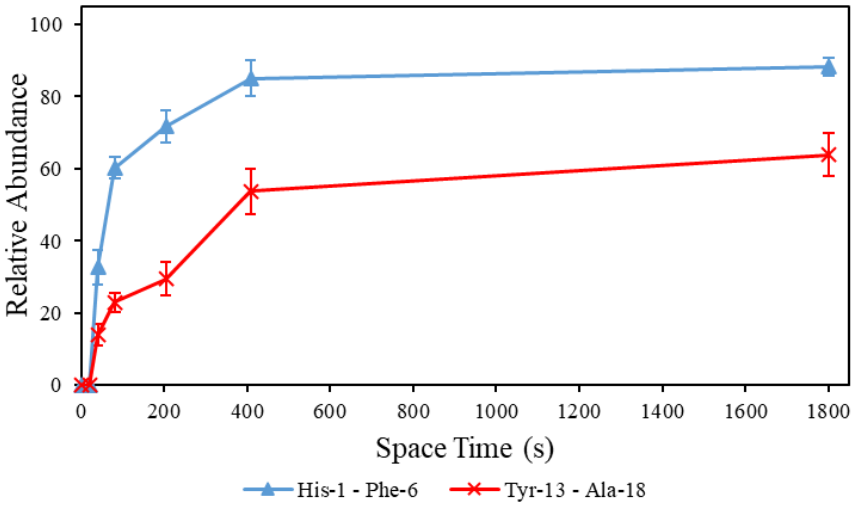


Figure 9.10: Reaction profiles of two of the proteolytic degradation fragments of semaglutide *upon elution of 0.1 mg/mL semaglutide at 37 °C and in pH 2.0 HCl through a 0.34 mL immobilised pepsin reactor at various flowrates*

As may be observed in Figure 9.9B and Figure 9.10, the degradation of His-7 – Ser-18 and His-7 – Ala-24 correlated very well with the formation of fragments His-7 – Phe-12 and Tyr-19 – Ala-24, respectively. The latter two smaller fragments were not observed at all upon elution of semaglutide at 1 mL/min (~20 s spacetime). Conversely, at the same flowrate, the larger fragments, the presence of which was partially dependent on the low rate of formation of the smaller fragments, were observed in their highest relative abundances. However, in Figure 9.9A, it may be observed that the relative abundance of both the His-7 – Ser-18 and His-7 – Ala-24 proteolytic degradation fragments reduced with increasing solution reaction time, but the His-7 – Phe-12 and Tyr-19 – Ala-24 fragments, expected upon their degradation, were unobserved in those reaction solutions. However, since in Figure 9.9A, the formation of His-7 – Ser-18 correlated strongly with the reduction in His-7 – Ala-24, Tyr-19 – Ala-24 production would also be expected unless this was also rapidly being degraded. Therefore, it can be concluded with a relatively high degree of certainty that for the Tyr-19 – Ala-24 fragment, it was either rapidly further degraded, (which was unlikely due to the fact that it is only a hexamer peptide species with only one hydrophobic residue at Leu-20 that would promote pepsin proteolytic cleavage), or that the

concentration of this sample fragment was lost during the pepsin separation procedure, perhaps due to reduced solubility in ethanol of this peptide fragment.

The suggestion of the loss of fragment sample peptides during pepsin precipitation was further supported by Figure 9.11. This figure shows an LC-MS base peak profile of the solution pepsin degradation products of semaglutide, which was required to be repeated due to the high loss of peptide fragments. The abundance of the Ala-25 – Glu-27 fragment was observed to be significantly higher than all others, likely due to their precipitation or aggregation during the pepsin precipitation with ethanol, skewing the relative abundances. This idea is supported by the use of ethanol as a precipitating agent for several short peptides by Chang et al., whilst the lipidated peptide fraction may have much higher solubility in ethanol [390]. If the Ala-25 – Glu-27 fragments were lost at a similar rate, then greater sample injection volumes may have been used, whilst ensuring not to overload the trap column. However, this may just be one of the difficulties with using a lipidated peptide and as such a different procedure may be required.

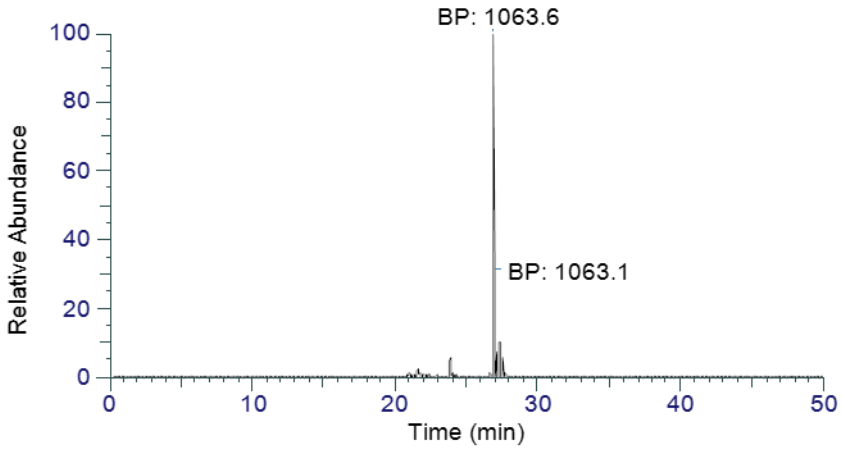


Figure 9.11: Example base peak of semaglutide proteolytic degradation products after 30 minutes incubation
at 0.1 mg/mL in 0.5 mg/mL pepsin at 37 °C in pH 2.0 HCl

9.6. Varied Column Loading

To further establish the relationship between the amount of proteolytic degradation and the observed semaglutide fractions, the pepsin immobilisation was repeated with varying concentrations to provide reactors of varying coverages. The reaction conditions and resultant enzyme coverages, as measured by UV-Vis spectroscopy of the reaction sample, are outlined in Table 9.1. Since the immobilisation was observed to proceed initially very quickly before slowly tending to an equilibrium, the initial concentrations of pepsin were varied rather than reaction time for ease of control and to establish optimised immobilisation conditions. However, it demonstrated that initially the loading yield decreased with initial concentration before increasing again at 40 mg/mL. This suggested that either aggregation of the enzyme on the surface or an increase in physisorption were being observed above a specific concentration. Once packed into a column, the resultant immobilised pepsin reactors aimed to establish the relationship between enzyme loading and the qualitative rate of reaction to semaglutide in terms of the species observed. Furthermore, this would aid in confirming whether the species not observed after solution reactions had indeed been lost during the precipitation separation.

Table 9.1: Varied reaction conditions and resultant pepsin loading concentrations

Immobilisation Pepsin Concentration (mg/mL)	Pepsin Resin Loading (mg/g.resin)	Error (mg.g/resin)
0	0.00	0.00
2	56.6	1.1
10	119	1.4
20	199	3.7
40	445	11.0

The relative abundance of a selection of the resultant observed peptide fragments after elution of 0.1 mg/mL semaglutide through each of the immobilised pepsin reactors of various loadings at 0.1 mL/min are outlined in Figure 9.12. Regardless of coverage, the larger fragments, such as the original substrate and the His-7 – Ala-24 fragment, that were readily observed in solution, were observed for none of the loadings. This was assumed to be due to the immobilised enzyme

still being in a large excess relative to the solution reactions, despite the reduced coverages tested. Therefore, specific species that required only one or two interactions to form, such as the His-7 – Ala-24 fragment, were readily degraded further and so required significantly lower loading than those tested to be observed. This evidence was further supported by the decreasing amount of the large His-7 – Ser-18 fragment observed with increasing enzyme loading. This observation suggested that the increase in pepsin loading perhaps increased the probability of an interaction between the substrate and the surface immobilised enzyme occurring and so the average number of interactions of a given fragment. Nevertheless, this probability may have remained sufficiently high at a coverage of 56.6 mg/g.resin pepsin, such that there was limited observation of His-7 – Ala-24 or the original substrate. However, it was unexpected that a short fragment, such as the Ala-25 – Phe-28 peptide, was observed to decrease in relative abundance, as the amount of Ala-25 – Glu-27 fragment increased with increasing pepsin loading. This suggests that either the increased crowding of the resin surface resulted in preferential fragmentation and binding at the Glu-27 – Phe-28 bond or that the proximity to the conjugated lipid species at the Lys-26 group provided an increased relative hydrophobicity and so stabilised binding at the active site to enable pepsin to display exopeptidase action.

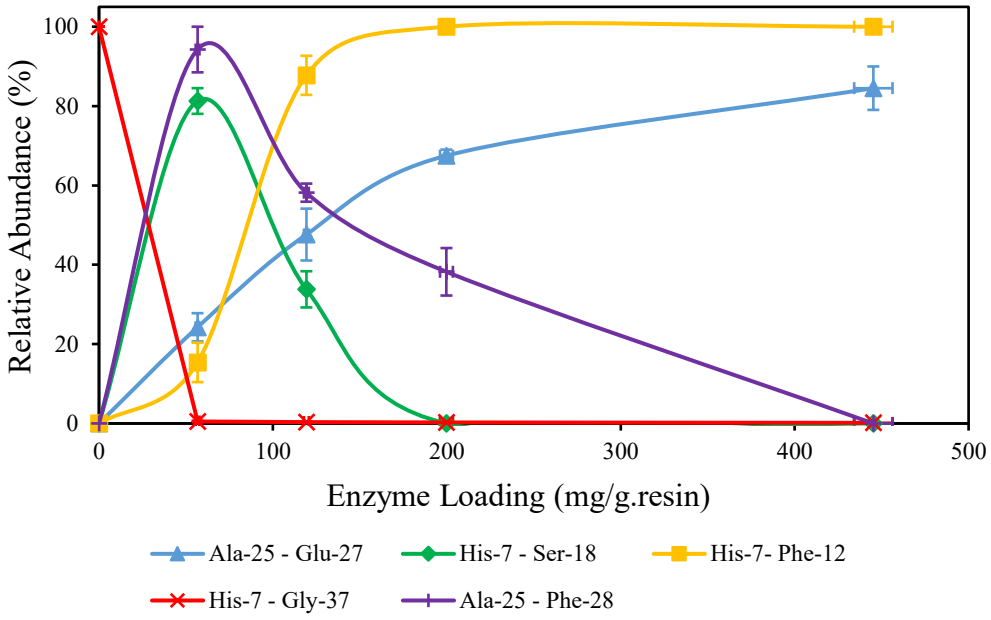


Figure 9.12: Variation in LC-MS detected relative abundance of semaglutide fragments *fractionally collected after elution of 0.1 mg/mL solution through immobilised pepsin reactors of various loadings at 0.1 mL/min*

Unfortunately, however, issues with LC-MS at the time prevented this study from being expanded into a more detailed analysis of how the enzyme coverage affected the relative activity of the enzyme.

9.7. Immobilised versus Free Solution Activity

To establish the change in proteolytic activity after immobilisation, the relative rates of proteolytic cleavage of P6 and semaglutide (the two peptides most susceptible to degradation of those previously tested), were eluted through three 30 mm L x 4.6 mm ID immobilised enzyme reactors (with pepsin loadings of 210.6 ± 8.6 mg/g.resin). They were also exposed to 0.5 mg/mL pepsin solutions at varying concentrations. The resultant rates of proteolytic degradation were calculated via measuring the change in concentration of the original substrate after 30 s in solution and after elution at 0.5 mL/min (~41 s residence time) for the immobilised pepsin reactors. These values were then used in the Lineweaver-Burk plots in Figure 9.13 to establish

the maximum velocity and Michaelis-Menten constants of the pepsin under each condition to each substrate via the following equations:

$$y - intercept = \frac{1}{V_{max,app}} \qquad \textbf{Equation 9.1}$$

$$Gradient = \frac{K_m}{V_{max,app}} \qquad \textbf{Equation 9.2}$$

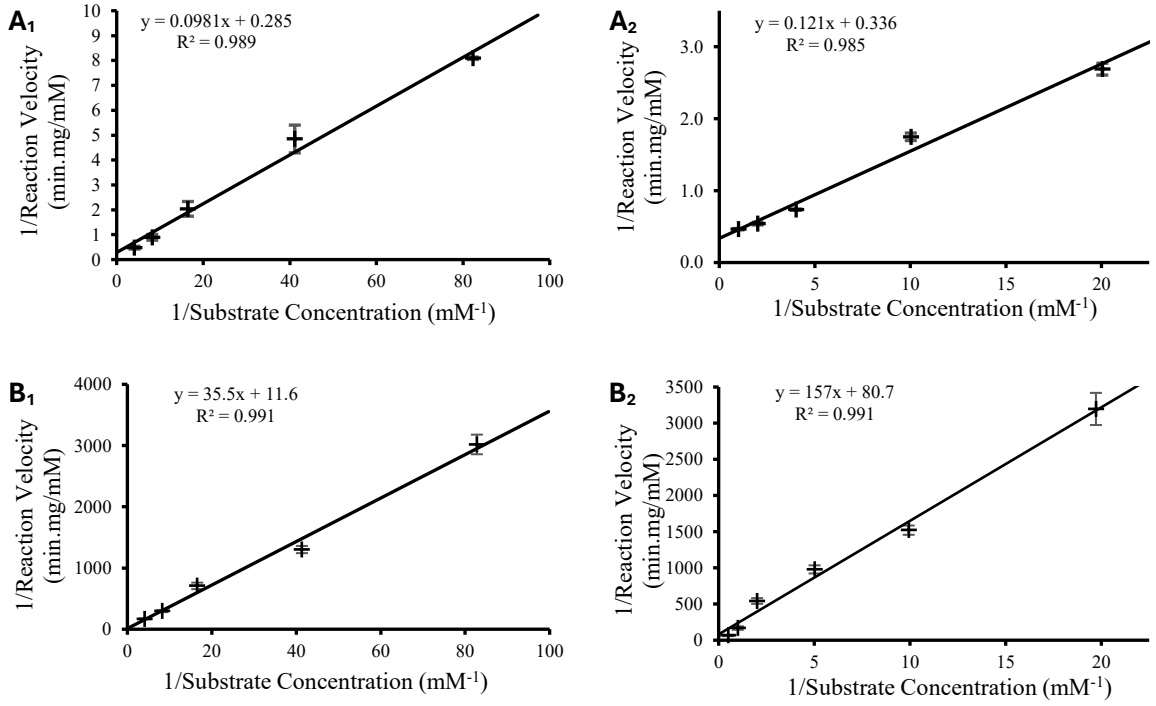


Figure 9.13: Lineweaver-Burk plots to establish the proteolytic activity of (A) free solution and (B) immobilised pepsin
to varying concentrations of (1) semaglutide and (2) P6

From an initial analysis of Figure 9.13, by noting the difference in scale of the y-axes by orders of magnitude, it is apparent that there was a significant loss in pepsin activity after immobilisation. The apparent maximum velocity of proteolytic activity of pepsin reduced from approximately 3.51 and 2.98 mM/min/mg.pepsin in solution to 0.086 and 0.01298 mM/min/mg.pepsin after immobilisation for semagltuide and P6, respectively. The reaction velocity parameters were normalised via the estimated mass of enzyme to account for the significantly higher enzyme concentration in the immobilised enzyme reactor relative to the solution reactions, which aimed

to mimic the gastric environment concentrations. Furthermore, the Michaelis-Menten constant increased from 0.34 and 0.36 mM in solution to 3.07 and 1.95 mM after immobilisation for semaglutide and P6, respectively. This indicated an over 5-fold increase in the substrate concentration required to reach half of the maximum rate of proteolytic activity and as such highlighted the a significant reduction in binding-affinity of pepsin to the substrates.

The reduction in activity was primarily on account of three potential factors: first, the concentration of the immobilised enzyme being orders of magnitudes higher relative to in solution, leading to potential overcrowding, which has been a discussed potential issue in glutaraldehyde immobilisation procedures [281]; second, a potential reduction in mobility or binding capacity of the pepsin subsites and active site, caused by structural changes during immobilisation; and thirdly potential interactions between the resin and the substrate, reducing the strength and number of enzyme-substrate interactions. The various effects the enzyme immobilisation strategy may have on the immobilised enzyme activity have been reviewed by Holyavka and Artyukhov [391], but these were seen as most likely due to the high coverage, use of glutaraldehyde-mediated covalent immobilisation procedure and a silica support phase with hydrogen-bond forming surface, silanol groups. When considering the effects of the concentration of the immobilised enzyme, it must first be noted that the value used for the pepsin loading throughout was that which was estimated via UV-Vis analysis immediately following the immobilisation reaction. Therefore, it did not factor in any reduction due to enzyme leaching from the resin surface. Since the apparent maximum reaction velocity was normalised by the amount of pepsin immobilised, any reduction in pepsin-loading from this initial measurement throughout the experimental process was not factored into the final reaction velocities. However, the leaching was not likely to be sufficient to account for the entire reduction in activity relative to in solution. As previously mentioned, due to the high concentration of the enzyme, there was likely to be crowding of the pepsin molecules on the resin surface, thereby reducing availability of active sites and binding subsites. Furthermore, due to the high pepsin loading, it was assumed

that there was a significant amount of physisorption, for which, unlike the covalent immobilisation, it cannot be guaranteed that the immobilisation is in an orientation which leaves the active site available for proteolytic action. Additionally, whilst it was assumed that there was an even dispersion of substrate through the column upon sample injection, the high pepsin loading may result in specific pores and so diffusion pathways being blocked, thereby reducing the amount of immobilised enzyme available relative to the estimated total amount immobilised. However, this was the case with all chromatographic resin surfaces. Furthermore, without detailed computational analysis estimating the number of potential substrate-enzyme interactions per column volume, it is not possible to truly quantify these effects.

In addition to the high enzyme loading, by increasing the rigidity of the enzyme structure, immobilisation, especially covalent immobilisation, reduces the mobility and flexibility of the enzyme subsites and active sites to bind to the substrate for proteolytic action [267]. Therefore, this provides additional explanation for the reduction in relative activity of the enzyme, and in particular the great increase in the observed Michaelis-Menten constants. Other factors that are not often discussed but may be relevant are the effects of the immobilisation resin. Firstly, whilst there appeared to be little to no retention of the peptide substrate during dead column experiments, nor any significant size exclusion effects, it cannot be ruled out whether interactions in the resin environment resulted in a change in conformation or orientation of the peptide substrates. Such changes may have resulted in reduced accessibility or exposure of their hydrophobic residues. Since pepsin preferentially cleaves at these hydrophobic sites, any such change in conformation would result in a reduction in apparent proteolytic action. Since the propylamino silica surface was assumed to be highly hydrophilic even after end-capping with ethanolamine, this may have resulted in an increase in peptide secondary structure conformational-folding to reduce the exposure of hydrophobic sites relative to solution reactions, and so reduced the observed pepsin activity.

9.8. Column Ageing Study

To establish the temporal stability of pepsin after its immobilisation, three new columns were prepared and at varying concentrations (0.05-1.00 mg/mL),P4-P6 and semaglutide were eluted through the columns at 1-4, 8 and 12 weeks after the initial preparation of the pepsin immobilisation reaction. The calculated maximum reaction velocity and Michelis-Menten numbers for the proteolytic activity of the immobilised pepsin to each substrate at each timepoint are plotted in Figure 9.14. Upon analysis of the graphs, three clear trends were apparent. Firstly, the Michaelis-Menten constant of the immobilised pepsin action to the P4 substrate increased greatly over the course of the initial weeks; second, between weeks 1 and 2 the greatest decrease in maximum reaction velocity was observed for all substrates; third, the reduction in maximum reaction velocity and increase in Michaelis-Menten constant (aside from P5 between weeks 8 and 12) were relatively linear with time. The first of these observations was assumed to be due to the poor activity of the immobilised pepsin to P4 (as discussed in Section 9.4); therefore, any reduction in activity resulted in a much greater observed change in the relative amount of original substrate remaining. Therefore, after week 4, the previously low level of activity observed reduced to being undetectable due to the poor binding of the enzyme to the substrate peptide as indicated by the increasing Michaelis-Menten constant. The second observation was assumed to be due to enzyme leaching from the propylamino silica resin, on account of a potentially significant fraction of the immobilised enzyme being physisorbed onto the highly hydrophilic surface, densely-populated with hydrogen-bonding silanol groups. This limitation of adsorbed enzymes species has been previously discussed by Robescu and Bavaro [280] and has been widely observed even for enzymes adsorbed and entrapped within the support phase [288]. Therefore, whilst the physisorbed species may have maintained their relative proteolytic activity to a greater extent than the covalently-bound species, as their structure had been altered to a lesser degree, unlike the covalently-bound enzymes, they may have leached from the column with time and repeated cycles of exposure to increased pressures

and highly acidic conditions (pH 2.0), breaking the intermolecular forces. However, since there were not high levels of pepsin detected in the eluted sample fractions, it was perhaps more likely that since physisorption did not provide the same stabilising effects as the increased rigidity caused by covalent immobilisation, a fraction of the immobilised pepsin denatured or deactivated with time, as reviewed by Iyer and Ananthanarayan [392], thereby reducing the observed overall activity. It may have also been due to column fouling or blocking of the pepsin active sites by substrate build-up, degradation products or impurities in the mobile phase or the sample injections. This final suggestion seemed more likely to be representative of the cause of third observed property of the results, the approximately linear reduction in activity and increase in Michaelis-Menten constant exhibited between weeks 2 and 12, since the column was consistently being used at the same rate over the course of the entire period of work. Analysing the linear region, between week 2 and week 8, with 57 samples being eluted through each column per week, the Michaelis-Menten constant was observed to increase by on average 5.9% (±0.3%), whilst the maximum reaction velocity decreased by 4.8% (±0.8%). These results were similar to those observed by Zhang et al. [386], who observed a retention of 95.9% activity after 40 runs. However, given these results were over a greater number of runs and a longer time period, it may be that this technique of immobilisation shows even greater stability.

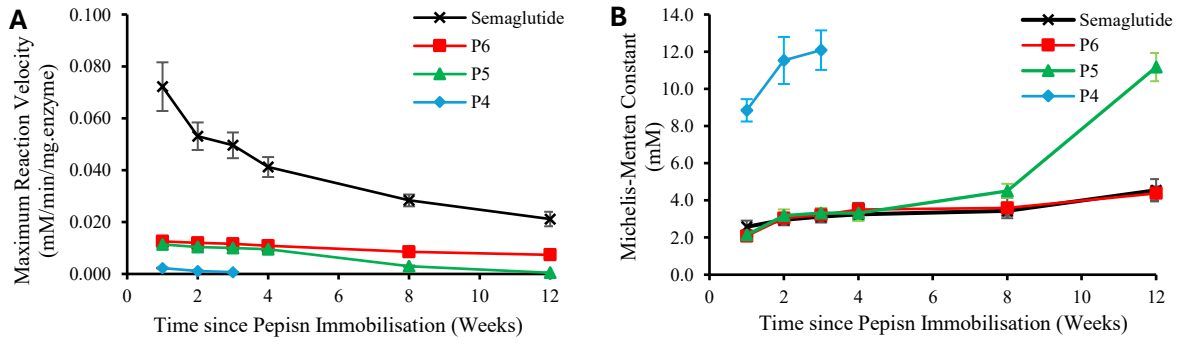


Figure 9.14: Variation in (A) maximum reaction velocity and (B) Michaelis-Menten constant of immobilised pepsin
with substrates P4, P5, P6 and semaglutide

9.9. Expanded Peptide study

Upon evaluating the activity of the immobilised pepsin under various conditions with respect to specific reference peptides and semaglutide, the study was broadened to a wider range of peptide substrates (as outlined in Sections 5.1.1, 5.1.2.1 and 5.1.4). These species included peptides with specific sequences or structural motifs that are suggested by literature to inhibit enzymatic degradation. They were selected to establish whether the immobilisation of pepsin results in a change in the selectivity to these structures. Therefore, pepsin was immobilised again, and three further columns were packed to study the susceptibility of these additional peptides to proteolytic degradation via immobilised pepsin. The results and the relevance of their structures are outlined in the following sections.

9.9.1. GLP-1 Analogues

In addition to semaglutide, two additional GLP-1 analogues, namely liraglutide and taspoglutide, were later eluted through the immobilised pepsin reactors. The differences in amino acid sequence from semaglutide and both new peptides were not anticipated to have a major impact in the enzymatic degradation pathway for one of two reasons. Firstly, they are proximal to the N- (Aib-8 to Ala-8 in liraglutide) or C-terminals (Gly-35 to Aib-35, loss of Gly-37 and amidation of C-terminal Arg-36 in taspoglutide) and since pepsin is an endopeptidase, the specific residue differences would likely not significantly impact enzyme binding at potential cleavage sites [340]. Secondly, the change from Arg-34 to Lys-34 in taspoglutide was not expected to provide sufficient change to the enzyme degradation pathway due to both amino acids being inhibitory to pepsin activity (as outlined in Table 5.3). However, the difference in lipidation of these peptides relative to semaglutide was considered to be relevant since large lipid motifs may sterically hinder the enzyme from readily binding to the peptide substrate. However, since pepsin is more selective to proteolytic degradation at more hydrophobic amino acid residues and the lipid motif may also increase the relative hydrophobicity of the sequence via the positive inductive effect, the presence and length of the lipid chain may also affect the selectivity of the protease.

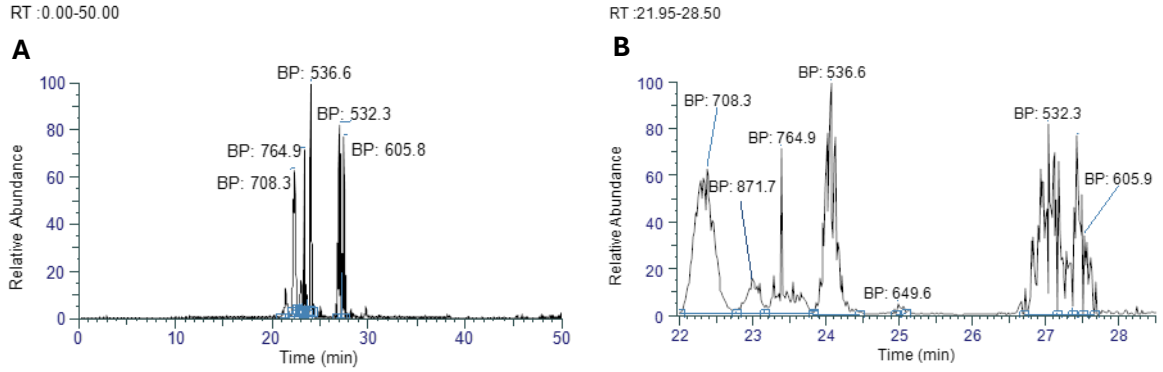


Figure 9.15: LC-MS base peak chromatogram of semaglutide after elution through an immobilised pepsin reactor

(A) full and (B) magnified view of the subsection with elution time 21.95-28.50 mins

Table 9.2: Observed MS peaks of semaglutide after elution through an immobilised pepsin reactor *and their corresponding fragment sequences in order of their retention time*

m/z $(M+nH)^{n+}$	Fragment Sequence	Fragment M_w (g/mol)
708.3	His-7 – Tyr-19	1414.4
871.7	Thr-13 – Leu-20	870.9
764.9	His-7 – Leu-20	1527.6
536.6	Phe-28 – Trp-31	535.6
649.6	Phe-28 – Leu-32	648.8
532.3	Ala-25 – Glu-27	1062.9
605.9	Ala-25– Phe-28	1208.9

As previously observed and indicated in Figure 9.15, after elution through the pepsin immobilised column, several semaglutide fragments were detected indicating its proteolytic degradation. Exemplary individual mass spectra of each labelled peak in Figure 9.15B are represented in Figure 12.14. These spectra and their corresponding MS^2 fragmental analysis were utilised to identify the fragment correlating to each individual MS peak. These fragments and their molecular weights are outlined in Table 9.2 and show good consistency with previous experimental analysis of semaglutide degradation.

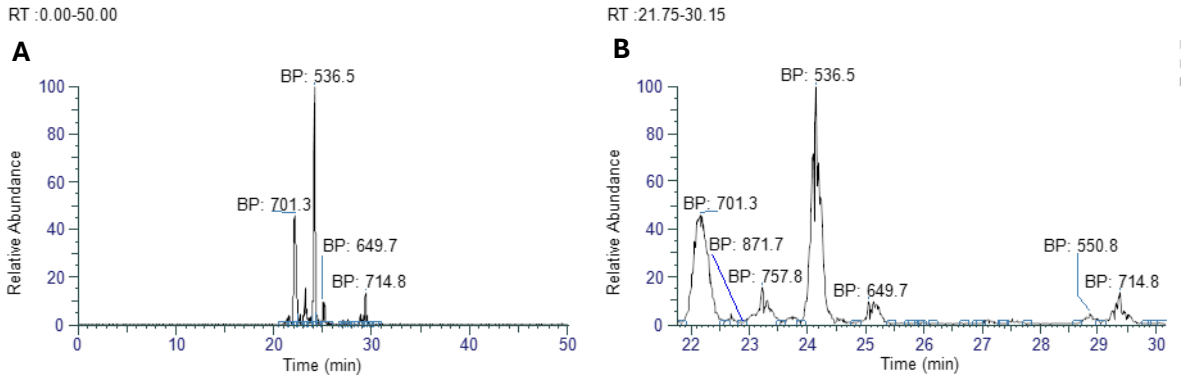


Figure 9.16: LC-MS base peak chromatogram of liraglutide after elution through an immobilised pepsin reactor

(A) full and (B) magnified view of the subsection with elution time 21.75-30.15 mins

Table 9.3: Observed MS peaks of liraglutide after elution through an immobilised pepsin reactor *and their corresponding fragment sequences in order of their retention time*

m/z $(M+nH)^{n+}$	Fragment Sequence	Fragment M_w (g/mol)
701.3	His-7 – Tyr-19	1400.3
871.7	Thr-13 – Leu-20	870.9
757.8	His-7 – Leu-20	1513.5
536.6	Phe-28 – Trp-31	535.6
649.6	Phe-28 – Leu-32	648.8
550.8	Glu-21 – Glu-27	1099.3
714.8	Ala-25 – Glu-27	713.9

The LC-MS base peak chromatogram of the fraction collected after liraglutide elution through the immobilised pepsin reactor showed very similar results to semaglutide, as may be observed by the similarity in number and retention times of each of the peaks in Figure 9.15B and Figure 9.16B. Furthermore, the m/z of the peaks observed were in some cases identical and in other cases, as observed in Table 9.2 and Table 9.3, related to the same cleavage sites but with differing mass values, due to the Ala-8/Aib-8 substitution and the shorter conjugated lipid motif at the Lys-26. In fact, the only major difference was the distinct observation of a Glu-21-Glu-27 fragment in liraglutide and absence of a clear Ala-25 – Phe-28. This indicated that either perhaps the susceptibility of the Ala-Ala or Phe-Ile bonds to proteolytic cleavage were reduced by the more hydrophobic, but shorter, lipid chain in liraglutide relative to semaglutide or that the structure

adopted resulted in increased steric hinderance of the binding of the protease active site to these sites. Nevertheless, given the proteolytic cleavage still occurred at the Glu-27 – Phe-28 bond, it is more likely that the shorter more hydrophobic chain resulted in an increased rate of proteolysis at this site reducing the relative amount of cleavage at the other sites. This is based on the fact that upon proteolytic cleavage at the Glu-21 – Phe-22 due to the other sites in question being proximal to the new C- or N-terminal, the immobilised pepsin significantly reduced binding capacity to those substrate sites.

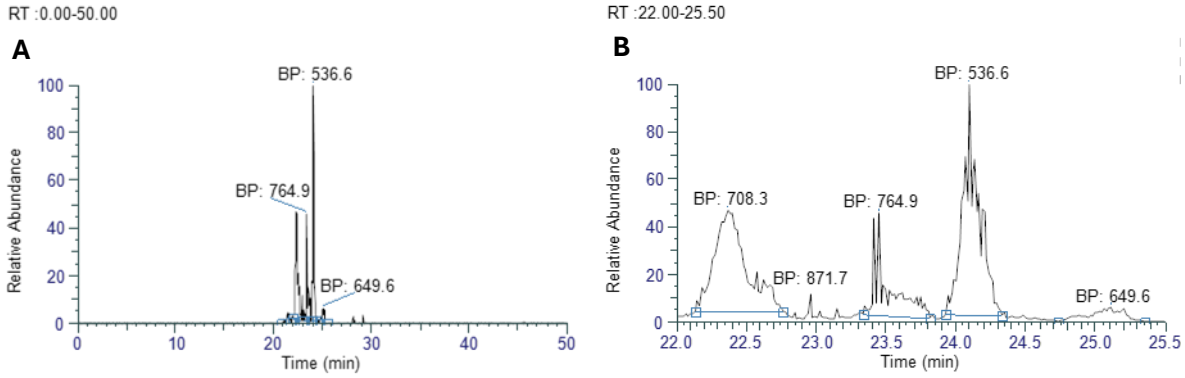


Figure 9.17: LC-MS base peak chromatogram of taspoglutide after elution through an immobilised pepsin reactor
(A) full and (B) magnified view of the subsection with elution time 22.00-25.50 mins

Table 9.4: Observed MS peaks of taspoglutide after elution through an immobilised pepsin reactor *and their corresponding fragment sequences in order of their retention time*

m/z $(M+nH)^{n+}$	Fragment Sequence	Fragment M_w (g/mol)
708.3	His-7 – Tyr-19	1414.4
871.7	Thr-13 – Leu-20	870.9
764.9	His-7 – Leu-20	1527.6
536.6	Phe-28 – Trp-31	535.6
649.6	Phe-28 – Leu-32	648.8

As may be observed through comparisons of Figure 9.15B and Figure 9.17B, all the primary LC-MS peaks observed for the elution of taspoglutide were similarly detected for semaglutide. Prior to this point, the absence of any peaks containing fragments Val-33 to the C-terminal for both liraglutide and semaglutide has not been discussed. This was on account of the fact that the cause of this was assumed to be the high charge and hydrophilicity of this fragment resulting in

it being poorly retained and concentrated in the PepMap-100 trap column during sample loading and due to its high charge resulting in a lower m/z value in the range of high noise. However, in the case of taspoglutide there was also an absence of any primary peaks containing the amino acids in the fragment Glu-21 – Glu-27, as may be observed in Table 9.4. This absence was due to the fact that species containing this fragment were coeluted with other species and as such were not observed as primary labelled peaks in the base peak curve. For example, m/z values of 867.0 and 579.2, indicating Thr-7 – Phe-28 and Tyr-19 – Phe-28 fragments were observed in the mass spectra of 764.9 and 708.3 m/z primary peaks, respectively. However, this established that there was limited proteolytic cleavage between the Ala-24 – Ala-25 in the absence of a lipid motif conjugated to Lys-26, indicating that the group rather than inhibiting degradation via steric hinderance, may have actually increased the proclivity of neighbouring sites, specifically the Ala-25 – Ala-26 site, to proteolytic cleavage.

The effect of the addition of the lipid motif and the consistency in the selectivity of the immobilised pepsin to specific proteolytic sites were in agreement with suggested theory and were clear when the observed and predicted cleavage sites for each of the three GLP-1 sequences were juxtaposed as outlined in Figure 9.18. All other cleavage sites aligned very well with the literature suggestions of promoting and inhibiting residues (as outlined in Table 5.3). However, the only slight discrepancy is that proteolytic cleavage of liraglutide would perhaps be expected to have preferentially occurred at the Phe-28 – Ile-29 peptide bond rather than the Glu-27 – Phe-28 peptide bond. This is due to the fact that Ile due to its higher hydrophobicity may have been expected to promote pepsin binding at the S'_1 substrate site more than Glu at the S_1 site. However, given the proximity of these sites of proteolytic cleavage to the conjugated lipid motif, the difference in preferential binding may have been due to steric hinderance or increased hydrophobicity due to positive inductive effects.

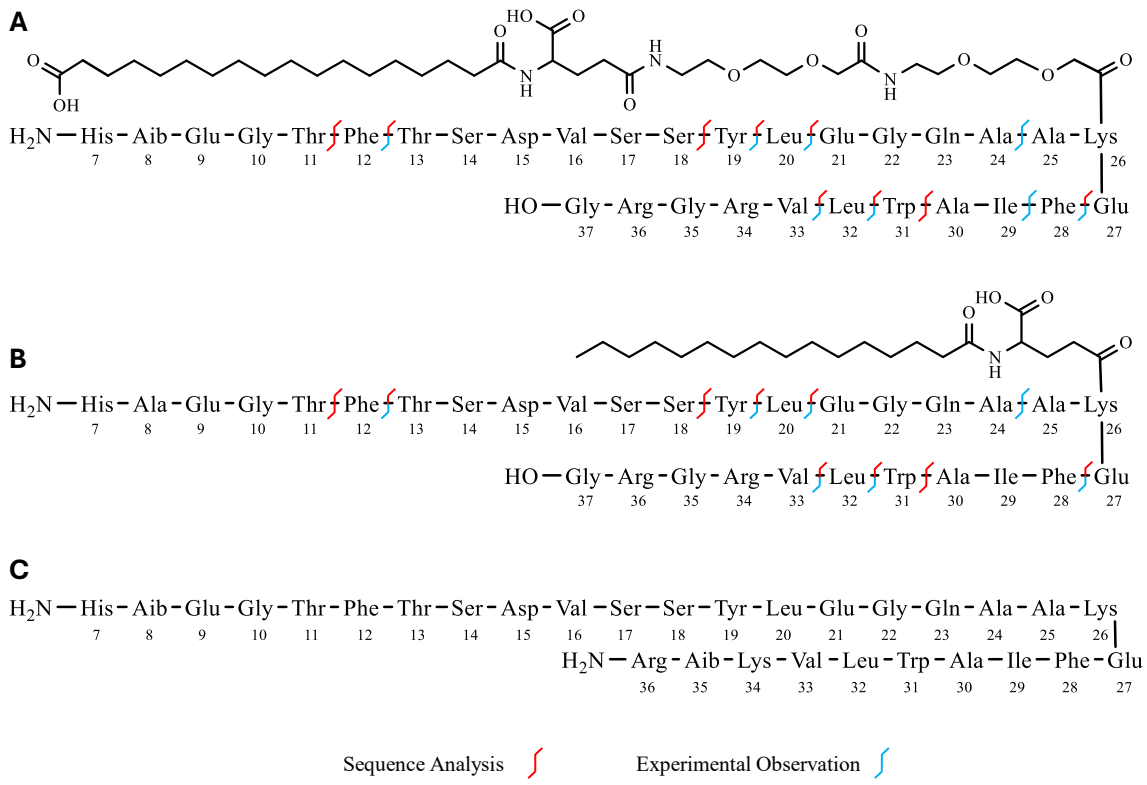


Figure 9.18: Amino acid sequences with immobilised pepsin proteolytic cleavage sites highlighted of GLP-1 analogues

Predicted by sequence analysis (see Table 5.3 for rules) and observed after elution through immobilised pepsin reactor for (A) semaglutide, (B) liraglutide and (C) taspoglutide

9.9.2.Insulin

In addition to GLP-1 analogues, insulin was also eluted through the immobilised pepsin reactors. Given its importance in the treatment of type-1 diabetes and the significant amount of research that has been invested into its oral stability, including from companies such as Oramed Pharmaceuticals, understanding the gastric degradation pathway of insulin is of great interest. Furthermore, its structure, consisting of two amino acid chains linked by two cysteine disulfide bonds with an additional disulfide bond on chain A linking Cys-6 and Cys-11 groups, provides a complex cyclic-like scaffold which may inhibit pepsin binding and so proteolytic cleavage. Therefore, it is of interest to study the stabilising effects of the disulfide cyclic-like properties of its structure and establish their inhibiting effects. In addition, due to insulin being large for a peptide (51 amino acids), it is important to establish whether the increase in size-exclusion

effects would result in it diffusing more poorly into the porous structure of the immobilised enzyme reactor packing and therefore from being able to access the full enzyme coverage.

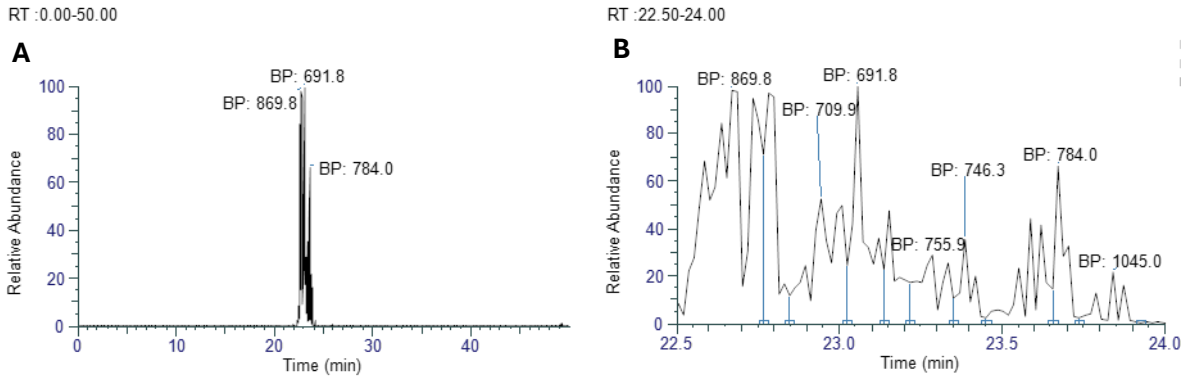


Figure 9.19: LC-MS base peak chromatogram of insulin after elution through an immobilised pepsin reactor
(A) full and (B) magnified view of the subsection with elution time 22.50 -24.0 mins

Table 9.5: Observed MS peaks of insulin after elution through an immobilised pepsin reactor *and their corresponding fragment sequences in order of their retention time*

m/z $(M+nH)^{n+}$	Fragment Sequence	Fragment M_w (g/mol)
869.8	Chain A: Gly-1 - Leu-13 + Chain B: Phe-1 - Leu-11	2606.4
709.9	Chain A: Gly-1 - Leu-13 + Chain B: Phe-1 - Glu-13	2834.1
691.8	Chain A: Gln-15 - Asn-21 + Chain B: Leu-17 - Tyr-26	2071.0
755.9	Chain A: Gly-1 - Leu-13 + Chain B: Phe-1 - Leu-15	3018.4
746.3	Chain A: Gln-15 - Asn-21 + Chain B: Tyr-16 - Tyr-26	2234.3
784.0	Chain A: Gln-15 - Asn-21 + Chain B: Leu-15 - Tyr-26	2347.3
1045.0	Chain A: Gly-1 - Asn-21 + Chain B: Phe-1 - Phe-25	5218.6

An example LC-MS base peak chromatogram of the fraction collected after a 10 µL injection of 0.1 mg/mL insulin into the immobilised pepsin reactors is outlined in Figure 9.19A, with the relevant peak region focused on in Figure 9.19B. It was firstly evident that the RPLC method or the PepMap-100 column applied to separate the fragments may not have been entirely appropriate for these samples since they were almost all eluted within 1.5 minutes of each other

in a 50 minute method. However, the similarity in sequence of several of these fragments must be noted, as they would likely have required a more tailored separation approach to provide greater clarification. This work would be appropriate when applying the technique for the study of an individual peptide's proteolytic cleavage, but since the purpose of this work was to establish the activity of the column for a range of peptides, it was deemed an unnecessary effort. The poor separation also implied that some of the less concentrated species may not have been observed as clearly, since they were obscured by the more concentrated (higher peak) fractions. However, it was possible to identify other coeluting fragments since they had large molecular weights and high charges resulting in multiple m/z peaks correlating to the same species. For example, m/z values of 969.4 and 727.6, correlating to a two-chain species of Gly-A1 – Leu-A13 + Phe-B1 – Ala-B14, were observed to coelute with the 709.8 peak. Additionally, 649.1 and 973.5 m/z peaks correlating to another two-chain species of Leu-A16 – Asn-A21 + Leu-B17 – Tyr-B26 was also observed to coelute with the 746.3 peak. Therefore, despite being a greater challenge due to the lack of MS^2 spectra, it was still possible to identify specific lower concentration species. Nevertheless, as observed in Table 9.5, the fragments correlating to each of the MS peaks aligned with the literature statements on pepsin preferential cleavage sites (as outlined in Section 5.1.5), with the majority of proteolytic cleavage occurring at hydrophobic residues.

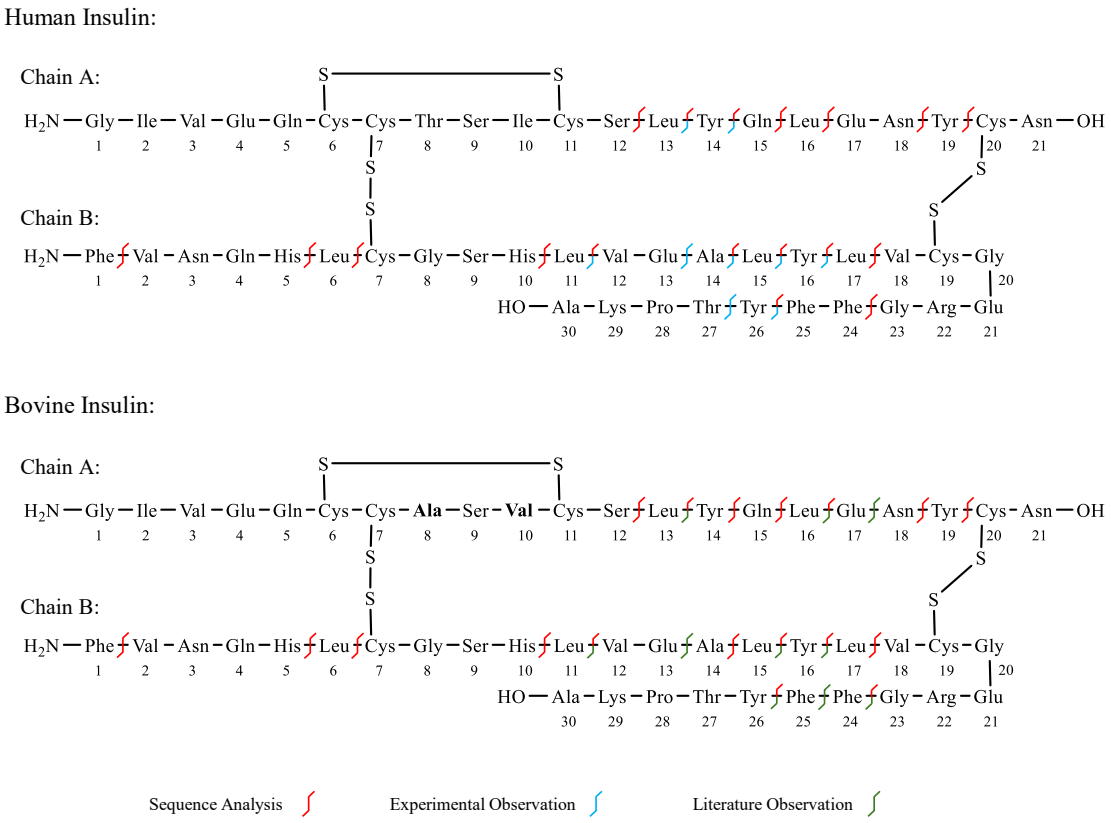


Figure 9.20: Human and bovine insulin sequences and their proteolytic degradation sites *as predicted by sequence analysis (see Table 5.3) and observed for human insulin after elution through immobilised pepsin reactor and using bovine insulin as a comparison with results from Kolinski et al. [393]*

Figure 9.20 outlines more clearly each of the identified proteolytic cleavage sites observed after elution of insulin through the immobilised pepsin reactor. Firstly, it is evident that insulin was preferentially cleaved by pepsin at the hydrophobic residue sites aligning with literature theory and so some of the sites predicted by sequence analysis. However, proteolytic cleavage did not occur at all hydrophobic sites predicted as may be observed by the differences between the sequence analysis and observed experimental results. This was primarily due to the conformation of insulin. Firstly, Phe-B1 is in too close proximity to the N-terminal for sufficient enzyme binding to ensure it was readily cleaved. Additionally, residues such as Tyr-A19 and Leu-B6, despite being located in the random coil secondary conformation segments of their respective chains, which generally allow for high substrate flexibility and stronger active site binding, exhibited no observed cleavage since they are located directly next to cysteine-bridges. These disulfide bonds result in a more rigid structure with increased steric hinderance of the

enzyme to bind to the bond due to the proximity of the second amino acid chain shielding the specific residues on one side, as such these two factors resulted in reduced substrate binding at these sites and so a lower rate of proteolytic cleavage. Similar reasoning can be applied to explain why there was limited/no proteolytic cleavage observed at the Tyr-A19 and Leu-B18 residues. Furthermore, based on the amino acid sequence alone, it would be predicted that pepsin degradation at Phe-B24 at the P_1 site would occur, especially due to the presence of Phe-B25 residue at the P'_1 site. However, it was not readily observed. This was likely due to a combination of three factors: first, the relatively higher rate of proteolytic activity at the Phe-B25 and Tyr-B26 sites, resulting in the Phe-B24 residue being at the penultimate or C-terminal residue and as such having limited enzyme binding; second, the Arg-B22 at the P_3 may have slightly inhibited enzymatic binding [340, 341]; third, in insulin's native conformation, it is located at the beginning of the random coil secondary conformation in a pocket between Asn-A21, Tyr-B16 and Val-B12, resulting in there being potential significant steric hinderance to pepsin binding. The second of these factors likely contributed the least since Pro-B28 should have also inhibited pepsin proteolytic cleavage the Phe-B25 and Tyr-B26 residues, which was not observed clearly in the results.

Typically, the Glu-B13 and Ala-B14 sites would not be expected to be readily cleaved by pepsin, since these amino acids are not suggested to bind strongly to pepsin substrate at the P_1 or P'_1 residue sites. However, upon analysis of the conformation of insulin (as seen in Figure 5.1), it may be observed that these two residues are in the middle of the alpha helix segment of the insulin B chain, with highest conformational flexibility due to their distance from the disulfide bridges. Therefore, it may be that these residues were simply the least sterically hindered from effectively binding to the active site, and so resulted in the maximal rate of proteolytic degradation. This observation was supported by observations from Kolinski et al. who observed proteolytic cleavage by pepsin at the same respective site on bovine insulin (see Figure 9.20) [393]. To conclude, the results observed, whilst they may not have been immediately those that

would have been predicted simply based on the amino acid sequence, were easily explained upon a more detailed examination of the peptide's conformation.

9.9.3.Oxytocin and Vasopressin Analogues

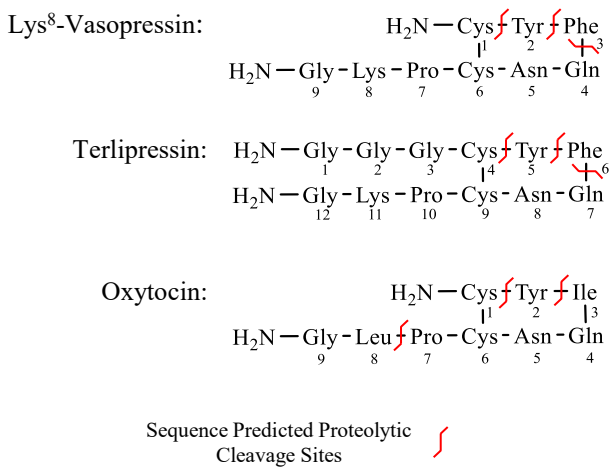


Figure 9.21: Lys[8]-vasopressin, terlipressin and oxytocin sequences and their predicted pepsin proteolytic cleavage sites
As predicted by sequence analysis (see Table 5.3)

To further evaluate the activity and selectivity of the immobilised pepsin, three neuropeptides with disulfide-bridged cyclic structures were also eluted through the immobilised pepsin reactor. The rigidity of cyclic structures, especially in short peptides typically results in poor binding at the enzyme active site and as such limited degradation. Additionally, the similarity in the sequences may allow for increased understanding of the factors that affect the susceptibility of cyclic peptides to pepsin degradation. The substitution of Lys-8 for Leu-8 from [Lys[8]]-vasopressin to oxytocin on the linear peptide trimer near the C-terminal and not within the disulfide-bridged ring should increase the probability of cleavage at this site since leucine promotes pepsin activity at the site whilst lysine inhibits it. However, given the proximity to the C-terminal and the proline residue at P_2 residue, this substitution is not likely to result in a significant change [340, 341]. The addition of the N-terminal triglycine in terlipressin relative to [Lys[8]]-vasopressin, increases the likelihood of pepsin binding at the pepsin activity promoting Tyr-5 and Phe-6 residues.

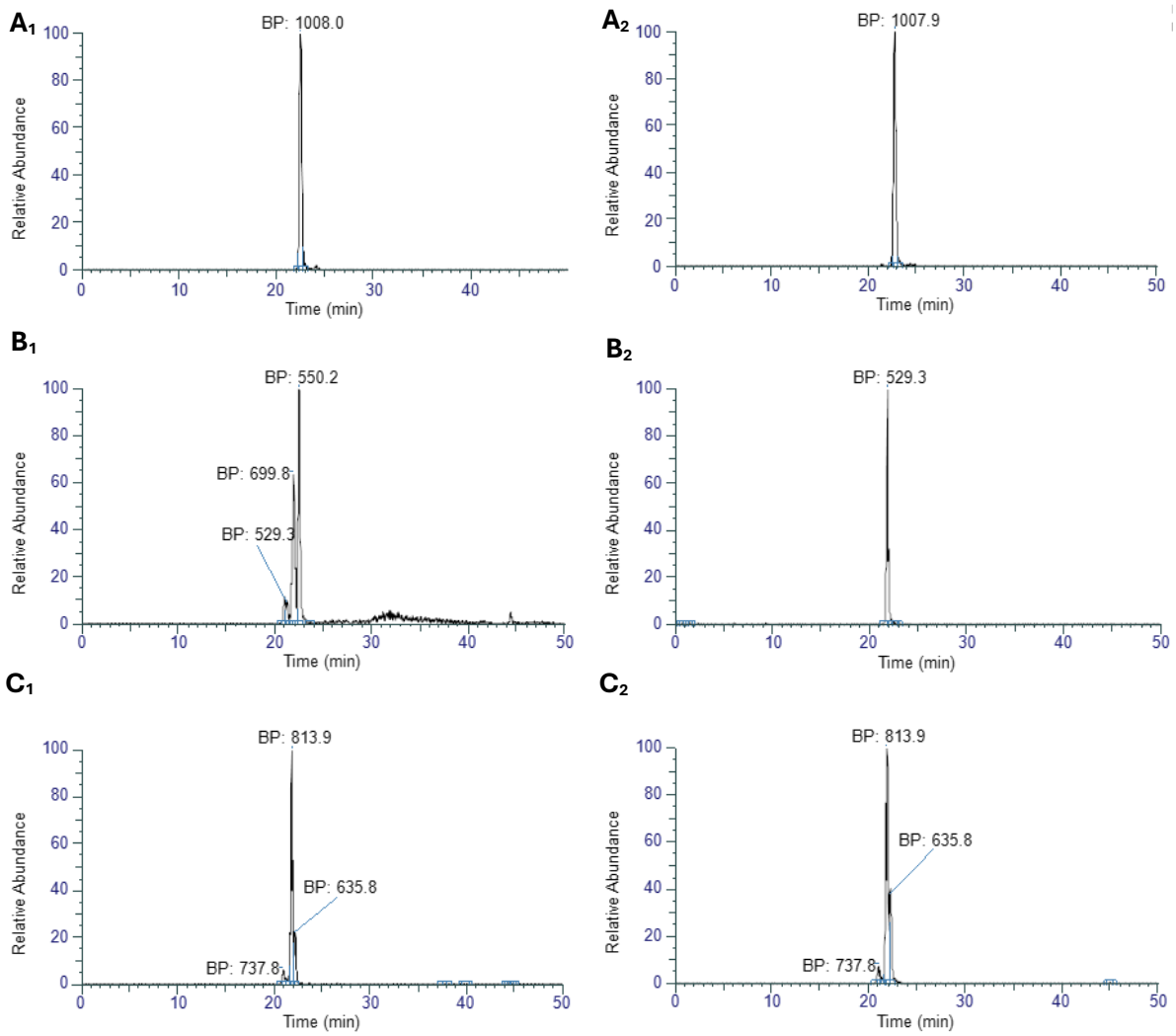


Figure 9.22: LC-MS base peak chromatograms of (A) oxytocin, (B) [Lys^8]-vasopressin and (C) terlipressin
after elution through a column packed with (1) an unfunctionalised and (2) a pepsin immobilised propylamino silica resin

The labelled base peaks LC-MS chromatograms of the fractions collected after elution of each of the cyclic neuropeptides through both a dead column and a pepsin immobilised reactor are outlined in Figure 9.22. Upon initial examination of the spectra, it was apparent that no species detected after exposure to the immobilised pepsin was not observed after elution through an unfunctionalised column. In the case of oxytocin (Figure 9.22A$_{1\text{-}2}$), a single 1008.0 m/z value representing the peptide monomer was detected, but for both [Lys^8]-vasopressin and terlipressin, multiple peaks were observed. In the case of vasopressin (Figure 9.22B$_{1\text{-}2}$), the 529.3 and 550.2 m/z values represented the 2+ monomer peak both without and with an acetonitrile

adduct, respectively, whilst the 699.8 m/z value represented a reacted dimer. The difference between the dead and pepsin functionalised column results were assumed to be due to the difference in time between column elution and LC-MS analysis allowing for dimerization and forming of configurations that resulted in increased acetonitrile binding. Terlipressin was observed to show no difference between elution through the pepsin-functionalised and unfunctionalised columns (Figure 9.22$C_{1\text{-}2}$). The 737.8 m/z value was representative of the 5+ charge of the terlipressin trimer; however, several other oligomers were observed to coelute with that species. It was assumed that their observation is due to the presence of the peptide building up in the detector causing the misdetection of several different oligomers. The subsequent 813.9 and 635.8 m/z values were, similarly to vasopressin, the reacted dimer 3+ and the 2+ charge of the monomer with acetonitrile adduct, respectively. Therefore, it can be simply assumed that their cyclic structures prevented pepsin proteolytic degradation, since the presence of the immobilised pepsin had no effect on the observed species.

9.9.4. Bradykinin Fragments

In addition to previous tested peptides, two bradykinin fragment variants were also eluted through the immobilised pepsin reactor. The species were of interest to study the activity and selectivity of the immobilised pepsin as they are of similar size, in terms of primary amino acid sequence length, to P5 and P6, whilst containing a Phe-Ser peptide bond which should promote pepsin cleavage. The difference in position of the Phe-Ser peptide bond, change in location of the arginine and the additional lysine should allow for the study to establish the impacts of these structural changes on the peptide structure susceptibility to pepsin enzymatic cleavage. However, due to the fact that both bradykinin fragments are rich in proline and contain an arginine residue, both of which inhibit pepsin binding, it is predicted that there may be limited observed reaction.

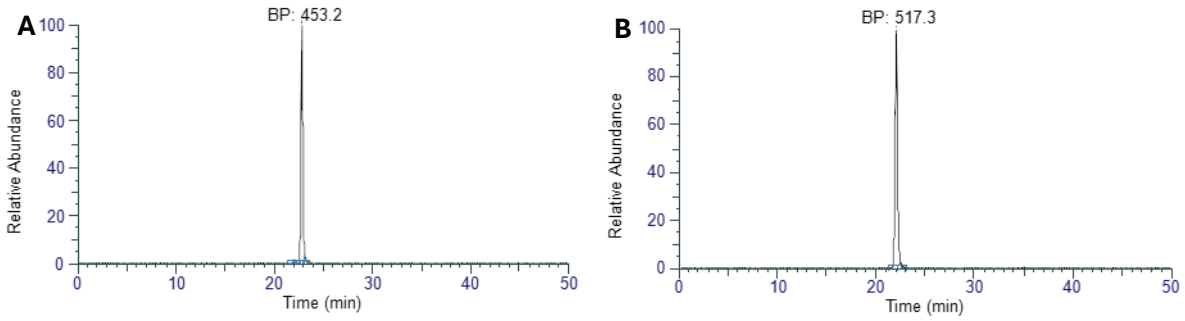


Figure 9.23: LC-MS base peak chromatogram of (A) bradykinin(2-9) and (B) Lys-des-Arg^9-bradykinin after elution through an immobilised pepsin reactor

As observed in the exemplary LC-MS base peak chromatograms in Figure 9.23A and B, neither bradykinin fragment displayed any noticeable reaction upon elution through the immobilised pepsin reactor, with the 2+ m/z value of the original peptide being the most abundant peak for both peptides. As previously discussed, this was likely due to the inhibiting effects of both the positively charged arginine (and lysine in the case of lys-des-arg^9-bradykinin) and the three proline residues [340, 341]. Given their presence and similar location in both sequences, the proline residues were the most likely cause of the lack of enzyme activity on the peptide structures, due to the absence of a hydrogen-bonding proton within its amide bond and the lack of rotational flexibility of its N-C$^{\alpha}$ bond [342]. Furthermore, the location of the Pro-Phe bond in both species despite being a potential cleavage site as observed in Figure 9.24 in both cases was too close to the C-terminus, resulting in poorer pepsin binding at this site.

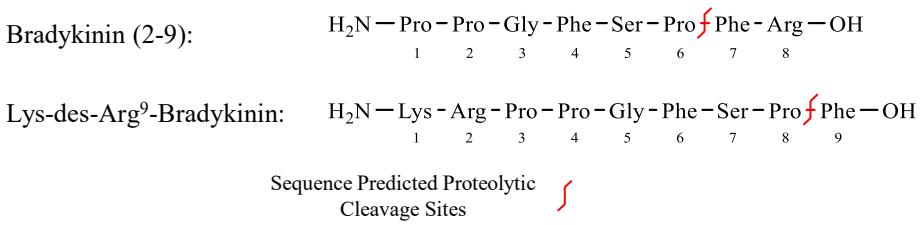


Figure 9.24: Bradykinin (2-9) and Lys-des-Arg^9-bradykinin sequences and their predicted pepsin proteolytic cleavage sites
as predicted by sequence analysis (see Table 5.3)

9.9.5.Substance P

Substance P was also eluted through the immobilised pepsin reactor to provide a comparison to the bradykinin fragments, wherein it similarly contains two proline and two positively charged (arginine and lysine) residues; however, they are all located in sequence towards the N-terminus, separated from the hydrophobic residues in proximity to the C-terminus by two glutamine residues. It is further of interest due to its amphipathic structure, which is similar to cell-penetrating structures that may be employed in oral peptide delivery strategies. Therefore, understanding the susceptibility of these sequences to enzymatic degradation and as such their stability within the gastrointestinal (GI) tract may be of importance to peptide delivery strategies.

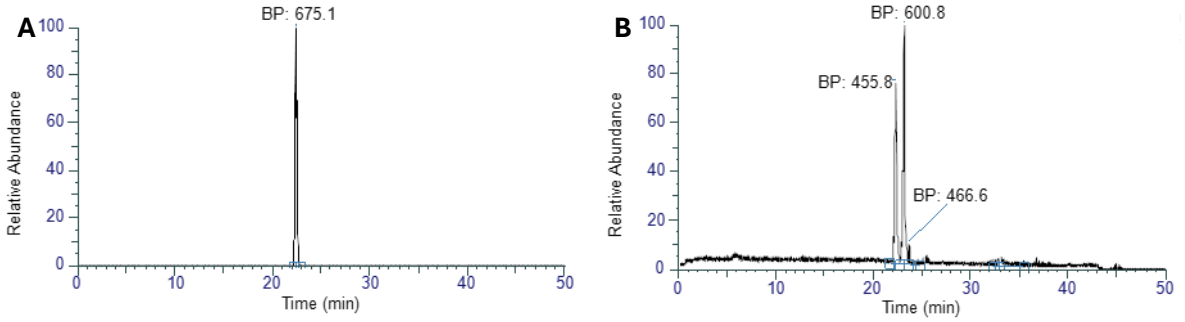


Figure 9.25: LC-MS base peak chromatograms of substance P after elution through a column packed with (A) an unfunctionalised and (B) a pepsin immobilised propylamino silica resin

As seen in Figure 9.25, relative to elution through an unfunctionalised resin, after which only the 2+ monomer m/z is detected, substance P undergoes a reaction when exposed to the immobilised pepsin. The two primary products observed were the 3+ charge of an Arg-1 – Phe7 dimer (m/z = 600.8) and the associated remaining fragment Phe-8 – Met-11 (m/z = 466.6). This was the point where pepsin cleavage would have been expected to occur, given the Phe-7 and Phe-8 residues at the P_1 and P'_1 sites respectively, both promote pepsin amide bond cleavage. Furthermore, these residues are sufficiently distanced from the proline and positively charged residue-rich region at the N-terminal to prevent their inhibition and are not at the C-terminus allowing for binding at all seven subsites of pepsin. Furthermore, similar proteolytic degradation

of substance P using an immobilised pepsin reactor was observed by Vorauer et al. [394]. This supports the observation that the preferred Phe-Phe cleavage site resulted in little observed degradation at the Gln-Phe bond, whilst the other sites are theorised to be too close to the C-terminal. However, the additional 455.8 m/z peak observed was the 3+ residue of the remaining original sample with an additional water adduct close to the N-terminus. Similar limited activity was observed by Vorauer et al. supporting the idea that despite the higher activity relative to the bradykinin fragments, there perhaps remained some relative enzymatic inhibiting action of the N-terminus residues.

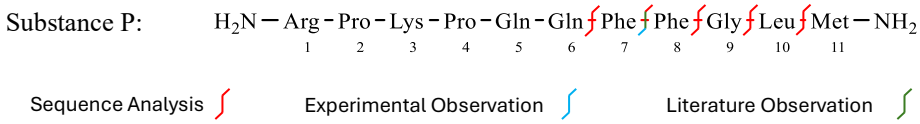


Figure 9.26: Substance P amino acid sequence and its proteolytic degradation sites
as predicted by sequence analysis (see Table 5.3) and observed after elution through immobilised pepsin reactor and by Vorauer et al. [394]

9.9.6. Gonadotropin-releasing hormone analogues

Finally, gonadotropin (luteinizing hormone) releasing hormone (LHRH) and its synthetic analogue Goserelin were also eluted through the immobilised enzyme reactor. This was to establish whether any stabilising effects would be observed due to the insertion of a tert-butylated d-amino acid close to the predicted cleavage site Tyr-5. Given the substitution of glycine with aza-glycine occurs at the C-terminal residue, it was not expected to greatly impact proteolytic cleavage rate since pepsin is an endopeptidase and the residues in proximity are Arg-8 and Pro-9, at which proteolytic cleavage was not expected. Furthermore, these residues may also inhibit the pepsin's action at the Leu-7 promoting residue, whilst the N-terminal pyroglutamic acid and subsequent histidine residues may inhibit proteolytic cleavage at the Trp-3 promoting residue due to their reduced hydrogen bonding capacity (due to the pyrrole ring similar to proline) and positive charge, respectively. Furthermore, the proximity of the Trp-3 residue to the N-terminus

likely inhibits proteolytic action by pepsin at this residue (contrary to sequence predictions observed in Figure 9.28). Therefore, the Tyr-5 – Gly-6 bond was the primary predicted site of proteolytic action, and so comparison with Goserelin should provide valuable feedback on the stabilising effects of the amino acid substitution to proteolytic action.

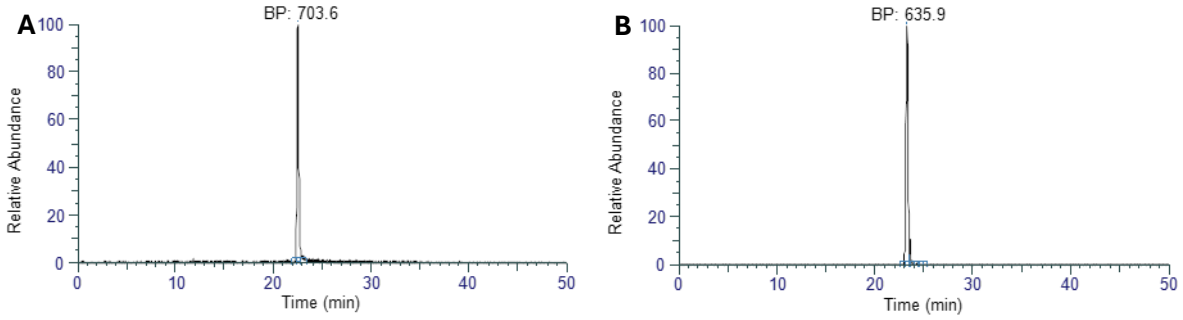


Figure 9.27: LC-MS base peak chromatogram of (A) Luteinizing Hormone releasing hormone (LHRH) and (B) Goserelin after elution through an immobilised pepsin reactor

As observed in Figure 9.27, after elution through the immobilised pepsin reactor, only one species was detected for both species. In the case of Goserelin (Figure 9.27B), the single 635.9 m/z value was indicative of the $(M+2H)^{2+}$ peak of the original peptide. By contrast, for LHRH the 703.6 m/z value was representative of Pyr-1 – Tyr-5, indicating proteolytic action of pepsin at the Tyr-5 – Gly-6 amide bond, the site predicted to be most susceptible. Therefore, this clearly establishes the stabilising effects that the substitution of Gly-6 in LHRH for d-Ser(tBu)-6 in Goserelin, likely due to the significantly reduced binding capacity of the enzyme due to the change in rotational bond angle of the amide bond for the d-amino acid. D-amino acid substitution has previously been shown to stabilise peptide protease degradation by Zhao et al. [257]. However, the additional tert-butyl group likely also sterically hindered pepsin's proteolytic action despite the increase in hydrophobicity that it provides, which (as previously outlined in Table 5.3) typically increases pepsin site proclivity. This reduction in activity due to tert-butylation, whilst not previously reported may be similar to the means by which methylation of alanine sterically hinders the proteolytic activity of dipeptidyl peptidase IV on semaglutide compared to native GLP-1 [140].

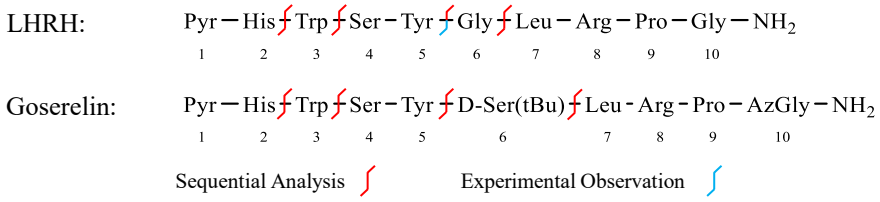


Figure 9.28: LHRH and Goserelin sequences and their proteolytic degradation sites *as predicted by sequence analysis (see Table 5.3) and observed after elution through immobilised pepsin reactor*

9.9.7.Summary of Observations

Table 9.6: Summary of observed immobilised pepsin reactions with each peptide compared to sequence analysis

Peptide	Sequence Analysis	Experimental Results	Comment on Differences
Semaglutide	☑	☑	
Liraglutide	☑	☑	
Taspoglutide	☑	☑	
Insulin	☑	☑	
Oxytocin	☑	☒	Cyclisation stabilises potential sites
Vasopressin	☑	☒	Cyclisation stabilises potential sites
Terlipressin	☑	☒	Cyclisation stabilises potential sites
Bradykinin (2-9)	☑	☒	Predicted site too close to C-terminus
Lys-des-Arg9-Bradykinin	☑	☒	Predicted site too close to C-terminus
Substance P	☑	☑	
LHRH	☑	☑	
Goserelin	☑	☒	Stabilised by D-amino acid

The results comparing the observation and prediction of proteolytic cleavage by the immobilised pepsin reactor for each peptide tested in Section 9.9 are summarised and compared in Table 9.6. Based on the differences in cleavage sites observed for each peptide (other than those where it was assumed that reaction was not observed due to a higher rate of degradation at a nearby residue inhibiting cleavage), the inhibiting and promoting effects outlined in Table 5.3, may be updated to incorporate the suggested reasons. Therefore, an updated list of the effects of the

amino acids within the sequence have on pepsin activity was produced and is outlined in Table 9.7

Table 9.7: Updated list of the effects substrate residue location has on the likelihood of proteolytic activity of pepsin

Residue	Promoting	Inhibiting
P_4	-	-
P_3	**Lys-Lipid**	Arg, Lys, His, **Cys-Cys**
P_2	-	Pro, **Cys-Cys**
P_1	Phe, Tyr, Trp, Leu, Met	Lys, His, Arg, Pro, **n<4, n>N-3, Cys-Cys**
P'_1	Phe, Tyr, Trp, Leu	**d-Xxx, Cys-Cys**
P'_2	**Lys-Lipid**	Pro, **Cys-Cys**
P'_3	-	Pro

n = residue number, N = total number of residues, Xxx = Any amino acid

9.10. Conclusions

The work in this Chapter developed and determined the applicability of a pepsin immobilised reactor to study the stability of peptides to pepsin proteolytic degradation under gastric conditions. UV-Vis spectroscopy first established that pepsin could be immobilised onto a propylamino silica resin with a very high loading (>150 mg/g.resin) with it being confirmed that it was at least in part due to covalent immobilisation via Raman spectroscopy. Initial studies with peptide hydrophobicity standards showed that the immobilised pepsin exhibited similar selectivity to solution reactions with the most hydrophobic peptides exhibiting the greatest proteolytic degradation at the same sites. Further qualitative studies with semaglutide showed that whilst similar cleavage site observations were made, some fragments were not observed from the solution reactions which were present in the immobilised enzyme reactions. These were assumed to be lost during enzyme precipitation of the solution reactions and so not observed during LC-MS analysis. Columns of various loadings between 0 and 450 mg/g.resin were synthesised by varying the initial concentration of pepsin in the reaction solution. These were applied to semaglutide to establish if any showed a similar product profile to the solution reactions was observed, although none did, which confirmed that the previous difference was

due to fragment loss during precipitation. However, the correlation between the loading and the product profile established that higher loadings, possibly exhibited lower overall activities, but increased the probability of additional substrate fragment enzyme interactions and so exhibited greater relative abundancies of smaller degradation fragments.

Quantitative comparison of the reaction velocity and Michaelis-Menten constant of the immobilised and free pepsin to P6 and semaglutide showed a greater than 50-fold decrease in the relative activity of the enzyme with a greater than 5x reduction in affinity for the substrate. However, due to the >500x increase in relative concentration of the enzyme when immobilised, this reduction in activity was compensated for. The affinity for and activity to P5, P6 and semaglutide of the immobilised pepsin columns were further found to decrease over the course of 12 weeks, with an initial larger decrease followed by a relatively linear reduction between weeks 2 and week 8 after pepsin immobilisation, represented by an average increase of 5.9% in the Michaelis-Menten constant and decrease of 4.8% in the maximum reaction velocities. However, the activity to P4 was observed to be almost undetectable after only 3 weeks. An expanded peptide study using the immobilised pepsin reactor showed other GLP-1 analogues exhibited similar proteolytic degradation site specificity to semaglutide, with the conjugated lipid increasing the susceptibility of the Ala-24-Ala-25 amide bond. Insulin was also observed to be degraded. However, cyclic species oxytocin and vasopressin exhibited resistance to proteolytic degradation. Similarly, proline and lysine rich positively charged species such as bradykinin fragments and the amphipathic structure of substance P exhibited stability to pepsin degradation. D-amino acid substitution was found to be important in protecting Goserelin from proteolytic degradation relative to LHRH, which was proteolytically cleaved at the specific substitution site.

Therefore, the immobilised pepsin reactor showed good potential for the basis of a high-throughput method to analyse the gastric stability of peptides over an extended period of time

with a similar activity profile to solution reactions but without the requirement for enzyme separation post reaction. However, in the future a more detailed study of different conditions to establish its applicability under more detailed gastric conditions may be required to examine if it maintains its good correlation with solution reactions. Furthermore, since insufficient data was collated during the varied pepsin loading experiments to study the changes in relative activity of the enzyme with increased immobilisation levels, these experiments may require repeating to correlate K_m and V_{max} with pepsin loading. An additional future study of the immobilised pepsin reactor to establish whether it permanently loses its activity after exposure to enzyme inhibitors would also be beneficial to determine whether the method may be applicable to the study of complex formulations. The method nevertheless shows potential and could in future be optimised by being integrated directly into an LC-MS set-up to remove the requirement for fractional collection, such that reaction, separation and analysis all occur in one step.

CHAPTER 10

10. CONCLUSIONS AND FUTURE WORK

10.1. Summary and Conclusions

This thesis examines the feasibility and accuracy of chromatographic and packed-bed methods to assess the physical and proteolytic stability of peptide therapeutics under gastrointestinal conditions.

The thesis begins with a review of the literature pertaining to mechanisms of peptide aggregation, the factors affecting it and the methods of detecting it. This Chapter helps to establish the applicability of self-interaction chromatography in the research of peptide aggregation propensity, outlining the advantages of its automated data collection for multiple conditions, lower peptide sample consumption. The subsequent Chapter reviews the literature regarding the conditions within the gastrointestinal tract with a focus on the proteolytic enzymes present and the methods of measuring their activity. The various techniques of enzyme immobilisation and the benefits of immobilised enzyme reactors compared to solution reactions are outlined: the lack of requirement for substrate separation, the reduced sample consumption and automated processing.

In Chapter 5, the materials consumed throughout the experimental work of the thesis are described with particular focus on the structures of insulin, the glucagon-like peptide-1 (GLP-1) analogues, semaglutide and liraglutide, and pepsin, including a detailed description of its mechanism. The Chapter then continues to describe the experimental procedures carried out to collect data including the HPLC setup, peptide and protein immobilisation, self-interaction

chromatography second virial coefficient calculation and the liquid chromatography-mass spectrometry (LC-MS) methods.

The experimental section of the thesis begins in Chapter 6, in which the development of the self-interaction chromatography protocol is discussed. It was found that the AF-Tresyl-650M resin with Tris end-capping, which has been applied in previous self-interaction chromatography studies on monoclonal antibodies, showed applicability as a stationary phase for a model peptide, semaglutide. The difference in retention times of the various end-capping methods tested established that 60 µL column volume was sufficient for semaglutide self-interaction chromatography, whilst a superficial flowrate of 0.12 mm/s was optimal. These results would be carried forward to the future self-interaction chromatography (SIC) experiments and provide understanding in the development of immobilised enzyme reactors.

Chapter 7, builds upon the previous, testing self-interaction chromatography of semaglutide under various intestinal conditions. The second virial coefficient results showed a similar trend to size distribution data collected via dynamic light scattering with pH showing the greatest effect on stability, with pH 8.0 being most optimal for minimising semaglutide aggregation. Higher lipid and lower salt concentrations were also found to stabilise the lipid conjugated peptide. However, there was a discrepancy between the magnitude of the SIC and dynamic light scattering (DLS) values, as the second virial coefficients predicted monomeric structures, whilst oligomers were measured after a few hours in solution. This was assumed to be due to SIC measuring the strength of the interactions between the immobilised monomer and the solution oligomers, rather than monomer-monomer interactions. Therefore, it was deemed that the technique showed potential to be applied in predicting the formation of larger aggregates but required prior understanding of the solution conformation to understand the results. The unexpected challenges in understanding the SIC data sets, was significantly linked the complexity of the peptide solution association processes.

Chapter 8 follows on from this work, assessing whether self-interaction chromatography would show comparable results and limitations with another GLP-1 peptide, liraglutide, and a larger peptide of interest, insulin. In the case of liraglutide, the technique predicted more negative second virial coefficients (increased aggregation propensity) at pH ≥7.0, which contradicted size exclusion and dynamic light scattering data, which showed the presence of 8-mers, which began to dissociate at pH <7.0 to form 13-mers. This was hypothesised to be due to the monomer interacting more strongly with the 8-mer than the 13-mer in the mobile phase, since what was being calculated was not the monomer second virial coefficient but the oligomer interaction strength with the monomer. Therefore, the results, similarly to those for semaglutide, agreed with those using standard techniques and those in literature, since liraglutide is known to adopt those two oligomeric structures prior to aggregation, with the octamer being more stable but with increased interaction strength with the monomer. By contrast, since insulin was observed to form hexamers instantly upon dissolution and large aggregates at pH <6.5, self-interaction chromatography was found to be an unsuitable technique. The results did not correlate well with dynamic light scattering and size exclusion chromatography particle size distributions. Peak splitting and strong retention were observed increasingly at lower pHs such that there was no observable peak at pH <7.0. However, for pH >7.0, the correct trend of increasing stability with pH, was observed via SIC, despite the differences in overall predicted values.

Therefore, the overall results from Chapters 7 and 8 suggest that whilst SIC may be a useful tool in analysing the stability of a peptide therapeutics under given conditions, it requires detailed understanding of the molecular association processes for specific peptides, including the relationship to the solution conditions especially pH. Therefore, at concentrations where the peptides are known to maintain a monomeric structure under intestinal conditions, the technique may show more proficiency. However, in its current form, this is not the case and much of the SIC data generated was difficult to interpret reliably. Nevertheless, with further

development the technique may still be applicable as a higher throughput method that may be applied during early-stage oral peptide formulation screening.

In Chapter 9, the work focused on the use of immobilised pepsin reactors to study the proteolytic stability of peptides under gastric conditions. A propylamino silica resin was functionalised with pepsin via a Schiff's base reaction as quantified and confirmed via UV-Vis and Raman spectroscopy, respectively. Initial experiments showed that qualitatively the immobilised pepsin showed similar selectivity to the free solution enzyme for a library of peptide hydrophobicity standards. However, for a larger peptide, semaglutide more products were observed relative to solution-based approaches, assumed to be due to specific sample loss during precipitation of solution enzymes and showing a benefit to the immobilised reactor setup. A greater than 50x decrease in maximum reaction velocity per mass enzyme and a 5x increase in the Michaelis-Menten constant, and so reduction in enzyme affinity were observed after immobilisation. However, this was compensated for by the >500x increase in concentration that immobilisation allowed for. The immobilised pepsin was shown to still show respectable levels of activity after 12 weeks, with an average reduction of 5.9% in maximum reaction rate and increase of 4.8% in the Michaelis-Menten constant occurring per week (57 sample elutions). Furthermore, the immobilised pepsin exhibited strong proteolytic reaction predicative behaviour for a range of peptides including the stability of cyclic and proline-rich peptides and the importance of the d-amino acid substitution which stabilises Goserelin relative to luteneizing hormone releasing hormone. Therefore, the method was shown that it had good promise to be used as a more high-throughput technique for studying peptide susceptibility to pepsin-catalysed degradation, which maintains selectivity but removes the requirement for sample-enzyme separation post-reaction.

10.2. Future Work

As outlined previously in the conclusions, this thesis discusses testing and developing the application of self-interaction chromatography and immobilised enzyme reactors in studying peptide species under gastrointestinal conditions. However, both methods could still be developed substantially further.

One of the major studies that would need to be carried out to improve self-interaction chromatography as a method for determining peptide aggregation propensity is analysing the effects of peptide coverage and injection sample concentration. From the experimental work carried out in Chapters 7 and 8, it was evident that whilst the technique could detect changes in peptide physical behaviour, under slightly unstable conditions the high concentration of peptide immobilised within the column initiated oligomerisation or aggregation of the mobile phase species. Therefore, as outlined in Chapter 8 with respect to insulin, this in-column aggregation resulted in peak splitting or a lack of peptide elution entirely at the concentration injected and surface coverage tested. Since peptides are more sensitive than larger proteins (such as monoclonal antibodies) to abrupt changes in aggregation propensity with changing conditions, due to their stronger interactions for their respective sizes, the technique needs to be adjusted. This would ensure the column fixated with peptide maintains the ability to differentiate between stable monomer species and slow aggregation or the formation of oligomers without initiating large-scale aggregation. Therefore, as previously mentioned, this would require a study analysing the effects of column loading and injected sample concentration. The injected sample would need to be in the concentration range which is sufficiently above the limit of detection but significantly below that which instantaneously forms aggregates under the range of conditions being tested. This concentration would then be eluted through columns of several different loadings under various conditions from stable to slightly unstable to highly unstable. This would assist in establishing the level of loading that is not so low that it does not resolve the differences in stability and so may not reliably measure the change in second virial coefficient for differing

conditions. It would also determine the level of loading which is too high resulting in oversensitivity of the column due to rapid aggregation due to interaction with the monomer species on the surface.

To confirm whether the hypothesis of in-column oligomerisation or aggregation was in fact what was being observed during previous experiments in this thesis, future experiments could be performed to study the secondary and tertiary structures of the sample pre- and post-injection. Circular dichroism could be performed on the original injection sample and fractionally collected self-interaction chromatography eluted peaks. However, an assumption that the initial sample concentration measurement is accurate and that the total injected amount is collected in the fractionally collected peak to ensure that the concentration of the fractionally collected sample is accurate. Furthermore, due to salt concentration limitations of circular dichroism, this may only be performed on conditions at low salt concentrations. Nevertheless, since the buffer concentrations were only 20 mM in most cases, this would allow for changes in secondary structure of the peptide during column elution to be established. For example, if a decrease in negative signal at 208 nm or 222 nm is observed and replaced by a slight positive peak at 212 nm between the original and fractionally collected sample, then a transition from more α-helix behaviour to more random coil structures has been observed. Furthermore, this could be correlated to the aggregation behaviour by analysing the hydrodynamic radii of the fractionally collected species via dynamic light scattering. Changes in the tertiary structure of the peptide may be outlined via nuclear Overhauser effect spectroscopy nuclear magnetic resonance experiments which indicate molecular protons that are in proximity to each other and so the folding structure of the peptide. However, if details of the tertiary structure aren't required and only the relative amount of folding, then heteronuclear single quantum coherence nuclear magnetic resonance experiments may be carried out instead. The relative amount of chemical shift dispersion for amide protons will indicate the diversity of the environments they are in, and so the relative amount of peptide folding.

To optimise the self-interaction chromatography setup further, a longer column of lower loading may be applied to provide sufficient surface peptide interaction for peak resolution but not sufficiently high to initiate rapid aggregation. To ensure that the self-interaction chromatography method is adaptable, this experiment would need to be performed for a range of varied peptide structures, to create a tabulated range of loadings appropriate for various different peptide structures such that the study is not required for each new entity. After establishing optimal resin loading for specific peptide structure types, the method may then be ready to screen formulations for a specific peptide under various simulated gastrointestinal conditions depending on the target absorption site. However, since it will have been tuned to potentially rapid aggregation conditions, it may require further analysis prior to use. Therefore, there is still a large amount of work that is required before self-interaction chromatography may be ready for oral peptide formulation screening. Nevertheless, in its current state it may still be a useful tool for studying peptide physical stability under a specific condition for which the peptide aggregation is well-understood, but these experiments are required to allow for more versatile implementation with the potential view to being used in screening oral formulations as described in Chapter 1.

To further establish the efficacy of the propylamino silica fixated pepsin reactor, a few further experiments need to be explored to ensure a detailed understanding of its predicative ability. Initially more detailed gastric conditions, including variations in osmolytes (e.g. sugars to remain suitable for mass spectrometry) and higher pHs to represent fed state conditions. These experiments have already been initiated but are not reported in this thesis. They will aim to establish whether the fixated pepsin maintains activity and selectivity relative to the free solution enzyme such that it may be applied more reliably. Furthermore, the stability of the immobilised pepsin will as in Chapter 9 be tested for an extended time period under this more broad range of conditions to establish whether the various conditions affects the rate of loss of activity and what the limits of the range of conditions are. In addition to these studies, it would be apt to study in

more detail (more quantitatively) the effects varied pepsin loadings have upon the activity of the enzyme and its stability with time. Additional studies comparing the columns performance against any commercially available or previously reported in literature may also be carried out to verify the pros and cons of different stationary phases and immobilisation techniques. Furthermore, the entire study, as outlined in Chapter 9 may be repeated to investigate the immobilisation of intestinal proteolytic enzymes and their respective conditions. Initial studies, not reported in this thesis, have begun on the immobilisation of trypsin to establish the effect covalent immobilisation has on its activity.

Upon completion of studying the effects various parameters have on the quality of the immobilised pepsin reactor, it may then be applied to study various formulations for a given peptide targeting gastric delivery including those with pepsin inhibitors. These studies would allow for it to be established whether any of the additives would permanently, especially any pepsin inhibitors, cause a loss of proteolytic activity and if the activity could be regenerated. If the immobilised enzyme shows a similar drop in activity to the free enzyme upon exposure to formulation additives but its activity is able to be regenerated or the loss is not permanent, then the reactor would be ideal as a potential screening tool for oral peptide formulations targeting gastric delivery.

An additional idea would be then to investigate whether the size of the immobilised pepsin reactor could be reduced to form a microscale reactor, thereby requiring less sample and so reducing potential costs further. This would require fewer experiments to establish its effectiveness as a range of conditions could be tested and compared to previous data to determine any changes in proteolytic activity. However, it may require a pre-packed column or a different packing procedure if very low diameter, microvolume reactors are being applied. Nevertheless, this strategy could be then further implemented directly into an LC-MS or nano-LC-MS setup so as to remove the requirement for fractional collection. It would require a system

with a mobile phase switching valve such that the injected sample may pass through the reactor under a simulated gastric condition and being adsorbed onto the reversed-phase column, prior to the mobile phase switching to an acetonitrile/water gradient to separate the degradation products. In this way, the immobilised pepsin would not be exposed to organic solutions likely resulting in its permanent loss of activity but would enable an all-in-one reaction, separation and detection setup.

CHAPTER 11

11. REFERENCES

[1] N. Sewald and H. Jukubke, "Fundamental Chemical and Structural Principles," in *Peptides: Chemistry and Biology*, 2 ed. Weinheim: Wiley-VCH, 2009, ch. 2, pp. 5-62.

[2] D. J. Drucker, "The biology of incretin hormones," *Cell Metab,* vol. 3, no. 3, pp. 153-65, Mar 2006, doi: 10.1016/j.cmet.2006.01.004.

[3] S. Suvas, "Role of Substance P Neuropeptide in Inflammation, Wound Healing, and Tissue Homeostasis," *J Immunol,* vol. 199, no. 5, pp. 1543-1552, Sep 1 2017, doi: 10.4049/jimmunol.1601751.

[4] Q. Hua, "Insulin: a small protein with a long journey," *Protein Cell,* vol. 1, no. 6, pp. 537-51, Jun 2010, doi: 10.1007/s13238-010-0069-z.

[5] J. Lau *et al.*, "Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide," *J Med Chem,* vol. 58, no. 18, pp. 7370-80, Sep 24 2015, doi: 10.1021/acs.jmedchem.5b00726.

[6] N. W. Shammas, "Bivalirudin: pharmacology and clinical applications," *Cardiovasc Drug Rev,* vol. 23, no. 4, pp. 345-60, Winter 2005, doi: 10.1111/j.1527-3466.2005.tb00177.x.

[7] K. Fosgerau and T. Hoffmann, "Peptide therapeutics: current status and future directions," *Drug Discov Today,* vol. 20, no. 1, pp. 122-8, Jan 2015, doi: 10.1016/j.drudis.2014.10.003.

[8] A. Mocsai, L. Kovacs, and P. Gergely, "What is the future of targeted therapy in rheumatology: biologics or small molecules?," *BMC Med,* vol. 12, p. 43, Mar 13 2014, doi: 10.1186/1741-7015-12-43.

[9] D. E. Scott, A. R. Bayly, C. Abell, and J. Skidmore, "Small molecules, big targets: drug discovery faces the protein-protein interaction challenge," *Nat Rev Drug Discov,* vol. 15, no. 8, pp. 533-50, Aug 2016, doi: 10.1038/nrd.2016.29.

[10] Z. Farooq, L. A. Howell, and P. J. McCormick, "Probing GPCR Dimerization Using Peptides," *Front Endocrinol (Lausanne),* vol. 13, p. 843770, 2022, doi: 10.3389/fendo.2022.843770.

[11] L. Pelz, S. Göbel, K. Jaen, U. Reichl, and Y. Genzel, "Upstream processing for viral vaccines–General aspects," in *Bioprocessing of Viral Vaccines*, 2022, pp. 79-135.

[12] P. Poduval, S. Parsekar, and S. N. Meena, "Small molecules vs biologics," in *New Horizons in Natural Compound Research*, 2023, pp. 179-199.

[13] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," *Adv Drug Deliv Rev,* vol. 46, no. 1-3, pp. 3-26, Mar 1 2001, doi: 10.1016/s0169-409x(00)00129-0.

[14] Y. Heled, "The Case for Disclosure of Biologics Manufacturing Information," *J Law Med Ethics,* vol. 47, no. 4_suppl, pp. 54-78, Dec 2019, doi: 10.1177/1073110519898043.

[15] M. C. Smith and J. E. Gestwicki, "Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity," *Expert Rev Mol Med,* vol. 14, p. e16, Jul 26 2012, doi: 10.1017/erm.2012.10.

[16] R. F. Beall, T. J. Hwang, and A. S. Kesselheim, "Pre-market development times for biologic versus small-molecule drugs," *Nat Biotechnol,* vol. 37, no. 7, pp. 708-711, Jul 2019, doi: 10.1038/s41587-019-0175-2.

[17] S. Sachdeva, "Peptides as 'Drugs': The Journey so Far," *International Journal of Peptide Research and Therapeutics,* vol. 23, no. 1, pp. 49-60, 2016, doi: 10.1007/s10989-016-9534-8.
[18] N. Fotouhi, "Peptide Therapeutics," in *Peptide Chemistry and Drug Design*, 2015, pp. 1-10.
[19] A. Henninot, J. C. Collins, and J. M. Nuss, "The Current State of Peptide Drug Discovery: Back to the Future?," *J Med Chem,* vol. 61, no. 4, pp. 1382-1414, Feb 22 2018, doi: 10.1021/acs.jmedchem.7b00318.
[20] R. B. Merrifield, "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide," *Journal of the American Chemical Society,* vol. 85, no. 14, pp. 2149-2154, 2002, doi: 10.1021/ja00897a025.
[21] L. Diao and B. Meibohm, "Pharmacokinetics and pharmacokinetic-pharmacodynamic correlations of therapeutic peptides," *Clin Pharmacokinet,* vol. 52, no. 10, pp. 855-68, Oct 2013, doi: 10.1007/s40262-013-0079-0.
[22] F. G. Banting, C. H. Best, J. B. Collip, W. R. Campbell, and A. A. Fletcher, "Pancreatic Extracts in the Treatment of Diabetes Mellitus," *Canadian Medical Association Journal,* vol. 12, no. 3, pp. 141–146, March 1922 1922.
[23] U.S. Food & Drug Administration. "Orange Book: Approved Drug Products with Therapeutic Equivalence Evaluations." https://www.accessdata.fda.gov/scripts/cder/ob/index.cfm (accessed.
[24] European Medicines Agency. Medicine Data [Online] Available: https://www.ema.europa.eu/en/medicines/download-medicine-data
[25] IQVIA, "Trends in the Prescription Medicines Market," in "IQVIA Quarterly Pharmaceutical Market Outlook Report, Q3 2024," 2024. [Online]. Available: https://www.iqvia.com/-/media/iqvia/pdfs/library/white-papers/iqpmo-newsletter-f3q24-iqvia-quarterly-update-trends-in-the-prescription-medicines-marketet.pdf
[26] Research And Markets, "Peptide Therapeutics Market Report 2025," 2025. [Online]. Available: https://www.researchandmarkets.com/reports/5767340/peptide-therapeutics-market-report
[27] Grand View Research, "Peptide Therapeutics Market Size, Industry Report, 2030," 2025. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/peptide-therapeutics-market
[28] Research And Markets, "Oral Proteins & Peptides Market by Molecule Global Forecast to 2030," 2025. Accessed: 23/06/2025. [Online]. Available: https://www.researchandmarkets.com/reports/6019674/oral-proteins-and-peptides-market-molecule?utm_source=GNE&utm_medium=PressRelease&utm_code=97dnrf&utm_campaign=2056005+-+Oral+Proteins+%26+Peptides+Market+Research+Report+2025%3a+Semaglutide%2c+Linaclotide%2c+Calcitonin%2c+GLP-1+Receptor+Agonist%2c+GEP%2c+CGRP%2c+Diabetes%2c+Gastroenterology%2c+Genetic+Disorder+Forecast+to+2029&utm_exec=chdomspi
[29] M. Speers and C. Bonnano, "Economic Aspects of Controlled Drug Delivery," in *Encyclopedia of Controlled Drug Delivery*, vol. 1, E. Mathiowitz Ed.: John Wiley & Sons, Inc., 1999, pp. 341-347.
[30] E. Thomaidou and Y. Ramot, "Injection site reactions with the use of biological agents," *Dermatol Ther,* vol. 32, no. 2, p. e12817, Mar 2019, doi: 10.1111/dth.12817.
[31] M. Peyrot, A. H. Barnett, L. F. Meneghini, and P. M. Schumm-Draeger, "Insulin adherence behaviours and barriers in the multinational Global Attitudes of Patients and Physicians in Insulin Therapy study," *Diabet Med,* vol. 29, no. 5, pp. 682-9, May 2012, doi: 10.1111/j.1464-5491.2012.03605.x.

[32] E. Tetteh and S. Morris, "Systematic review of drug administration costs and implications for biopharmaceutical manufacturing," *Appl Health Econ Health Policy,* vol. 11, no. 5, pp. 445-56, Oct 2013, doi: 10.1007/s40258-013-0045-x.
[33] A. Abramson, F. Halperin, J. Kim, and G. Traverso, "Quantifying the Value of Orally Delivered Biologic Therapies: A Cost-Effectiveness Analysis of Oral Semaglutide," *J Pharm Sci,* vol. 108, no. 9, pp. 3138-3145, Sep 2019, doi: 10.1016/j.xphs.2019.04.022.
[34] R. Ismail and I. Csoka, "Novel strategies in the oral delivery of antidiabetic peptide drugs - Insulin, GLP 1 and its analogs," *Eur J Pharm Biopharm,* vol. 115, pp. 257-267, Jun 2017, doi: 10.1016/j.ejpb.2017.03.015.
[35] R. E. Steinert *et al.*, "Orally administered glucagon-like peptide-1 affects glucose homeostasis following an oral glucose tolerance test in healthy male subjects," *Clin Pharmacol Ther,* vol. 86, no. 6, pp. 644-50, Dec 2009, doi: 10.1038/clpt.2009.159.
[36] E. Zijlstra, L. Heinemann, and L. Plum-Morschel, "Oral insulin reloaded: a structured approach," *J Diabetes Sci Technol,* vol. 8, no. 3, pp. 458-65, May 2014, doi: 10.1177/1932296814529988.
[37] S. Baliga, S. Muglikar, and R. Kale, "Salivary pH: A diagnostic biomarker," *J Indian Soc Periodontol,* vol. 17, no. 4, pp. 461-5, Jul 2013, doi: 10.4103/0972-124X.118317.
[38] M. Koziolek *et al.*, "Investigation of pH and Temperature Profiles in the GI Tract of Fasted Human Subjects Using the Intellicap((R)) System," *J Pharm Sci,* vol. 104, no. 9, pp. 2855-63, Sep 2015, doi: 10.1002/jps.24274.
[39] D. M. Mudie, G. L. Amidon, and G. E. Amidon, "Physiological parameters for oral delivery and in vitro testing," *Mol Pharm,* vol. 7, no. 5, pp. 1388-405, Oct 4 2010, doi: 10.1021/mp100149j.
[40] I. Pereira de Sousa and A. Bernkop-Schnurch, "Pre-systemic metabolism of orally administered drugs and strategies to overcome it," *J Control Release,* vol. 192, pp. 301-9, Oct 28 2014, doi: 10.1016/j.jconrel.2014.08.004.
[41] J. Wang, V. Yadav, A. L. Smart, S. Tajiri, and A. W. Basit, "Toward oral delivery of biopharmaceuticals: an assessment of the gastrointestinal stability of 17 peptide drugs," *Mol Pharm,* vol. 12, no. 3, pp. 966-73, Mar 2 2015, doi: 10.1021/mp500809f.
[42] R. Bansil and B. S. Turner, "Mucin structure, aggregation, physiological functions and biomedical applications," *Current Opinion in Colloid & Interface Science,* vol. 11, no. 2-3, pp. 164-170, 2006, doi: 10.1016/j.cocis.2005.11.001.
[43] V. Bonis, C. Rossell, and H. Gehart, "The Intestinal Epithelium - Fluid Fate and Rigid Structure From Crypt Bottom to Villus Tip," *Front Cell Dev Biol,* vol. 9, p. 661931, 2021, doi: 10.3389/fcell.2021.661931.
[44] M. A. Odenwald and J. R. Turner, "The intestinal epithelial barrier: a therapeutic target?," *Nat Rev Gastroenterol Hepatol,* vol. 14, no. 1, pp. 9-21, Jan 2017, doi: 10.1038/nrgastro.2016.169.
[45] C. J. Watson, M. Rowland, and G. Warhurst, "Functional modeling of tight junctions in intestinal cell monolayers using polyethylene glycol oligomers," *Am J Physiol Cell Physiol,* vol. 281, no. 2, pp. C388-97, Aug 2001, doi: 10.1152/ajpcell.2001.281.2.C388.
[46] S. M. Pond and T. N. Tozer, "First-pass elimination. Basic concepts and clinical consequences," *Clin Pharmacokinet,* vol. 9, no. 1, pp. 1-25, Jan-Feb 1984, doi: 10.2165/00003088-198409010-00001.
[47] A. Abramson *et al.*, "A luminal unfolding microneedle injector for oral delivery of macromolecules," *Nat Med,* vol. 25, no. 10, pp. 1512-1518, Oct 2019, doi: 10.1038/s41591-019-0598-9.
[48] A. Abramson *et al.*, "An ingestible self-orienting system for oral delivery of macromolecules," *Science,* vol. 363, no. 6427, pp. 611-615, Feb 8 2019, doi: 10.1126/science.aau2277.

[49] A. K. Dhalla *et al.*, "A robotic pill for oral delivery of biotherapeutics: safety, tolerability, and performance in healthy subjects," *Drug Deliv Transl Res,* vol. 12, no. 1, pp. 294-305, Jan 2022, doi: 10.1007/s13346-021-00938-1.
[50] S. Thakral, N. K. Thakral, and D. K. Majumdar, "Eudragit: a technology evaluation," *Expert Opin Drug Deliv,* vol. 10, no. 1, pp. 131-49, Jan 2013, doi: 10.1517/17425247.2013.736962.
[51] Q. Zhu, Z. Chen, P. K. Paul, Y. Lu, W. Wu, and J. Qi, "Oral delivery of proteins and peptides: Challenges, status quo and future perspectives," *Acta Pharm Sin B,* vol. 11, no. 8, pp. 2416-2448, Aug 2021, doi: 10.1016/j.apsb.2021.04.001.
[52] J. M. Smith, M. Dornish, and E. J. Wood, "Involvement of protein kinase C in chitosan glutamate-mediated tight junction disruption," *Biomaterials,* vol. 26, no. 16, pp. 3269-3276, 2005, doi: 10.1016/j.biomaterials.2004.06.020.
[53] N. Hoshyar, S. Gray, H. Han, and G. Bao, "The effect of nanoparticle size on in vivo pharmacokinetics and cellular interaction," *Nanomedicine (Lond),* vol. 11, no. 6, pp. 673-92, Mar 2016, doi: 10.2217/nnm.16.5.
[54] M. Hofmann-Amtenbrink, D. W. Grainger, and H. Hofmann, "Nanoparticles in medicine: Current challenges facing inorganic nanoparticle toxicity assessments and standardizations," *Nanomedicine,* vol. 11, no. 7, pp. 1689-94, Oct 2015, doi: 10.1016/j.nano.2015.05.005.
[55] N. G. Lamson, A. Berger, K. C. Fein, and K. A. Whitehead, "Anionic nanoparticles enable the oral delivery of proteins by enhancing intestinal permeability," *Nat Biomed Eng,* vol. 4, no. 1, pp. 84-96, Jan 2020, doi: 10.1038/s41551-019-0465-5.
[56] P. Ganesan and D. Narayanasamy, "Lipid nanoparticles: Different preparation techniques, characterization, hurdles, and strategies for the production of solid lipid nanoparticles and nanostructured lipid carriers for oral drug delivery," *Sustainable Chemistry and Pharmacy,* vol. 6, pp. 37-56, 2017, doi: 10.1016/j.scp.2017.07.002.
[57] X. Hu *et al.*, "Evidence does not support absorption of intact solid lipid nanoparticles via oral delivery," *Nanoscale,* vol. 8, no. 13, pp. 7024-35, Apr 7 2016, doi: 10.1039/c5nr07474f.
[58] S. Martins, S. Costa-Lima, T. Carneiro, A. Cordeiro-da-Silva, E. B. Souto, and D. C. Ferreira, "Solid lipid nanoparticles as intracellular drug transporters: an investigation of the uptake mechanism and pathway," *Int J Pharm,* vol. 430, no. 1-2, pp. 216-27, Jul 1 2012, doi: 10.1016/j.ijpharm.2012.03.032.
[59] S. Alqahtani, A. Alayoubi, S. Nazzal, P. W. Sylvester, and A. Kaddoumi, "Nonlinear absorption kinetics of self-emulsifying drug delivery systems (SEDDS) containing tocotrienols as lipophilic molecules: in vivo and in vitro studies," *AAPS J,* vol. 15, no. 3, pp. 684-95, Jul 2013, doi: 10.1208/s12248-013-9481-7.
[60] G. Leonaviciute and A. Bernkop-Schnurch, "Self-emulsifying drug delivery systems in oral (poly)peptide drug delivery," *Expert Opin Drug Deliv,* vol. 12, no. 11, pp. 1703-16, 2015, doi: 10.1517/17425247.2015.1068287.
[61] J. M. Kovarik, E. A. Mueller, J. B. van Bree, W. Tetzloff, and K. Kutz, "Reduced inter- and intraindividual variability in cyclosporine pharmacokinetics from a microemulsion formulation," *J Pharm Sci,* vol. 83, no. 3, pp. 444-6, Mar 1994, doi: 10.1002/jps.2600830336.
[62] D. A. Subramanian, R. Langer, and G. Traverso, "Mucus interaction to improve gastrointestinal retention and pharmacokinetics of orally administered nano-drug delivery systems," *J Nanobiotechnology,* vol. 20, no. 1, p. 362, Aug 6 2022, doi: 10.1186/s12951-022-01539-x.
[63] S. Takatsuka, T. Kitazawa, T. Morita, Y. Horikiri, and H. Yoshino, "Enhancement of intestinal absorption of poorly absorbed hydrophilic compounds by simultaneous use

of mucolytic agent and non-ionic surfactant," *Eur J Pharm Biopharm,* vol. 62, no. 1, pp. 52-8, Jan 2006, doi: 10.1016/j.ejpb.2005.07.008.
[64] A. Banerjee, K. Ibsen, T. Brown, R. Chen, C. Agatemor, and S. Mitragotri, "Ionic liquids for oral insulin delivery," *Proc Natl Acad Sci U S A,* vol. 115, no. 28, pp. 7296-7301, Jul 10 2018, doi: 10.1073/pnas.1722338115.
[65] K. Palanisamy and M. Prakash, "The molecular mechanism behind the stabilization of insulin by choline and geranate (CAGE) ionic liquids - computational insights into oral insulin drug formulation," *Phys Chem Chem Phys,* vol. 23, no. 44, pp. 25298-25307, Nov 17 2021, doi: 10.1039/d1cp03349b.
[66] S. Haddadzadegan, F. Dorkoosh, and A. Bernkop-Schnurch, "Oral delivery of therapeutic peptides and proteins: Technology landscape of lipid-based nanocarriers," *Adv Drug Deliv Rev,* vol. 182, p. 114097, Mar 2022, doi: 10.1016/j.addr.2021.114097.
[67] S. Maher, R. J. Mrsny, and D. J. Brayden, "Intestinal permeation enhancers for oral peptide delivery," *Adv Drug Deliv Rev,* vol. 106, no. Pt B, pp. 277-319, Nov 15 2016, doi: 10.1016/j.addr.2016.06.005.
[68] S. T. Buckley *et al.*, "Transcellular stomach absorption of a derivatized glucagon-like peptide-1 receptor agonist," *Sci Transl Med,* vol. 10, no. 467, Nov 14 2018, doi: 10.1126/scitranslmed.aar7047.
[69] S. Tuvia *et al.*, "A novel suspension formulation enhances intestinal absorption of macromolecules via transient and reversible transport mechanisms," *Pharm Res,* vol. 31, no. 8, pp. 2010-21, Aug 2014, doi: 10.1007/s11095-014-1303-9.
[70] International Federation of Pharmaceutical Manufacturers & Associations (IFPMA). "Facts and Figures 2022: The Pharmaceutical Industry and Global Health." https://www.ifpma.org/publications/facts-and-figures-2022-the-pharmaceutical-industry-and-global-health/ (accessed.
[71] D. Proudman *et al.*, "Public sector replacement of privately funded pharmaceutical R&D: cost and efficiency considerations," *J Med Econ,* vol. 27, no. 1, pp. 1253-1266, Jan-Dec 2024, doi: 10.1080/13696998.2024.2405407.
[72] W. Xiao *et al.*, "Advance in peptide-based drug development: delivery platforms, therapeutics and vaccines," *Signal Transduct Target Ther,* vol. 10, no. 1, p. 74, Mar 5 2025, doi: 10.1038/s41392-024-02107-5.
[73] H. Zhang *et al.*, "Deep Learning-Based Bioactive Therapeutic Peptide Generation and Screening," *J Chem Inf Model,* vol. 63, no. 3, pp. 835-845, Feb 13 2023, doi: 10.1021/acs.jcim.2c01485.
[74] C. F. Deacon, L. B. Knudsen, K. Madsen, F. C. Wiberg, O. Jacobsen, and J. J. Holst, "Dipeptidyl peptidase IV resistant analogues of glucagon-like peptide-1 which have extended metabolic stability and improved biological activity," *Diabetologia,* vol. 41, pp. 271-278, Oct. 20 1997 1998, doi: doi.org/10.1007/s001250050903.
[75] W. K. Feng, Sudipan; Zhou, Maotang, "Quality Considerations in Solid Phase Peptide Synthesis," in *2021 FDA Science Forum: Science as the Foundation for Protecting and Promoting Public Health*, Silver Spring, MD, USA, 2021: U.S. Food and Drug Administration,. [Online]. Available: https://www.fda.gov/media/149068/download. [Online]. Available: https://www.fda.gov/media/149068/download
[76] Y. Li, "Common Deficiencies Associated with Comparative Peptide Impurity Profile Studies and Qualification of Impurity Levels and Proposed Limits," in *FDA-SBIA Workshop: Advancing Generic Drug Development: Translating Science to Approval*, 2022: U.S. Food & Drug Adminisatration.
[77] J. Bowes *et al.*, "Reducing safety-related drug attrition: the use of in vitro pharmacological profiling," *Nat Rev Drug Discov,* vol. 11, no. 12, pp. 909-22, Dec 2012, doi: 10.1038/nrd3845.

[78] S. Fontana *et al.*, "Discovery Toxicology in Lead Optimization," in *The Handbook of Medicinal Chemistry*, 2023, pp. 533-596.
[79] A. Singh and C. Haskell-Luevano, "Peptide Design Strategies for G-Protein Coupled Receptors (GPCR)," in *Peptide Chemistry and Drug Design*, 2015, pp. 75-112.
[80] M. Rahban, F. Ahmad, M. A. Piatyszek, T. Haertle, L. Saso, and A. A. Saboury, "Stabilization challenges and aggregation in protein-based therapeutics in the pharmaceutical industry," *RSC Adv,* vol. 13, no. 51, pp. 35947-35963, Dec 8 2023, doi: 10.1039/d3ra06476j.
[81] S. Patel, V. K. Vyas, and P. J. Mehta, "A Review on Forced Degradation Strategies to Establish the Stability of Therapeutic Peptide Formulations," *International Journal of Peptide Research and Therapeutics,* vol. 29, no. 2, 2023, doi: 10.1007/s10989-023-10492-8.
[82] C. Dahley, T. Bockmann, A. Ebert, and K. U. Goss, "Predicting the intrinsic membrane permeability of Caco-2/MDCK cells by the solubility-diffusion model," *Eur J Pharm Sci,* vol. 195, p. 106720, Apr 1 2024, doi: 10.1016/j.ejps.2024.106720.
[83] G. V. Kryukov *et al.*, "Characterization of mammalian selenoproteomes," *Science,* vol. 300, no. 5624, pp. 1439-1443, 2003, doi: 10.1126/science.1083516.
[84] S. Damodaran, "Amino Acids, Peptides and Proteins," in *Fennema's Food Chemistry*, S. Damodaran and K. Parkin Eds., 5 ed.: CRC Press, 2017, ch. 5, pp. 235-356.
[85] K. Linderstrøm-Lang, *Lane Medical Lectures: Proteins and Enzymes*. Stanford University, 1952, p. 115.
[86] P. Faller and C. Hureau, "Reproducibility Problems of Amyloid-β Self-Assembly and How to Deal With Them," *Frontiers in Chemistry,* vol. 8, 2021, doi: 10.3389/fchem.2020.611227.
[87] T. M. Takahashi, Hisakazu, "Peptide and Protein Mimetics Inhibiting Amyloid β-Peptide Aggregation," *Accounts of Chemical Research,* vol. 41, no. 10, pp. 1309-1318, 2008, doi: 10.1021/ar8000475.
[88] W. M. Tay, D. Huang, T. L. Rosenberry, and A. K. Paravastu, "The Alzheimer's amyloid-beta(1-42) peptide forms off-pathway oligomers and fibrils that are distinguished structurally by intermolecular organization," *J Mol Biol,* vol. 425, no. 14, pp. 2494-508, Jul 24 2013, doi: 10.1016/j.jmb.2013.04.003.
[89] F. Chiti and C. M. Dobson, "Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade," *Annual Review of Biochemistry,* vol. 86, no. 1, pp. 27-68, 2017, doi: 10.1146/annurev-biochem-061516-045115.
[90] L. Xu, S. Shan, Y. Chen, X. Wang, R. Nussinov, and B. Ma, "Coupling of Zinc-Binding and Secondary Structure in Nonfibrillar Abeta40 Peptide Oligomerization," *J Chem Inf Model,* vol. 55, no. 6, pp. 1218-30, Jun 22 2015, doi: 10.1021/acs.jcim.5b00063.
[91] I. W. Hamley, "Self-assembly of amphiphilic peptides," *Soft Matter,* vol. 7, no. 9, 2011, doi: 10.1039/c0sm01218a.
[92] I. W. Fu, C. B. Markegard, B. K. Chu, and H. D. Nguyen, "The role of electrostatics and temperature on morphological transitions of hydrogel nanostructures self-assembled by peptide amphiphiles via molecular dynamics simulations," *Adv Healthc Mater,* vol. 2, no. 10, pp. 1388-400, Oct 2013, doi: 10.1002/adhm.201200400.
[93] A. Dehsorkhi, V. Castelletto, and I. W. Hamley, "Self-assembling amphiphilic peptides," *J Pept Sci,* vol. 20, no. 7, pp. 453-67, Jul 2014, doi: 10.1002/psc.2633.
[94] S. Kirkham *et al.*, "Self-Assembly of the Cyclic Lipopeptide Daptomycin: Spherical Micelle Formation Does Not Depend on the Presence of Calcium Chloride," *Chemphyschem,* vol. 17, no. 14, pp. 2118-22, Jul 18 2016, doi: 10.1002/cphc.201600308.

[95] J. A. Hutchinson, S. Burholt, and I. W. Hamley, "Peptide hormones and lipopeptides: from self-assembly to therapeutic applications," *J Pept Sci,* vol. 23, no. 2, pp. 82-94, Feb 2017, doi: 10.1002/psc.2954.
[96] M. Fändrich, "On the structural definition of amyloid fibrils and other polypeptide aggregates," *Cellular and Molecular Life Sciences,* vol. 64, no. 16, pp. 2066-2078, 2007, doi: 10.1007/s00018-007-7110-2.
[97] J. Lee, E. K. Culyba, E. T. Powers, and J. W. Kelly, "Amyloid-beta forms fibrils by nucleated conformational conversion of oligomers," *Nat Chem Biol,* vol. 7, no. 9, pp. 602-9, Jul 31 2011, doi: 10.1038/nchembio.624.
[98] N. Cremades *et al.*, "Direct observation of the interconversion of normal and toxic forms of alpha-synuclein," *Cell,* vol. 149, no. 5, pp. 1048-59, May 25 2012, doi: 10.1016/j.cell.2012.03.037.
[99] M. Bouchard, J. Zurdo, E. J. Nettleton, C. M. Dobson, and C. V. Robinson, "Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy," *Protein Sci,* vol. 9, no. 10, pp. 1960-7, Oct 2000, doi: 10.1110/ps.9.10.1960.
[100] G. M. Meisl, Thomas; Arosio, Paolo; Vendruscolo, Michele; Dobson, Christoph; Knowles, Tuomas, "Dynamics and Control of Peptide Self-Assembly and Aggregation," in *Biological and Bio-inspired Nanomaterials*, vol. 1174, S. B. Perrett, Alexander; Knowles, Tuomas Ed. Singapore: Springer, 2019, ch. 1-34.
[101] J. Adamcik and R. Mezzenga, "Amyloid Polymorphism in the Protein Folding and Aggregation Energy Landscape," *Angew Chem Int Ed Engl,* vol. 57, no. 28, pp. 8370-8382, Jul 9 2018, doi: 10.1002/anie.201713416.
[102] A. K. Srivastava *et al.*, "β-Amyloid aggregation and heterogeneous nucleation," *Protein Science,* vol. 28, no. 9, pp. 1567-1581, 2019, doi: 10.1002/pro.3674.
[103] D. Kashchiev and S. Auer, "Nucleation of amyloid fibrils," *J Chem Phys,* vol. 132, no. 21, p. 215101, Jun 7 2010, doi: 10.1063/1.3447891.
[104] T. Christopeit, P. Hortschansky, V. Schroeckh, K. Guhrs, G. Zandomeneghi, and M. Fandrich, "Mutagenic analysis of the nucleation propensity of oxidized Alzheimer's beta-amyloid peptide," *Protein Sci,* vol. 14, no. 8, pp. 2125-31, Aug 2005, doi: 10.1110/ps.051470405.
[105] M. Fandrich, V. Forge, K. Buder, M. Kittler, C. M. Dobson, and S. Diekmann, "Myoglobin forms amyloid fibrils by association of unfolded polypeptide segments," *Proceedings of the National Academy of Sciences (PNAS),* vol. 100, no. 26, pp. 15463-15468, Dec. 9 2003 2003, doi: 10.1073/pnas.0303758100.
[106] P. Marek *et al.*, "Aromatic Interactions Are Not Required for Amyloid Fibril Formation by Islet Amyloid Polypeptide but Do Influence the Rate of Fibril Formation and Fibril Morphology," *Biochemistry,* vol. 46, no. 11, pp. 3255-3261, 02/12/2007 2007, doi: 10.1021/bi0621967.
[107] K. E. Marshall *et al.*, "Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability," *Biochemistry,* vol. 50, no. 12, pp. 2061-71, Mar 29 2011, doi: 10.1021/bi101936c.
[108] S. M. Tracz, A. Abedini, M. Driscoll, and D. P. Raleigh, "Role of Aromatic Interactions in Amyloid Formation by Peptides Derived from Human Amylin," *Biochemistry,* vol. 43, no. 50, pp. 15901-15908, 11 Aug 2004 2004, doi: 10.1021/bi048812l.
[109] S. Maurer-Stroh *et al.*, "Exploring the sequence determinants of amyloid structure using position-specific scoring matrices," *Nature Methods,* vol. 7, no. 3, pp. 237-242, 2010, doi: 10.1038/nmeth.1432.
[110] A. M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, and L. Serrano, "Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins," *Nat Biotechnol,* vol. 22, no. 10, pp. 1302-6, Oct 2004, doi: 10.1038/nbt1012.

[111] A. Esteras-Chopo, L. Serrano, and M. Lopez de la Paz, "The amyloid stretch hypothesis: Recruiting proteins toward the dark site," *Proceedings of the National Academy of Sciences (PNAS),* vol. 102, no. 46, pp. 16672-16677, Nov 15 2005 2005, doi: 10.1073/pnas.0505905102.
[112] P. K. Teng and D. Eisenberg, "Short protein segments can drive a non-fibrillizing protein into the amyloid state," *Protein Engineering Design and Selection,* vol. 22, no. 8, pp. 531-536, 2009, doi: 10.1093/protein/gzp037.
[113] J. S. Beerten, J. ; Rousseau, F., "Aggregation Prone Regions and Gatekeeping Residues in Protein Sequences," *Current Topics in Medicinal Chemistry,* vol. 12, no. 22, pp. 2470-2478, 2012, doi: 10.2174/1568026611212220003.
[114] G. De Baets, J. Schymkowitz, and F. Rousseau, "Predicting aggregation-prone sequences in proteins," *Essays Biochem,* vol. 56, pp. 41-52, 2014, doi: 10.1042/bse0560041.
[115] C. M. Dobson, "Protein folding and misfolding," *Insight Review Articles,* vol. 426, 18 Dec 2003 2003, doi: 10.1038/nature02261.
[116] F. Chiti and C. M. Dobson, "Amyloid formation by globular proteins under native conditions," *Nat Chem Biol,* vol. 5, no. 1, pp. 15-22, Jan 2009, doi: 10.1038/nchembio.131.
[117] A. P. Pawar, K. F. Dubay, J. Zurdo, F. Chiti, M. Vendruscolo, and C. M. Dobson, "Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases," *J Mol Biol,* vol. 350, no. 2, pp. 379-92, Jul 8 2005, doi: 10.1016/j.jmb.2005.04.016.
[118] R. Linding, J. Schymkowitz, F. Rousseau, F. Diella, and L. Serrano, "A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins," *J Mol Biol,* vol. 342, no. 1, pp. 345-53, Sep 3 2004, doi: 10.1016/j.jmb.2004.06.088.
[119] D. Otzen, O. Kristensen, and M. Oliveberg, "Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: A structural clue to amyloid assembly," *Proceedings of the National Academy of Sciences (PNAS),* vol. 97, no. 18, pp. 9907-9912, 29 Aug 2000 2000, doi: 10.1073/pnas.160086297.
[120] F. Rousseau, L. Serrano, and J. W. Schymkowitz, "How evolutionary pressure against protein aggregation shaped chaperone specificity," *J Mol Biol,* vol. 355, no. 5, pp. 1037-47, Feb 3 2006, doi: 10.1016/j.jmb.2005.11.035.
[121] F. Rousseau, J. Schymkowitz, and L. Serrano, "Protein aggregation and amyloidosis: confusion of the kinds?," *Curr Opin Struct Biol,* vol. 16, no. 1, pp. 118-26, Feb 2006, doi: 10.1016/j.sbi.2006.01.011.
[122] M. C. Rodriguez and M. Cudic, "Optimization of physicochemical and pharmacological properties of peptide drugs by glycosylation," *Methods Mol Biol,* vol. 1081, pp. 107-36, 2013, doi: 10.1007/978-1-62703-652-8_8.
[123] P. L. Turecek, M. J. Bossard, F. Schoetens, and I. A. Ivens, "PEGylation of Biopharmaceuticals: A Review of Chemistry and Nonclinical Safety Information of Approved Drugs," *J Pharm Sci,* vol. 105, no. 2, pp. 460-475, Feb 2016, doi: 10.1016/j.xphs.2015.11.015.
[124] G. Lippens *et al.*, "Tau aggregation in Alzheimer's disease: what role for phosphorylation?," *Prion,* vol. 1, no. 1, pp. 21-5, Jan-Mar 2007, doi: 10.4161/pri.1.1.4055.
[125] N. M. Valette, S. E. Radford, S. A. Harris, and S. L. Warriner, "Phosphorylation as a tool to modulate aggregation propensity and to predict fibril architecture," *Chembiochem,* vol. 13, no. 2, pp. 271-81, Jan 23 2012, doi: 10.1002/cbic.201100607.

[126] N. McTiernan, I. Kjosas, and T. Arnesen, "Illuminating the impact of N-terminal acetylation: from protein to physiology," *Nat Commun,* vol. 16, no. 1, p. 703, Jan 15 2025, doi: 10.1038/s41467-025-55960-5.
[127] C. J. White and A. K. Yudin, "Contemporary strategies for peptide macrocyclization," *Nat Chem,* vol. 3, no. 7, pp. 509-24, Jun 23 2011, doi: 10.1038/nchem.1062.
[128] Y. H. Lau, P. de Andrade, Y. Wu, and D. R. Spring, "Peptide stapling techniques based on different macrocyclisation chemistries," *Chem Soc Rev,* vol. 44, no. 1, pp. 91-102, Jan 7 2015, doi: 10.1039/c4cs00246f.
[129] L. B. Zhang, G., "Converting Peptides into Drug Leads by Lipidation," *Current Medicinal Chemistry,* vol. 19, no. 11, pp. 1602-1618, 27/01/2012 2012, doi: 10.2174/092986712799945003.
[130] J. A. Hutchinson *et al.*, "The Effect of Lipidation on the Self-Assembly of the Gut-Derived Peptide Hormone PYY(3-36)," *Bioconjug Chem,* vol. 29, no. 7, pp. 2296-2308, Jul 18 2018, doi: 10.1021/acs.bioconjchem.8b00286.
[131] R. Kowalczyk, P. W. R. Harris, G. M. Williams, S.-H. Yang, and M. A. Brimble, "Peptide Lipidation – A Synthetic Strategy to Afford Peptide Based Therapeutics," in *Peptides and Peptide-based Biomaterials and their Biomedical Applications*, (Advances in Experimental Medicine and Biology, 2017, ch. Chapter 9, pp. 185-227.
[132] S. Ostergaard *et al.*, "The effect of fatty diacid acylation of human PYY(3-36) on Y(2) receptor potency and half-life in minipigs," *Sci Rep,* vol. 11, no. 1, p. 21179, Oct 27 2021, doi: 10.1038/s41598-021-00654-3.
[133] L. Maletinska, W. Neugebauer, J. Perodin, M. Lefebvre, and E. Escher, "Angiotensin Analogues Palmitoylated in Positions 1 and 4," *Journal of Medicinal Chemistry,* vol. 40, no. 20, pp. 3271-3279, 1997, doi: 10.1021/jm9608669.
[134] L. Yuan, J. Wang, and W. C. Shen, "Reversible lipidization prolongs the pharmacological effect, plasma duration, and liver retention of octreotide," *Pharm Res,* vol. 22, no. 2, pp. 220-7, Feb 2005, doi: 10.1007/s11095-004-1189-z.
[135] J. Wang, D. Chow, H. Heiati, and W. C. Shen, "Reversible lipidization for the oral delivery of salmon calcitonin," *J Control Release,* vol. 88, no. 3, pp. 369-80, Mar 26 2003, doi: 10.1016/s0168-3659(03)00008-7.
[136] A. Rawat and R. Nagaraj, "Covalently attached fatty acyl chains alter the aggregation behavior of an amyloidogenic peptide derived from human β2-microglobulin," *Journal of Peptide Science,* vol. 19, no. 12, pp. 770-783, 2013, doi: 10.1002/psc.2575.
[137] I. W. Hamley, "Lipopeptides: from self-assembly to bioactivity," *Chem Commun (Camb),* vol. 51, no. 41, pp. 8574-83, May 21 2015, doi: 10.1039/c5cc01535a.
[138] R. S. Tu and M. Tirrell, "Bottom-up design of biomimetic assemblies," *Adv Drug Deliv Rev,* vol. 56, no. 11, pp. 1537-63, Sep 22 2004, doi: 10.1016/j.addr.2003.10.047.
[139] B. P. Ward *et al.*, "Peptide lipidation stabilizes structure to enhance biological function," *Mol Metab,* vol. 2, no. 4, pp. 468-79, 2013, doi: 10.1016/j.molmet.2013.08.008.
[140] S. Kalra and R. Sahay, "A Review on Semaglutide: An Oral Glucagon-Like Peptide 1 Receptor Agonist in Management of Type 2 Diabetes Mellitus," *Diabetes Ther,* vol. 11, no. 9, pp. 1965-1982, Sep 2020, doi: 10.1007/s13300-020-00894-y.
[141] J. R. Kumita, C. J. Weston, L. P. Choo-Smith, G. A. Woolley, and O. S. Smart, "Prevention of peptide fibril formation in an aqueous environment by mutation of a single residue to Aib," *Biochemistry,* vol. 42, no. 15, pp. 4492-8, Apr 22 2003, doi: 10.1021/bi026856t.
[142] J. Chabenne *et al.*, "A glucagon analog chemically stabilized for immediate treatment of life-threatening hypoglycemia," *Mol Metab,* vol. 3, no. 3, pp. 293-300, Jun 2014, doi: 10.1016/j.molmet.2014.01.006.
[143] Y. Zheng, K. Mao, S. Chen, and H. Zhu, "Chirality Effects in Peptide Assembly Structures," *Front Bioeng Biotechnol,* vol. 9, p. 703004, 2021, doi: 10.3389/fbioe.2021.703004.

[144] Y. Zheng, L. Yu, Y. Zou, Y. Yang, and C. Wang, "Steric Dependence of Chirality Effect in Surface-Mediated Peptide Assemblies Identified with Scanning Tunneling Microscopy," *Nano Lett,* vol. 19, no. 8, pp. 5403-5409, Aug 14 2019, doi: 10.1021/acs.nanolett.9b01904.
[145] Z. Luo, S. Wang, and S. Zhang, "Fabrication of self-assembling D-form peptide nanofiber scaffold d-EAK16 for rapid hemostasis," *Biomaterials,* vol. 32, no. 8, pp. 2013-20, Mar 2011, doi: 10.1016/j.biomaterials.2010.11.049.
[146] M. Góngora-Benítez, J. Tulla-Puche, and F. Albericio, "Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics," *Chemical Reviews,* vol. 114, no. 2, pp. 901-926, 2013, doi: 10.1021/cr400031z.
[147] D. Hamada *et al.*, "Competition between folding, native-state dimerisation and amyloid aggregation in beta-lactoglobulin," *J Mol Biol,* vol. 386, no. 3, pp. 878-90, Feb 27 2009, doi: 10.1016/j.jmb.2008.12.038.
[148] S. Pechmann, E. Levy, G. G. Tartaglia, and M. Vandruscolo, "Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins," *Proceedings of the National Academy of Sciences (PNAS),* vol. 106, no. 25, pp. 10159-10164, 2009, doi: 10.1073 pnas.0812414106.
[149] G. G. Tartaglia, A. P. Pawar, S. Campioni, C. M. Dobson, F. Chiti, and M. Vendruscolo, "Prediction of aggregation-prone regions in structured proteins," *J Mol Biol,* vol. 380, no. 2, pp. 425-36, Jul 4 2008, doi: 10.1016/j.jmb.2008.05.013.
[150] M. Wilczynska, S. Lobov, P. I. Ohlsson, and T. Ny, "A redox-sensitive loop regulates plasminogen activator inhibitor type 2 (PAI-2) polymerization," *EMBO J,* vol. 22, no. 8, pp. 1753-61, Apr 15 2003, doi: 10.1093/emboj/cdg178.
[151] A. Maekawa, B. Schmidt, B. Fazekas de St Groth, Y. H. Sanejouand, and P. J. Hogg, "Evidence for a domain-swapped CD4 dimer as the coreceptor for binding to class II MHC," *J Immunol,* vol. 176, no. 11, pp. 6873-8, Jun 1 2006, doi: 10.4049/jimmunol.176.11.6873.
[152] T. Ganderton, J. W. Wong, C. Schroeder, and P. J. Hogg, "Lateral self-association of VWF involves the Cys2431-Cys2453 disulfide/dithiol in the C2 domain," *Blood,* vol. 118, no. 19, pp. 5312-8, Nov 10 2011, doi: 10.1182/blood-2011-06-360297.
[153] D. P. Hong, A. Ahmad, and A. L. Fink, "Fibrillation of human insulin A and B chains," *Biochemistry,* vol. 45, no. 30, pp. 9342-53, Aug 1 2006, doi: 10.1021/bi0604936.
[154] M. B. Borgia, A. A. Nickson, J. Clarke, and M. J. Hounslow, "A mechanistic model for amorphous protein aggregation of immunoglobulin-like domains," *J Am Chem Soc,* vol. 135, no. 17, pp. 6456-64, May 1 2013, doi: 10.1021/ja308852b.
[155] K. L. Zapadka, F. J. Becher, A. L. Gomes Dos Santos, and S. E. Jackson, "Factors affecting the physical stability (aggregation) of peptide therapeutics," *Interface Focus,* vol. 7, no. 6, p. 20170030, Dec 6 2017, doi: 10.1098/rsfs.2017.0030.
[156] L. K. Nielsen, R.; Coats, A.; Frojaer, S.; Brange, J.; Vyas, S.; Uversky, V.; Fink, A., "Effect of Environmental Factors on the Kinetics of Insulin Fibril Formation Elucidation of the Molecular Mechanism," *Biochemistry,* vol. 40, no. 20, pp. 6036-6046, March 20 2001 2001.
[157] G. Meisl *et al.*, "Molecular mechanisms of protein aggregation from global fitting of kinetic models," *Nat Protoc,* vol. 11, no. 2, pp. 252-72, Feb 2016, doi: 10.1038/nprot.2016.010.
[158] K. L. Zapadka *et al.*, "A pH-Induced Switch in Human Glucagon-like Peptide-1 Aggregation Kinetics," *J Am Chem Soc,* vol. 138, no. 50, pp. 16259-16265, Dec 21 2016, doi: 10.1021/jacs.6b05025.
[159] S. Yun, B. Urbanc, L. Cruz, G. Bitan, D. B. Teplow, and H. E. Stanley, "Role of electrostatic interactions in amyloid beta-protein (A beta) oligomer formation: a

discrete molecular dynamics study," *Biophys J,* vol. 92, no. 11, pp. 4064-77, Jun 1 2007, doi: 10.1529/biophysj.106.097766.
[160] B. Raman *et al.*, "Critical Balance of Electrostatic and Hydrophobic Interactions Is Required for β2-Microglobulin Amyloid Fibril Growth and Stability," *Biochemistry,* vol. 44, no. 4, pp. 1288-1299, 1 Jul 2005 2005, doi: 10.1021/bi048029t.
[161] M. Kamihira, Y. Oshiro, S. Tuzi, A. Y. Nosaka, H. Saito, and A. Naito, "Effect of electrostatic interaction on fibril formation of human calcitonin as studied by high resolution solid state 13C NMR," *J Biol Chem,* vol. 278, no. 5, pp. 2859-65, Jan 31 2003, doi: 10.1074/jbc.M205285200.
[162] M. R. Krebs, K. R. Domike, and A. M. Donald, "Protein aggregation: more than just fibrils," *Biochem Soc Trans,* vol. 37, no. Pt 4, pp. 682-6, Aug 2009, doi: 10.1042/BST0370682.
[163] W. Wang and C. J. Roberts, "Non-Arrhenius protein aggregation," *AAPS J,* vol. 15, no. 3, pp. 840-51, Jul 2013, doi: 10.1208/s12248-013-9485-3.
[164] J. M. Andrews and C. J. Roberts, "Non-Native Aggregation of α-Chymotrypsinogen Occurs through Nucleation and Growth with Competing Nucleus Sizes and Negative Activation Energies," *Biochemistry,* vol. 46, no. 25, pp. 7558-7571, May 27 2007 2007, doi: doi.org/10.1021/bi700296f.
[165] R. Carrotta, M. Manno, D. Bulone, V. Martorana, and P. L. San Biagio, "Protofibril formation of amyloid beta-protein at low pH via a non-cooperative elongation mechanism," *J Biol Chem,* vol. 280, no. 34, pp. 30001-8, Aug 26 2005, doi: 10.1074/jbc.M500052200.
[166] Y. Kusumoto, A. Lomakin, D. B. Teplow, and G. B. Benedek, "Temperature dependence of amyloid β-protein fibrillization," *Proceedings of the National Academy of Sciences (PNAS),* vol. 95, no. 21, pp. 12277-12282, Oct 13 1998 1998, doi: 10.1073/pnas.95.21.12277.
[167] R. Sabate, M. Gallardo, and J. Estelrich, "Temperature dependence of the nucleation constant rate in beta amyloid fibrillogenesis," *Int J Biol Macromol,* vol. 35, no. 1-2, pp. 9-13, Mar 2005, doi: 10.1016/j.ijbiomac.2004.11.001.
[168] M. I. Smith, J. S. Sharp, and C. J. Roberts, "Nucleation and growth of insulin fibrils in bulk solution and at hydrophobic polystyrene surfaces," *Biophys J,* vol. 93, no. 6, pp. 2143-51, Sep 15 2007, doi: 10.1529/biophysj.107.105338.
[169] M. Manno *et al.*, "Kinetics of different processes in human insulin amyloid formation," *J Mol Biol,* vol. 366, no. 1, pp. 258-74, Feb 9 2007, doi: 10.1016/j.jmb.2006.11.008.
[170] M. Riès-kautt and A. Ducruix, "Inferences drawn from physicochemical studies of crystallogenesis and precrystalline state," *Methods in Enzymology,* vol. 276, pp. 23-59, 1997, doi: doi.org/10.1016/s0076-6879(97)76049-x.
[171] K. Klement, K. Wieligmann, J. Meinhardt, P. Hortschansky, W. Richter, and M. Fandrich, "Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer's abeta(1-40) amyloid fibrils," *J Mol Biol,* vol. 373, no. 5, pp. 1321-33, Nov 9 2007, doi: 10.1016/j.jmb.2007.08.068.
[172] D. T. Gjerde, J. S. Fritz, and G. Schmuckler, "Anion chromatography with low-conductivity eluents," *Journal of Chromatography A,* vol. 186, pp. 509-519, 30 Dec 11979 1979, doi: 10.1016/S0021-9673(00)95271-3.
[173] R. L. Baldwin, "How Hofmeister ion interactions affect protein stability," *Biophysical Journal,* vol. 71, no. 4, pp. 2056-2063, 1996, doi: 10.1016/S0006-3495(96)79404-3.
[174] W. Kunz, J. Henle, and B. W. Ninham, "'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers," *Current Opinion in Colloid & Interface Science,* vol. 9, no. 1-2, pp. 19-37, 2004, doi: 10.1016/j.cocis.2004.05.005.

[175] L. A. Munishkina, J. Henriques, V. N. Uversky, and A. L. Fink, "Role of Protein–Water Interactions and Electrostatics in α-Synuclein Fibril Formation," *Biochemistry,* vol. 43, no. 11, pp. 3289-3300, 27 Feb 2004 2004, doi: 10.1021/bi034938r.
[176] P. J. Marek, V. Patsalo, D. F. Green, and D. P. Raleigh, "Ionic strength effects on amyloid formation by amylin are a complicated interplay among Debye screening, ion selectivity, and Hofmeister effects," *Biochemistry,* vol. 51, no. 43, pp. 8478-90, Oct 30 2012, doi: 10.1021/bi300574r.
[177] A. Abelein, J. Jarvet, A. Barth, A. Graslund, and J. Danielsson, "Ionic Strength Modulation of the Free Energy Landscape of Abeta40 Peptide Fibril Formation," *J Am Chem Soc,* vol. 138, no. 21, pp. 6893-902, Jun 1 2016, doi: 10.1021/jacs.6b04511.
[178] H. J. Lee, A. McAuley, K. F. Schilke, and J. McGuire, "Molecular origins of surfactant-mediated stabilization of protein drugs," *Adv Drug Deliv Rev,* vol. 63, no. 13, pp. 1160-71, Oct 2011, doi: 10.1016/j.addr.2011.06.015.
[179] T. A. Khan, H. C. Mahler, and R. S. Kishore, "Key interactions of surfactants in therapeutic protein formulations: A review," *Eur J Pharm Biopharm,* vol. 97, no. Pt A, pp. 60-7, Nov 2015, doi: 10.1016/j.ejpb.2015.09.016.
[180] D. Otzen, "Protein-surfactant interactions: a tale of many states," *Biochim Biophys Acta,* vol. 1814, no. 5, pp. 562-91, May 2011, doi: 10.1016/j.bbapap.2011.03.003.
[181] R. S. Kishore *et al.*, "Degradation of polysorbates 20 and 80: studies on thermal autoxidation and hydrolysis," *J Pharm Sci,* vol. 100, no. 2, pp. 721-31, Feb 2011, doi: 10.1002/jps.22290.
[182] R. S. Kishore, S. Kiese, S. Fischer, A. Pappenberger, U. Grauschopf, and H. C. Mahler, "The degradation of polysorbates 20 and 80 and its potential impact on the stability of biotherapeutics," *Pharm Res,* vol. 28, no. 5, pp. 1194-210, May 2011, doi: 10.1007/s11095-011-0385-x.
[183] P. H. Yancey, M. E. Clark, S. C. Hand, D. Bowlus, and G. N. Somero, "Living with Water Stress: Evolution of Osmolyte Systems," *Science,* vol. 217, no. 4556, pp. 1214-1222, Sep. 24 1982 1982, doi: 10.1126/science.7112124.
[184] F. Macchi, M. Eisenkolb, H. Kiefer, and D. E. Otzen, "The Effect of Osmolytes on Protein Fibrillation," *International Journal of Molecular Sciences,* vol. 13, no. 3, pp. 3801-3819, 2012, doi: 10.3390/ijms13033801.
[185] D. W. Bolen and G. D. Rose, "Structure and energetics of the hydrogen-bonded backbone in protein folding," *Annu Rev Biochem,* vol. 77, pp. 339-62, 2008, doi: 10.1146/annurev.biochem.77.061306.131357.
[186] W. Qi, A. Zhang, T. A. Good, and E. J. Fernandez, "Two disaccharides and trimethylamine N-oxide affect Abeta aggregation differently, but all attenuate oligomer-induced membrane permeability," *Biochemistry,* vol. 48, no. 37, pp. 8908-19, Sep 22 2009, doi: 10.1021/bi9006397.
[187] Y. Maeda, H. Koga, H. Yamada, T. Ueda, and T. Imoto, "Effective renaturation of reduced lysozyme by gentle removal of urea," *Protein Engineering,* vol. 8, no. 2, pp. 201-205, 1995, doi: 10.1093/protein/8.2.201.
[188] R. Politi and D. Harries, "Enthalpically driven peptide stabilization by protective osmolytes," *Chem Commun (Camb),* vol. 46, no. 35, pp. 6449-51, Sep 21 2010, doi: 10.1039/c0cc01763a.
[189] D. E. Otzen, *Amyloid Fibrils and Prefibrillar Aggregates*. 2013.
[190] S. M. D'Addio *et al.*, "New and Evolving Techniques for the Characterization of Peptide Therapeutics," *J Pharm Sci,* vol. 105, no. 10, pp. 2989-3006, Oct 2016, doi: 10.1016/j.xphs.2016.06.011.
[191] M. F. Pignataro, M. G. Herrera, and V. I. Dodero, "Evaluation of Peptide/Protein Self-Assembly and Aggregation by Spectroscopic Methods," *Molecules,* vol. 25, no. 20, Oct 21 2020, doi: 10.3390/molecules25204854.

[192] Y. Wang, A. Lomakin, S. Kanai, R. Alex, and G. B. Benedek, "Transformation of oligomers of lipidated peptide induced by change in pH," *Mol Pharm,* vol. 12, no. 2, pp. 411-9, Feb 2 2015, doi: 10.1021/mp500519s.
[193] J. R. Bothe, A. Andrews, K. J. Smith, L. A. Joyce, Y. Krishnamachari, and S. Kashi, "Peptide Oligomerization Memory Effects and Their Impact on the Physical Stability of the GLP-1 Agonist Liraglutide," *Molecular Pharmaceutics,* vol. 16, no. 5, pp. 2153-2161, 2019, doi: 10.1021/acs.molpharmaceut.9b00106.
[194] T. M. Frederiksen *et al.*, "Oligomerization of a Glucagon-like Peptide 1 Analog: Bridging Experiment and Simulations," *Biophys J,* vol. 109, no. 6, pp. 1202-13, Sep 15 2015, doi: 10.1016/j.bpj.2015.07.051.
[195] Q. Li, V. Tangry, D. P. Allen, K. D. Seibert, K. K. Qian, and N. J. Wagner, "Surface-mediated spontaneous emulsification of the acylated peptide, semaglutide," *Proc Natl Acad Sci U S A,* vol. 121, no. 5, p. e2305770121, Jan 30 2024, doi: 10.1073/pnas.2305770121.
[196] M. Giannetti, A. Palleschi, B. Ricciardi, and M. Venanzi, "A Spectroscopic and Molecular Dynamics Study on the Aggregation Properties of a Lipopeptide Analogue of Liraglutide, a Therapeutic Peptide against Diabetes Type 2," *Molecules,* vol. 28, no. 22, Nov 11 2023, doi: 10.3390/molecules28227536.
[197] S. Dolui *et al.*, "Raman Spectroscopic Insights of Phase-Separated Insulin Aggregates," *ACS Phys Chem Au,* vol. 4, no. 3, pp. 268-280, May 22 2024, doi: 10.1021/acsphyschemau.3c00065.
[198] T. K. Karamanos, A. P. Kalverda, G. S. Thompson, and S. E. Radford, "Mechanisms of amyloid formation revealed by solution NMR," *Prog Nucl Magn Reson Spectrosc,* vol. 88-89, pp. 86-104, Aug 2015, doi: 10.1016/j.pnmrs.2015.05.002.
[199] J. Stetefeld, S. A. McKenna, and T. R. Patel, "Dynamic light scattering: a practical guide and applications in biomedical sciences," *Biophys Rev,* vol. 8, no. 4, pp. 409-427, Dec 2016, doi: 10.1007/s12551-016-0218-6.
[200] A. Lomakin, D. B. Teplow, and G. B. Benedek, "Quasielastic Light Scattering for Protein Assembly Studies," in *Amyloid Proteins*, vol. 299, E. M. Sigurdsson Ed., (Methods in Molecular Biology: Humana Press, 2005, ch. 10, pp. 153-174.
[201] B. Lorber, F. Fischer, M. Bailly, H. Roy, and D. Kern, "Protein analysis by dynamic light scattering: methods and techniques for students," *Biochem Mol Biol Educ,* vol. 40, no. 6, pp. 372-82, Nov-Dec 2012, doi: 10.1002/bmb.20644.
[202] P. N. Pusey, "Macromolecular Diffusion," *Photon Correlation and Light Beating Spectroscopy,* no. 3, pp. 387-428, 1974, doi: 10.1007/978-1-4615-8906-8_10.
[203] H.-C. Mahler, W. Friess, U. Grauschopf, and S. Kiese, "Protein aggregation: Pathways, induction factors and analysis," *Journal of Pharmaceutical Sciences,* vol. 98, no. 9, pp. 2909-2934, 2009, doi: 10.1002/jps.21566.
[204] V. M. Gun'ko, A. V. Klyueva, Y. N. Levchuk, and R. Leboda, "Photon correlation spectroscopy investigations of proteins," *Adv Colloid Interface Sci,* vol. 105, pp. 201-328, Sep 18 2003, doi: 10.1016/s0001-8686(03)00091-5.
[205] F. Varenne, J. Botton, C. Merlet, M. Beck-Broichsitter, F.-X. Legrand, and C. Vauthier, "Standardization and validation of a protocol of size measurements by dynamic light scattering for monodispersed stable nanomaterial characterization," *Colloids and Surfaces A: Physicochemical and Engineering Aspects,* vol. 486, pp. 124-138, 2015, doi: 10.1016/j.colsurfa.2015.08.043.
[206] S. Fekete, A. Beck, J. L. Veuthey, and D. Guillarme, "Theory and practice of size exclusion chromatography for the analysis of protein aggregates," *J Pharm Biomed Anal,* vol. 101, pp. 161-73, Dec 2014, doi: 10.1016/j.jpba.2014.04.011.

[207] P. Hong, S. Koza, and E. S. Bouvier, "Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates," *J Liq Chromatogr Relat Technol,* vol. 35, no. 20, pp. 2923-2950, Nov 2012, doi: 10.1080/10826076.2012.743724.
[208] T. Arakawa, D. Ejima, T. Li, and J. S. Philo, "The critical role of mobile phase composition in size exclusion chromatography of protein pharmaceuticals," *J Pharm Sci,* vol. 99, no. 4, pp. 1674-92, Apr 2010, doi: 10.1002/jps.21974.
[209] Y. Liu, W. Radke, and H. Pasch, "Coil-Stretch Transition of High Molar Mass Polymers in Packed-Column Hydrodynamic Chromatography," *Macromolecules,* vol. 38, no. 17, pp. 7476-7484, 2005, doi: https://doi.org/10.1021/ma050964g.
[210] D. Badgujar, S. Bawake, A. Chawathe, and N. Sharma, "Size-exclusion LC-UV/HRMS based method for the analysis of aggregates in synthetic GLP-1 analog liraglutide and evaluation of excipient impact on aggregation," *Biomedical Chromatography,* vol. 38, no. 10, 2024, doi: 10.1002/bmc.5983.
[211] E. Brichtova, "Studies on the physical stability of a C-terminally amidated variant of GLP-1," Master of Philosophy, Chemistry, University of Cambridge, Cambridge, 2019.
[212] P. M. Tessier, A. M. Lenhoff, and S. I. Sandler, "Rapid Measurement of Protein Osmotic Second Virial Coefficients by Self-Interaction Chromatography," *Biophys J,* vol. 82, no. 3, pp. 1620-1631, 2002, doi: 10.1016/S0006-3495(02)75513-6.
[213] X. Hu, S. Spada, S. White, S. Hudson, E. Magner, and J. G. Wall, "Adsorption and activity of a domoic acid binding antibody fragment on mesoporous silicates," *J Phys Chem B,* vol. 110, no. 37, pp. 18703-9, Sep 21 2006, doi: 10.1021/jp062423e.
[214] P. DePhillips and A. M. Lenhoff, "Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography," *J Chromatogr A,* vol. 883, no. 1-2, pp. 39-54, Jun 23 2000, doi: 10.1016/s0021-9673(00)00420-9.
[215] C. Haas, J. Drenth, and W. W. Wilson, "Relation between the Solubility of Proteins in Aqueous Solutions and the Second Virial Coefficient of the Solution," *Journal of Physical Chemistry B,* vol. 103, no. 14, pp. 2808-2811, Mar 20 1999 1999, doi: 10.1021/jp984035l.
[216] S. Ruppert, S. I. Sandler, and A. M. Lenhoff, "Correlation between the Osmotic Second Virial Coefficient and the Solubility of Proteins," *Biotechnology Process,* vol. 182, no. 1, 2001, doi: 10.1021/bp0001314.
[217] W. W. Wilson and L. J. Delucas, "Applications of the second virial coefficient: protein crystallization and solubility," *Acta Crystallogr F Struct Biol Commun,* vol. 70, no. Pt 5, pp. 543-54, May 2014, doi: 10.1107/S2053230X1400867X.
[218] R. W. Payne *et al.*, "Second virial coefficient determination of a therapeutic peptide by self-interaction chromatography," *Biopolymers,* vol. 84, no. 5, pp. 527-33, 2006, doi: 10.1002/bip.20554.
[219] S. H. M. Hedberg, D. Lee, Y. Mishra, J. M. Haigh, and D. R. Williams, "Mapping the mAb Aggregation Propensity Using Self-Interaction Chromatography as a Screening Tool," *Analytical Chemistry,* vol. 90, no. 6, pp. 3878-3885, 15 Feb 2018 2018, doi: 10.1021/acs.analchem.7b04605.
[220] J. Drewe, C. Beglinger, and T. Kissel, "The absorption site of cyclosporin in the human gastrointestinal tract," *Br J Clin Pharmacol,* vol. 33, no. 1, pp. 39-43, Jan 1992, doi: 10.1111/j.1365-2125.1992.tb03998.x.
[221] A. L. Lewis, N. McEntee, J. Holland, and A. Patel, "Development and approval of rybelsus (oral semaglutide): ushering in a new era in peptide delivery," *Drug Deliv Transl Res,* vol. 12, no. 1, pp. 1-6, Jan 2022, doi: 10.1007/s13346-021-01000-w.
[222] A. Fuchs and J. B. Dressman, "Composition and physicochemical properties of fasted-state human duodenal and jejunal fluid: a critical evaluation of the available data," *J Pharm Sci,* vol. 103, no. 11, pp. 3398-3411, Nov 2014, doi: 10.1002/jps.24183.

[223] P. J. Lu *et al.*, "Gastric juice acidity in upper gastrointestinal diseases," *World J Gastroenterol,* vol. 16, no. 43, pp. 5496-501, Nov 21 2010, doi: 10.3748/wjg.v16.i43.5496.
[224] J. Fallingborg, L. A. Christensen, B. A. Jacobsen, and S. N. Rasmussen, "Very low intraluminal colonic pH in patients with active ulcerative colitis," *Dig Dis Sci,* vol. 38, no. 11, pp. 1989-93, Nov 1993, doi: 10.1007/BF01297074.
[225] Y. Sasaki, R. Hada, H. Nakajima, S. Fukuda, and A. Munakata, "Improved localizing method of radiopill in measurement of entire gastrointestinal pH profiles: colonic luminal pH in normal subjects and patients with Crohn's disease," *The American Journal of Gastroenterology,* vol. 92, no. 1, pp. 114-118, 1997.
[226] N. Rouge, P. Buri, and E. Doelker, "Drug absorption sites in the gastrointestinal tract and dosage forms for site-specific delivery," *International Journal of Pharmaceutics,* vol. 136, no. 1-2, pp. 117-139, 1996, doi: 10.1016/0378-5173(96)85200-8.
[227] M. Koziolek *et al.*, "Navigating the human gastrointestinal tract for oral drug delivery: Uncharted waters and new frontiers," *Adv Drug Deliv Rev,* vol. 101, pp. 75-88, Jun 1 2016, doi: 10.1016/j.addr.2016.03.009.
[228] M. A. Kwiatek *et al.*, "Effect of meal volume and calorie load on postprandial gastric function and emptying: studies under physiological conditions by combined fiber-optic pressure measurement and MRI," *Am J Physiol Gastrointest Liver Physiol,* vol. 297, no. 5, pp. G894-901, Nov 2009, doi: 10.1152/ajpgi.00117.2009.
[229] A. Y. Abuhelwa, D. J. R. Foster, and R. N. Upton, "A Quantitative Review and Meta-Models of the Variability and Factors Affecting Oral Drug Absorption-Part I: Gastrointestinal pH," *AAPS J,* vol. 18, no. 5, pp. 1309-1321, Sep 2016, doi: 10.1208/s12248-016-9952-8.
[230] A. Y. Abuhelwa, D. J. R. Foster, and R. N. Upton, "A Quantitative Review and Meta-models of the Variability and Factors Affecting Oral Drug Absorption-Part II: Gastrointestinal Transit Time," *AAPS J,* vol. 18, no. 5, pp. 1322-1333, Sep 2016, doi: 10.1208/s12248-016-9953-7.
[231] R. Bennink *et al.*, "Evaluation of small-bowel transit for solid and liquid test meal in healthy men and women," *Eur J Nucl Med,* vol. 26, no. 12, pp. 1560-6, Dec 1999, doi: 10.1007/s002590050495.
[232] C. Litou *et al.*, "Characteristics of the Human Upper Gastrointestinal Contents in the Fasted State Under Hypo- and A-chlorhydric Gastric Conditions Under Conditions of Typical Drug - Drug Interaction Studies," *Pharm Res,* vol. 33, no. 6, pp. 1399-412, Jun 2016, doi: 10.1007/s11095-016-1882-8.
[233] L. Kalantzi, K. Goumas, V. Kalioras, B. Abrahamsson, J. B. Dressman, and C. Reppas, "Characterization of the human upper gastrointestinal contents under conditions simulating bioavailability/bioequivalence studies," *Pharm Res,* vol. 23, no. 1, pp. 165-76, Jan 2006, doi: 10.1007/s11095-005-8476-1.
[234] S. Kong, Y. H. Zhang, and W. Zhang, "Regulation of Intestinal Epithelial Cells Properties and Functions by Amino Acids," *Biomed Res Int,* vol. 2018, p. 2819154, 2018, doi: 10.1155/2018/2819154.
[235] D. F. Evans, G. Pye, R. Bramley, A. G. Clark, T. J. Dyson, and J. D. Hardcastle, "Measurement of gastrointestinal pH profiles in normal ambulant human subjects," *Gut,* vol. 29, no. 8, pp. 1035-41, Aug 1988, doi: 10.1136/gut.29.8.1035.
[236] D. Riethorst, R. Mols, G. Duchateau, J. Tack, J. Brouwers, and P. Augustijns, "Characterization of Human Duodenal Fluids in Fasted and Fed State Conditions," *J Pharm Sci,* vol. 105, no. 2, pp. 673-681, Feb 2016, doi: 10.1002/jps.24603.
[237] J. B. Dressman, M. Vertzoni, K. Goumas, and C. Reppas, "Estimating drug solubility in the gastrointestinal tract," *Adv Drug Deliv Rev,* vol. 59, no. 7, pp. 591-602, Jul 30 2007, doi: 10.1016/j.addr.2007.05.009.

[238] A. Fuchs, M. Leigh, B. Kloefer, and J. B. Dressman, "Advances in the design of fasted state simulating intestinal fluids: FaSSIF-V3," *Eur J Pharm Biopharm,* vol. 94, pp. 229-40, Aug 2015, doi: 10.1016/j.ejpb.2015.05.015.
[239] G. Ismailos, C. Reppas, J. B. Dressman, and P. Macheras, "Unusual solubility behaviour of cyclosporin A in aqueous media," *J Pharm Pharmacol,* vol. 43, no. 4, pp. 287-9, Apr 1991, doi: 10.1111/j.2042-7158.1991.tb06688.x.
[240] C. Markopoulos, C. J. Andreas, M. Vertzoni, J. Dressman, and C. Reppas, "In-vitro simulation of luminal conditions for evaluation of performance of oral drug products: Choosing the appropriate test media," *Eur J Pharm Biopharm,* vol. 93, pp. 173-82, Jun 2015, doi: 10.1016/j.ejpb.2015.03.009.
[241] E. Jantratid, N. Janssen, C. Reppas, and J. B. Dressman, "Dissolution media simulating conditions in the proximal human gastrointestinal tract: an update," *Pharm Res,* vol. 25, no. 7, pp. 1663-76, Jul 2008, doi: 10.1007/s11095-008-9569-4.
[242] T. J. Dening, J. T. Douglas, and M. J. Hageman, "Do Macrocyclic Peptide Drugs Interact with Bile Salts under Simulated Gastrointestinal Conditions?," *Mol Pharm,* vol. 18, no. 8, pp. 3086-3098, Aug 2 2021, doi: 10.1021/acs.molpharmaceut.1c00309.
[243] P. Langguth *et al.*, "The challenge of proteolytic enzymes in intestinal peptide delivery," *Journal of Controlled Release,* vol. 46, no. 1-2, pp. 39-57, 1997, doi: 10.1016/s0168-3659(96)01586-6.
[244] A. Bernkop-Schnurch, "The use of inhibitory agents to overcome the enzymatic barrier to perorally administered therapeutic peptides and proteins," *J Control Release,* vol. 52, no. 1-2, pp. 1-16, Mar 2 1998, doi: 10.1016/s0168-3659(97)00204-6.
[245] O. G. Famutimi, V. G. Adebiyi, B. G. Akinmolu, O. V. Dada, and I. O. Adewale, "Trypsin, chymotrypsin and elastase in health and disease," *Future Journal of Pharmaceutical Sciences,* vol. 10, no. 1, 2024, doi: 10.1186/s43094-024-00709-y.
[246] D. C. Whitcomb and M. E. Lowe, "Human pancreatic digestive enzymes," *Dig Dis Sci,* vol. 52, no. 1, pp. 1-17, Jan 2007, doi: 10.1007/s10620-006-9589-z.
[247] J. Woodley, "Enzymatic barriers for GI peptide and protein delivery," *Critical Reviews in Therapeutic Drug Carrier Systems,* vol. 11, no. 2-3, pp. 61-95, 1994.
[248] L. Michaelis and M. L. Menten, "Die kinetik der invertinwirkung," *Biochem. z,* vol. 49, no. 333-369, p. 352, 1913.
[249] H. Bisswanger, *Enzyme Kinetics*. 2017.
[250] K. Buccholz, V. Kasche, and U. T. Bornscheuer, *Biocatalysts and Enzyme Technology*, 2 ed. Wiley-Blackwell, 2012, p. 606.
[251] R. A. Copeland, "Reversible Modes of Inhibitor Interactions with Enzymes," in *Evaluation of Enzyme Inhibitors in Drug Discovery*, 2013, pp. 57-121.
[252] M. Werle and F. Foger, "Peroral peptide delivery: Peptidase inhibition as a key concept for commercial drug products," *Bioorg Med Chem,* vol. 26, no. 10, pp. 2906-2913, Jun 1 2018, doi: 10.1016/j.bmc.2017.08.028.
[253] J. Woodley, "Enzymatic Barriers," in *Oral Delivery of Macromolecular Drugs*, 2009, ch. Chapter 1, pp. 1-19.
[254] R. Eldor *et al.*, "Oral insulin (ORMD-0801) in type 2 diabetes mellitus: A dose-finding 12-week randomized placebo-controlled study," *Diabetes Obes Metab,* vol. 25, no. 4, pp. 943-952, Apr 2023, doi: 10.1111/dom.14901.
[255] M. F. Lee and C. L. Poh, "Strategies to improve the physicochemical properties of peptide-based drugs," *Pharm Res,* vol. 40, no. 3, pp. 617-632, Mar 2023, doi: 10.1007/s11095-023-03486-0.
[256] Y. Markert, J. Koditz, R. Ulbrich-Hofmann, and U. Arnold, "Proline versus charge concept for protein stabilization against proteolytic attack," *Protein Eng,* vol. 16, no. 12, pp. 1041-6, Dec 2003, doi: 10.1093/protein/gzg136.

[257] Y. Zhao *et al.*, "Antimicrobial activity and stability of the D-amino acid substituted derivatives of antimicrobial peptide polybia-MPI," *AMB Express,* vol. 6, no. 1, p. 122, Dec 2016, doi: 10.1186/s13568-016-0295-8.
[258] K. Tarchoun, M. a. Yousef, and Z. Bánóczi, "Azapeptides as an Efficient Tool to Improve the Activity of Biologically Effective Peptides," *Future Pharmacology,* vol. 2, no. 3, pp. 293-305, 2022, doi: 10.3390/futurepharmacol2030020.
[259] S. Hess *et al.*, "Effect of structural and conformation modifications, including backbone cyclization, of hydrophilic hexapeptides on their intestinal permeability and enzymatic stability," *J Med Chem,* vol. 50, no. 24, pp. 6201-11, Nov 29 2007, doi: 10.1021/jm070836d.
[260] S. Muranishi, M. Murakami, M. Hashidzume, K. Yamada, S. Tajima, and Y. Kiso, "Trials of lipid modification of peptide hormones for intestinal delivery," *Journal of Controlled Release,* vol. 19, no. 1-3, pp. 179-188, 1992, doi: 10.1016/0168-3659(92)90075-3.
[261] E. Yodoya *et al.*, "Enhanced permeability of tetragastrin across the rat intestinal membrane and its reduced degradation by acylation with various fatty acids," *The Journal of Pharmacology and Experimental Therapeutics,* vol. 271, no. 3, pp. 1509-1513, 1994, doi: 10.1016/s0022-3565(25)24037-x.
[262] R. A. Copeland, "Experimental Measures of Steady-State Enzyme Activity," in *Enzymes A Practical Introduction to Structure, Mechanism and Data Analysis*, 3rd ed.: John Wiley & Sons, Inc., 2023, ch. 7, pp. 193-260.
[263] H. Bisswanger, "Enzyme assays," *Perspectives in Science,* vol. 1, no. 1-6, pp. 41-55, 2014, doi: 10.1016/j.pisc.2014.02.005.
[264] M. Wu, E. Petryayeva, I. L. Medintz, and W. R. Algar, "Quantitative measurement of proteolytic rates with quantum dot-peptide substrate conjugates and Forster resonance energy transfer," *Methods Mol Biol,* vol. 1199, pp. 215-39, 2014, doi: 10.1007/978-1-4939-1280-3_17.
[265] P. B. Noto *et al.*, "Alternative stabilities of a proline-rich antibacterial peptide in vitro and in vivo," *Protein Sci,* vol. 17, no. 7, pp. 1249-55, Jul 2008, doi: 10.1110/ps.034330.108.
[266] X. D. Kong *et al.*, "De novo development of proteolytically resistant therapeutic peptides for oral administration," *Nat Biomed Eng,* vol. 4, no. 5, pp. 560-571, May 2020, doi: 10.1038/s41551-020-0556-3.
[267] J. M. Bolivar, J. M. Woodley, and R. Fernandez-Lafuente, "Is enzyme immobilization a mature discipline? Some critical considerations to capitalize on the benefits of immobilization," *Chem Soc Rev,* vol. 51, no. 15, pp. 6251-6290, Aug 1 2022, doi: 10.1039/d2cs00083k.
[268] N. A. Mohidem, M. Mohamad, M. U. Rashid, M. N. Norizan, F. Hamzah, and H. b. Mat, "Recent Advances in Enzyme Immobilisation Strategies: An Overview of Techniques and Composite Carriers," *Journal of Composites Science,* vol. 7, no. 12, 2023, doi: 10.3390/jcs7120488.
[269] L. Qiao, Y. Liu, S. P. Hudson, P. Yang, E. Magner, and B. Liu, "A nanoporous reactor for efficient proteolysis," *Chemistry,* vol. 14, no. 1, pp. 151-7, 2008, doi: 10.1002/chem.200701102.
[270] C. Mateo, J. M. Palomo, G. Fernandez-Lorente, J. M. Guisan, and R. Fernandez-Lafuente, "Improvement of enzyme activity, stability and selectivity via immobilization techniques," *Enzyme and Microbial Technology,* vol. 40, no. 6, pp. 1451-1463, 2007, doi: 10.1016/j.enzmictec.2007.01.018.
[271] Y. Jiang *et al.*, "Pros and Cons in Various Immobilization Techniques and Carriers for Enzymes," *Appl Biochem Biotechnol,* vol. 196, no. 9, pp. 5633-5655, Sep 2024, doi: 10.1007/s12010-023-04838-7.
[272] Microsoft CoPilot. ""Can you create for me a diagram showing the different immobilised enzyme types (i.e. physisorbed to a surface, covalently bound to a surface, entrapped

and cross-linked enzyme aggregates) in the format as if for a scientific journal so plenty of detail but no need for labels. Just try to make sure it has a biorender feel to it so it wouldn't look out of place in a scientific journal"." (accessed 7th July, 2025).
[273] T. Jesionowski, J. Zdarta, and B. Krajewska, "Enzyme immobilization by adsorption: a review," *Adsorption,* vol. 20, no. 5-6, pp. 801-821, 2014, doi: 10.1007/s10450-014-9623-y.
[274] E. T. Hwang and M. B. Gu, "Enzyme stabilization by nano/microsized hybrid materials," *Engineering in Life Sciences,* vol. 13, no. 1, pp. 49-61, 2012, doi: 10.1002/elsc.201100225.
[275] L. Thudi *et al.*, "Enzyme immobilization on epoxy supports in reverse micellar media: Prevention of enzyme denaturation," *Journal of Molecular Catalysis B: Enzymatic,* vol. 74, no. 1-2, pp. 54-62, 2012, doi: 10.1016/j.molcatb.2011.08.014.
[276] S. Hudson, J. Cooney, and E. Magner, "Proteins in mesoporous silicates," *Angew Chem Int Ed Engl,* vol. 47, no. 45, pp. 8582-94, 2008, doi: 10.1002/anie.200705238.
[277] S. Hudson, J. Cooney, B. K. Hodnett, and E. Magner, "Chloroperoxidase on Periodic Mesoporous Organosilanes: Immobilization and Reuse," *Chemistry of Materials,* vol. 19, no. 8, pp. 2049-2055, 2007, doi: 10.1021/cm070180c.
[278] A. Mohamed Nageib, A. Tariq Jameel, and F. B. Ahmad, "Immobilization of alpha-naphthyl acetate esterase (ANAE) on k-carrageenan for potential biosensor application in pesticide detection," *Materials Today: Proceedings,* vol. 57, pp. 1296-1300, 2022, doi: 10.1016/j.matpr.2021.11.601.
[279] E. Steen Redeker, D. T. Ta, D. Cortens, B. Billen, W. Guedens, and P. Adriaensens, "Protein engineering for directed immobilization," *Bioconjug Chem,* vol. 24, no. 11, pp. 1761-77, Nov 20 2013, doi: 10.1021/bc4002823.
[280] M. S. Robescu and T. Bavaro, "A Comprehensive Guide to Enzyme Immobilization: All You Need to Know," *Molecules,* vol. 30, no. 4, Feb 18 2025, doi: 10.3390/molecules30040939.
[281] O. Barbosa, C. Ortiz, Á. Berenguer-Murcia, R. Torres, R. C. Rodrigues, and R. Fernandez-Lafuente, "Glutaraldehyde in bio-catalysts design: a useful crosslinker and a versatile tool in enzyme immobilization," *RSC Adv.,* vol. 4, no. 4, pp. 1583-1600, 2014, doi: 10.1039/c3ra45991h.
[282] P. Gupta, T. Maqbool, and M. Saleemuddin, "Oriented immobilization of stem bromelain via the lone histidine on a metal affinity support," (in English), *J Mol Catal B-Enzym,* vol. 45, no. 3-4, pp. 78-83, Apr 30 2007, doi: 10.1016/j.molcatb.2006.12.002.
[283] L. Zhang *et al.*, "Site-specific, covalent immobilization of PNGase F on magnetic particles mediated by microbial transglutaminase," *Anal Chim Acta,* vol. 1250, p. 340972, Apr 15 2023, doi: 10.1016/j.aca.2023.340972.
[284] E. X. Ramalho and R. J. S. de Castro, "Covalent bonding immobilization of a Bacillus licheniformis protease on chitosan and its application in protein hydrolysis," *Biocatalysis and Agricultural Biotechnology,* vol. 50, 2023, doi: 10.1016/j.bcab.2023.102713.
[285] G. Ozyilmaz, "The effect of spacer arm on hydrolytic and synthetic activity of Candida rugosa lipase immobilized on silica gel," *Journal of Molecular Catalysis B: Enzymatic,* vol. 56, no. 4, pp. 231-236, 2009, doi: 10.1016/j.molcatb.2008.05.008.
[286] R. C. Rodrigues, A. Berenguer-Murcia, D. Carballares, R. Morellon-Sterling, and R. Fernandez-Lafuente, "Stabilization of enzymes via immobilization: Multipoint covalent attachment and other stabilization strategies," *Biotechnol Adv,* vol. 52, p. 107821, Nov 15 2021, doi: 10.1016/j.biotechadv.2021.107821.
[287] C. Rother and B. Nidetzky, "Enzyme Immobilization by Microencapsulation: Methods, Materials, and Technological Applications," in *Encyclopedia of Industrial Biotechnology*, 2014, pp. 1-21.

[288] H. T. Imam, P. C. Marr, and A. C. Marr, "Enzyme entrapment, biocatalyst immobilization without covalent attachment," *Green Chemistry,* vol. 23, no. 14, pp. 4980-5005, 2021, doi: 10.1039/d1gc01852c.
[289] N. R. Jaafar, N. Jailani, R. A. Rahman, E. T. Oner, A. M. A. Murad, and R. M. Illias, "Protein surface engineering and interaction studies of maltogenic amylase towards improved enzyme immobilisation," *Int J Biol Macromol,* vol. 213, pp. 70-82, Jul 31 2022, doi: 10.1016/j.ijbiomac.2022.05.169.
[290] P. Kaul, A. Stolz, and U. C. Banerjee, "Cross-Linked Amorphous Nitrilase Aggregates for Enantioselective Nitrile Hydrolysis," *Advanced Synthesis & Catalysis,* vol. 349, no. 13, pp. 2167-2176, 2007, doi: 10.1002/adsc.200700125.
[291] R. A. Sheldon, A. Basso, and D. Brady, "New frontiers in enzyme immobilisation: robust biocatalysts for a circular bio-based economy," *Chem Soc Rev,* vol. 50, no. 10, pp. 5850-5862, May 21 2021, doi: 10.1039/d1cs00015b.
[292] O. M. Bouguerra *et al.*, "An Overview of Crosslinked Enzyme Aggregates: Concept of Development and Trends of Applications," *Appl Biochem Biotechnol,* vol. 196, no. 9, pp. 5711-5739, Sep 2024, doi: 10.1007/s12010-023-04809-y.
[293] R. A. Herman *et al.*, "Magnetic Janus SiO(2) nanoparticles immobilized protease mutant T70I as a novel clarification agent for juice processing," *Int J Biol Macromol,* vol. 292, p. 139327, Mar 2025, doi: 10.1016/j.ijbiomac.2024.139327.
[294] Q. Ni, B. Chen, S. Dong, L. Tian, and Q. Bai, "Preparation of core-shell structure Fe3 O4 @SiO2 superparamagnetic microspheres immoblized with iminodiacetic acid as immobilized metal ion affinity adsorbents for His-tag protein purification," *Biomed Chromatogr,* vol. 30, no. 4, pp. 566-73, Apr 2016, doi: 10.1002/bmc.3584.
[295] H. S. Jamal, R. Raja, S. Ahmed, G. Yesiloz, and S. A. Ali, "Immobilization of collagenase in inorganic hybrid nanoflowers with enhanced stability, proteolytic activity, and their anti-amyloid potential," *Int J Biol Macromol,* vol. 274, no. Pt 1, p. 133114, Aug 2024, doi: 10.1016/j.ijbiomac.2024.133114.
[296] Z. Wang, P. Liu, Z. Fang, and H. Jiang, "Trypsin/Zn(3)(PO(4))(2) Hybrid Nanoflowers: Controlled Synthesis and Excellent Performance as an Immobilized Enzyme," *Int J Mol Sci,* vol. 23, no. 19, Oct 6 2022, doi: 10.3390/ijms231911853.
[297] A. Balcerak-Wozniak, M. Dzwonkowska-Zarzycka, and J. Kabatc-Borcz, "A Comprehensive Review of Stimuli-Responsive Smart Polymer Materials-Recent Advances and Future Perspectives," *Materials (Basel),* vol. 17, no. 17, Aug 28 2024, doi: 10.3390/ma17174255.
[298] H. M. Abdel-Mageed *et al.*, "Bioconjugation as a smart immobilization approach for alpha-amylase enzyme using stimuli-responsive Eudragit-L100 polymer: a robust biocatalyst for applications in pharmaceutical industry," *Artif Cells Nanomed Biotechnol,* vol. 47, no. 1, pp. 2361-2368, Dec 2019, doi: 10.1080/21691401.2019.1626414.
[299] S. Khan, B. Burciu, C. D. M. Filipe, Y. Li, K. Dellinger, and T. F. Didar, "DNAzyme-Based Biosensors: Immobilization Strategies, Applications, and Future Prospective," *ACS Nano,* vol. 15, no. 9, pp. 13943-13969, Sep 28 2021, doi: 10.1021/acsnano.1c04327.
[300] B. Chen *et al.*, "DNA directed immobilization of horseradish peroxidase on phase-transitioned lysozyme modified TiO2 for efficient degradation of phenol in wastewater," *Materials & Design,* vol. 201, 2021, doi: 10.1016/j.matdes.2021.109463.
[301] M. Munir, M. Zaman, M. A. Waqar, M. A. Khan, and M. N. Alvi, "Solid lipid nanoparticles: a versatile approach for controlled release and targeted drug delivery," *J Liposome Res,* vol. 34, no. 2, pp. 335-348, Jun 2024, doi: 10.1080/08982104.2023.2268711.
[302] A. Salawi, "Self-emulsifying drug delivery systems: a novel approach to deliver drugs," *Drug Deliv,* vol. 29, no. 1, pp. 1811-1823, Dec 2022, doi: 10.1080/10717544.2022.2083724.

[303] Y. Li, W. Zhang, R. Zhao, and X. Zhang, "Advances in oral peptide drug nanoparticles for diabetes mellitus treatment," *Bioact Mater,* vol. 15, pp. 392-408, Sep 2022, doi: 10.1016/j.bioactmat.2022.02.025.
[304] A. Abramson *et al.*, "Oral delivery of systemic monoclonal antibodies, peptides and small molecules using gastric auto-injectors," *Nat Biotechnol,* vol. 40, no. 1, pp. 103-109, Jan 2022, doi: 10.1038/s41587-021-01024-0.
[305] M. L. Merz *et al.*, "De novo development of small cyclic peptides that are orally bioavailable," *Nat Chem Biol,* vol. 20, no. 5, pp. 624-633, May 2024, doi: 10.1038/s41589-023-01496-y.
[306] J. L. Hickey, D. Sindhikara, S. L. Zultanski, and D. M. Schultz, "Beyond 20 in the 21st Century: Prospects and Challenges of Non-canonical Amino Acids in Peptide Drug Discovery," *ACS Med Chem Lett,* vol. 14, no. 5, pp. 557-565, May 11 2023, doi: 10.1021/acsmedchemlett.3c00037.
[307] J. Wu, S. Roesger, N. Jones, C. J. Hu, and S. D. Li, "Cell-penetrating peptides for transmucosal delivery of proteins," *J Control Release,* vol. 366, pp. 864-878, Feb 2024, doi: 10.1016/j.jconrel.2024.01.038.
[308] P. Kurtzhals, S. Ostergaard, E. Nishimura, and T. Kjeldsen, "Derivatization with fatty acids in peptide and protein drug discovery," *Nat Rev Drug Discov,* vol. 22, no. 1, pp. 59-80, Jan 2023, doi: 10.1038/s41573-022-00529-w.
[309] D. J. Brayden and S. Maher, "Transient Permeation Enhancer(R) (TPE(R)) technology for oral delivery of octreotide: a technological evaluation," *Expert Opin Drug Deliv,* vol. 18, no. 10, pp. 1501-1512, Oct 2021, doi: 10.1080/17425247.2021.1942838.
[310] A. Das, M. Shah, and I. Saraogi, "Molecular Aspects of Insulin Aggregation and Various Therapeutic Interventions," *ACS Bio & Med Chem Au,* vol. 2, no. 3, pp. 205-221, 2022, doi: 10.1021/acsbiomedchemau.1c00054.
[311] C. Panda, S. Kumar, S. Gupta, and L. M. Pandey, "Structural, kinetic, and thermodynamic aspects of insulin aggregation," *Phys Chem Chem Phys,* vol. 25, no. 36, pp. 24195-24213, Sep 20 2023, doi: 10.1039/d3cp03103a.
[312] M. S. Ivanova, S.; Sawaya, M.; Wall, J.; Eisenberg, D.;, "Molecular basis for insulin fibril assembly," *Proceedings of the National Academy of Sciences (PNAS),* vol. 106, no. 45, pp. 18990-18995, 2009, doi: 10.1073 pnas.0910080106.
[313] O. M. Selivanova, S. Y. Grishin, A. V. Glyakina, A. S. Sadgyan, N. I. Ushakova, and O. V. Galzitskaya, "Analysis of Insulin Analogs and the Strategy of Their Further Development," *Biochemistry (Mosc),* vol. 83, no. Suppl 1, pp. S146-S162, Jan 2018, doi: 10.1134/S0006297918140122.
[314] B. Vestergaard *et al.*, "A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils," *PLoS Biol,* vol. 5, no. 5, p. e134, May 2007, doi: 10.1371/journal.pbio.0050134.
[315] A. Noormagi, K. Valmsen, V. Tougu, and P. Palumaa, "Insulin Fibrillization at Acidic and Physiological pH Values is Controlled by Different Molecular Mechanisms," *Protein J,* vol. 34, no. 6, pp. 398-403, Dec 2015, doi: 10.1007/s10930-015-9634-x.
[316] M. Zander, S. Madsbad, J. L. Madsen, and J. J. Holst, "Effect of 6-week course of glucagon-like peptide 1 on glycaemic control, insulin sensitivity, and beta-cell function in type 2 diabetes: a parallel-group study," *Lancet,* vol. 359, no. 9309, pp. 824-30, Mar 9 2002, doi: 10.1016/S0140-6736(02)07952-7.
[317] J. J. Holst, "The physiology of glucagon-like peptide 1," *Physiol Rev,* vol. 87, no. 4, pp. 1409-39, Oct 2007, doi: 10.1152/physrev.00034.2006.
[318] F. J. Becher, "Studies of the Physical Stability of GLP-1 & Chemically Modified Forms of GLP-1," Doctor of Philosophy, Chemistry, University of Cambridge, Cambridge, 2020.

[319] L. Lima and L. P. Icart, "Amyloidogenicity of peptides targeting diabetes and obesity," *Colloids Surf B Biointerfaces,* vol. 209, no. Pt 1, p. 112157, Jan 2022, doi: 10.1016/j.colsurfb.2021.112157.
[320] M. Venanzi *et al.*, "A spectroscopic and molecular dynamics study on the aggregation process of a long-acting lipidated therapeutic peptide: the case of semaglutide," *Soft Matter,* vol. 16, no. 44, pp. 10122-10131, Nov 18 2020, doi: 10.1039/d0sm01011a.
[321] O. V. Krokhin and V. Spicer, "Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides," *Anal Chem,* vol. 81, no. 22, pp. 9522-30, Nov 15 2009, doi: 10.1021/ac9016693.
[322] J. Rosenstock *et al.*, "The fate of taspoglutide, a weekly GLP-1 receptor agonist, versus twice-daily exenatide for type 2 diabetes: the T-emerge 2 trial," *Diabetes Care,* vol. 36, no. 3, pp. 498-504, Mar 2013, doi: 10.2337/dc12-0709.
[323] E. Sebokova *et al.*, "Taspoglutide, an analog of human glucagon-like Peptide-1 with enhanced stability and in vivo potency," *Endocrinology,* vol. 151, no. 6, pp. 2474-82, Jun 2010, doi: 10.1210/en.2009-1459.
[324] C. Blais, Jr., F. Marceau, J. L. Rouleau, and A. Adam, "The kallikrein-kininogen-kinin system: lessons from the quantification of endogenous kinins," *Peptides,* vol. 21, no. 12, pp. 1903-40, Dec 2000, doi: 10.1016/s0196-9781(00)00348-x.
[325] U. Yuki, "Oxytocin: An expansive review of its mechanisms, functions, and therapeutic potential," *World Journal of Advanced Research and Reviews,* vol. 19, no. 1, pp. 1264-1272, 2023, doi: 10.30574/wjarr.2023.19.1.1499.
[326] M. Lozic, O. Sarenac, D. Murphy, and N. Japundzic-Zigon, "Vasopressin, Central Autonomic Control and Blood Pressure Regulation," *Curr Hypertens Rep,* vol. 20, no. 2, p. 11, Feb 26 2018, doi: 10.1007/s11906-018-0811-0.
[327] A. Ayyad and R. A. Al-Horani, "Terlipressin for the Prevention and Treatment of Renal Decline in Hepatorenal Syndrome: A Drug Profile," *Gastroenterol Insights,* vol. 14, no. 4, pp. 420-430, Dec 2023, doi: 10.3390/gastroent14040031.
[328] J. P. Moore, Jr. and S. Wray, "Luteinizing hormone-releasing hormone (LHRH) biosynthesis and secretion in embryonic LHRH," *Endocrinology,* vol. 141, no. 12, pp. 4486-95, Dec 2000, doi: 10.1210/endo.141.12.7814.
[329] N. LeVasseur, M. Manna, and K. J. Jerzak, "An Overview of Long-Acting GnRH Agonists in Premenopausal Breast Cancer Patients: Survivorship Challenges and Management," *Curr Oncol,* vol. 31, no. 8, pp. 4209-4224, Jul 25 2024, doi: 10.3390/curroncol31080314.
[330] T. Kageyama, "Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development," *Cell Mol Life Sci,* vol. 59, no. 2, pp. 288-306, Feb 2002, doi: 10.1007/s00018-002-8423-9.
[331] I. M. Samloff, "Slow Moving Protease and the Seven Pepsinogens," *Gastroenterology,* vol. 57, no. 6, pp. 659-669, 1969, doi: 10.1016/s0016-5085(19)33821-1.
[332] T. Kageyama, M. Ichinose, K. Miki, S. B. Athauda, M. Tanji, and K. Takahashi, "Difference of activation processes and structure of activation peptides in human pepsinogens A and progastricsin," *J Biochem,* vol. 105, no. 1, pp. 15-22, Jan 1989, doi: 10.1093/oxfordjournals.jbchem.a122610.
[333] A. T. Jones, K. K. Balan, S. A. Jenkins, R. Sutton, M. Critchley, and N. B. Roberts, "Assay of gastricsin and individual pepsins in human gastric juice," *J Clin Pathol,* vol. 46, no. 3, pp. 254-8, Mar 1993, doi: 10.1136/jcp.46.3.254.
[334] B. Veerapandian *et al.*, "Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases," *Protein Sci,* vol. 1, no. 3, pp. 322-8, Mar 1992, doi: 10.1002/pro.5560010303.
[335] M. Okoniewska, T. Tanaka, and R. Y. Yada, "The role of the flap residue, threonine 77, in the activation and catalytic activity of pepsin A," *Protein Eng,* vol. 12, no. 1, pp. 55-61, Jan 1999, doi: 10.1093/protein/12.1.55.

[336] T. J. Cottrell, L. J. Harris, T. Tanaka, and R. Y. Yada, "The sole lysine residue in porcine pepsin works as a key residue for catalysis and conformational flexibility," *J Biol Chem,* vol. 270, no. 34, pp. 19974-8, Aug 25 1995, doi: 10.1074/jbc.270.34.19974.
[337] A. R. Sielecki, A. A. Fedorov, A. Boodhoo, N. S. Andreeva, and M. N. James, "Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 A resolution," *J Mol Biol,* vol. 214, no. 1, pp. 143-70, Jul 5 1990, doi: 10.1016/0022-2836(90)90153-D.
[338] H. Kageyama *et al.*, "Differences in the P1' substrate specificities of pepsin A and chymosin," *J Biochem,* vol. 147, no. 2, pp. 167-74, Feb 2010, doi: 10.1093/jb/mvp158.
[339] B. Keil, *Specificity of Proteolysis*. 1992.
[340] J. C. Powers, A. D. Harley, and D. V. Myers, "Subsite specificity of porcine pepsin," *Adv Exp Med Biol,* vol. 95, pp. 141-57, 1977, doi: 10.1007/978-1-4757-0719-9_9.
[341] Y. Hamuro, S. J. Coales, K. S. Molnar, S. J. Tuske, and J. A. Morrow, "Specificity of immobilized porcine pepsin in H/D exchange compatible conditions," *Rapid Commun Mass Spectrom,* vol. 22, no. 7, pp. 1041-6, Apr 2008, doi: 10.1002/rcm.3467.
[342] P. R. Schimmel and P. J. Flory, "Conformational energies and configurational statistics of copolypeptides containing L-proline," *J Mol Biol,* vol. 34, no. 1, pp. 105-20, May 28 1968, doi: 10.1016/0022-2836(68)90237-4.
[343] K. J. Stanforth *et al.*, "Pepsin properties, structure, and its accurate measurement: a narrative review," *Annals of Esophagus,* vol. 5, pp. 31-31, 2022, doi: 10.21037/aoe-20-95.
[344] L. A. Campos and J. Sancho, "The active site of pepsin is formed in the intermediate conformation dominant at mildly acidic pH," *FEBS Lett,* vol. 538, no. 1-3, pp. 89-95, Mar 13 2003, doi: 10.1016/s0014-5793(03)00152-2.
[345] S. A. Jacobs, S. J. Wu, Y. Feng, D. Bethea, and K. T. O'Neil, "Cross-interaction chromatography: a rapid method to identify highly soluble monoclonal antibody candidates," *Pharm Res,* vol. 27, no. 1, pp. 65-71, Jan 2010, doi: 10.1007/s11095-009-0007-z.
[346] S. H. M. Hedberg, L. G. Brown, A. Meghdadi, and D. R. Williams, "Improved adsorption reactions, kinetics and stability for model and therapeutic proteins immobilised on affinity resins," *Adsorption (Boston),* vol. 25, no. 6, pp. 1177-1190, 2019, doi: 10.1007/s10450-019-00106-5.
[347] P. Monsan, "Optimization of glutaraldehyde activation of a support for enzyme immobilization," *Journal of Molecular Catalysis,* vol. 3, no. 5, pp. 371-384, 1978, doi: 10.1016/0304-5102(78)80026-1.
[348] W. Han *et al.*, "Pepsin immobilization on an aldehyde-modified polymethacrylate monolith and its application for protein analysis," *J Biosci Bioeng,* vol. 119, no. 5, pp. 505-10, May 2015, doi: 10.1016/j.jbiosc.2014.10.018.
[349] A. Quigley, J. Y. Heng, J. M. Liddell, and D. R. Williams, "The accurate measurement of second virial coefficients using self-interaction chromatography: experimental considerations," *Eur J Pharm Biopharm,* vol. 85, no. 3 Pt B, pp. 1103-11, Nov 2013, doi: 10.1016/j.ejpb.2013.04.004.
[350] Z. Papai and T. L. Pap, "Analysis of peak asymmetry in chromatography," *J Chromatogr A,* vol. 953, no. 1-2, pp. 31-8, Apr 12 2002, doi: 10.1016/s0021-9673(02)00121-8.
[351] S. H. Hedberg, J. Y. Heng, D. R. Williams, and J. M. Liddell, "Self-Interaction Chromatography of mAbs: Accurate Measurement of Dead Volumes," *Pharm Res,* vol. 32, no. 12, pp. 3975-85, Dec 2015, doi: 10.1007/s11095-015-1758-3.
[352] C. A. Teske, H. W. Blanch, and J. M. Prausnitz, "Measurement of Lysozyme-Lysozyme Interactions with Quantitative Affinity Chromatography," *Journal of Physical Chemistry B,* vol. 108, no. 22, pp. 7437–7444, May 12 2004 2004, doi: 10.1021/jp0361576.
[353] H. Barth, "Chromatography Fundamentals, Part V: Theoretical Plates: Significance, Properties, and Uses," *LCGC North America,* vol. 36, no. 11, 2018.

[354] Tosoh Bioscience. "TSKgel G3000SWxl." https://www.separations.us.tosohbioscience.com/HPLC_Columns/id-8327/TSKgel_G3000SWxl (accessed.
[355] S. Y. Patro and T. M. Przybycien, "Self-interaction chromatography: A tool for the study of protein–protein interactions in bioprocessing environments," presented at the Biotechnology and Bioengineering, 20 October 1996, 1996.
[356] V. Le Brun, W. Friess, S. Bassarab, S. Muhlau, and P. Garidel, "A critical evaluation of self-interaction chromatography as a predictive tool for the assessment of protein-protein interactions in protein formulation development: a case study of a therapeutic monoclonal antibody," *Eur J Pharm Biopharm,* vol. 75, no. 1, pp. 16-25, May 2010, doi: 10.1016/j.ejpb.2010.01.009.
[357] R. A. Lewus, P. A. Darcy, A. M. Lenhoff, and S. I. Sandler, "Interactions and phase behavior of a monoclonal antibody," *Biotechnol Prog,* vol. 27, no. 1, pp. 280-9, Jan-Feb 2011, doi: 10.1002/btpr.536.
[358] S. H. M. Hedberg, R. J, J. M. Haigh, and D. R. Williams, "Cross-interaction chromatography as a rapid screening technique to identify the stability of new antibody therapeutics," *Eur J Pharm Biopharm,* vol. 133, pp. 131-137, 2018, doi: 10.1016/j.ejpb.2018.10.009.
[359] I. Dmitriev, I. Kuryndin, N. Bobrova, and M. Smirnov, "Swelling behavior and network characterization of hydrogels from linear polyacrylamide crosslinked with glutaraldehyde," *Materials Today Communications,* vol. 4, pp. 93-100, 2015, doi: 10.1016/j.mtcomm.2015.06.005.
[360] F. Horkay and P. J. Basser, "Ionic and pH effects on the osmotic properties and structure of polyelectrolyte gels," *J Polym Sci B Polym Phys,* vol. 46, no. 24, pp. 2803-2810, Dec 15 2008, doi: 10.1002/polb.21590.
[361] BenchChem Technical Support Team, "Technical Support Center: Troubleshooting HPLC Peak Tailing for Modified Peptides," 2025. [Online]. Available: https://pdf.benchchem.com/1579/Technical_Support_Center_Troubleshooting_HPLC_Peak_Tailing_for_Modified_Peptides.pdf
[362] J. J. van Deemter, F. J. Zuiderweg, and A. Klinkenberg, "Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography," *Chemical Engineering Science,* vol. 5, no. 6, pp. 271-289, 1956, doi: 10.1016/0009-2509(56)80003-1.
[363] T. D. Brown, K. A. Whitehead, and S. Mitragotri, "Materials for oral delivery of proteins and peptides," *Nature Reviews Materials,* vol. 5, no. 2, pp. 127-148, 2019, doi: 10.1038/s41578-019-0156-6.
[364] K. C. Baral and K. Y. Choi, "Barriers and Strategies for Oral Peptide and Protein Therapeutics Delivery: Update on Clinical Advances," *Pharmaceutics,* vol. 17, no. 4, Mar 21 2025, doi: 10.3390/pharmaceutics17040397.
[365] C. Maderuelo, J. M. Lanao, and A. Zarzuelo, "Enteric coating of oral solid dosage forms as a tool to improve drug bioavailability," *Eur J Pharm Sci,* vol. 138, p. 105019, Oct 1 2019, doi: 10.1016/j.ejps.2019.105019.
[366] W. Wang, S. K. Singh, N. Li, M. R. Toler, K. R. King, and S. Nema, "Immunogenicity of protein aggregates--concerns and realities," *Int J Pharm,* vol. 431, no. 1-2, pp. 1-11, Jul 15 2012, doi: 10.1016/j.ijpharm.2012.04.040.
[367] (2022). *Semaglutide (Rybelsus®)*. [Online] Available: https://www.hhs.texas.gov/sites/default/files/documents/semaglutide-rybelsus-monograph.pdf
[368] I. K. Mkam Tsengam *et al.*, "Transformation of Lipid Vesicles into Micelles by Adding Nonionic Surfactants: Elucidating the Structural Pathway and the Intermediate

Structures," *J Phys Chem B,* vol. 126, no. 11, pp. 2208-2216, Mar 24 2022, doi: 10.1021/acs.jpcb.1c09685.
[369] K. Yamada, K. Yagishita, and T. Sato, "Dispersion states of lecithin–soybean oil mixtures in water," *Colloid and Polymer Science,* vol. 300, no. 2, pp. 95-102, 2022, doi: 10.1007/s00396-021-04932-4.
[370] S. Hvidt, "Insulin association in neutral solutions studied by light scattering," *Biophys Chem,* vol. 39, no. 2, pp. 205-13, Feb 1991, doi: 10.1016/0301-4622(91)85023-j.
[371] C. F. Hjorth *et al.*, "Structure, Aggregation, and Activity of a Covalent Insulin Dimer Formed During Storage of Neutral Formulation of Human Insulin," *J Pharm Sci,* vol. 105, no. 4, pp. 1376-86, Apr 2016, doi: 10.1016/j.xphs.2016.01.003.
[372] M. Balamurugan and S. Sarumathy, "Advancements in Oral Drug Delivery of Insulin: A Review," *International Journal of Diabetes in Developing Countries,* 2024, doi: 10.1007/s13410-024-01416-3.
[373] C. Y. Low *et al.*, "Critical updates on oral insulin drug delivery systems for type 2 diabetes mellitus," *J Nanobiotechnology,* vol. 23, no. 1, p. 16, Jan 15 2025, doi: 10.1186/s12951-024-03062-7.
[374] E. Smirnova, I. Safenkova, V. Stein-Margolina, V. Shubin, and B. Gurvits, "Can aggregation of insulin govern its fate in the intestine? Implications for oral delivery of the drug," *Int J Pharm,* vol. 471, no. 1-2, pp. 65-8, Aug 25 2014, doi: 10.1016/j.ijpharm.2014.05.025.
[375] H. P. Erickson, "Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy," *Biol Proced Online,* vol. 11, pp. 32-51, May 15 2009, doi: 10.1007/s12575-009-9008-x.
[376] A. Scotti *et al.*, "The CONTIN algorithm and its application to determine the size distribution of microgel suspensions," *J Chem Phys,* vol. 142, no. 23, p. 234905, Jun 21 2015, doi: 10.1063/1.4921686.
[377] H. G. Barth and F. J. Carlin, "A Review of Polymer Shear Degradation in Size-Exclusion Chromatography," *Journal of Liquid Chromatography,* vol. 7, no. 9, pp. 1717-1738, 2006, doi: 10.1080/01483918408068832.
[378] I. Javeri *et al.*, "Use of Cyclodextrin as a Novel Agent in the SEC-HPLC Mobile Phase to Mitigate the Interactions of Proteins or Peptide or their Impurities with the Residual Silanols of Commercial SEC-HPLC Columns with Improved Separation and Resolution," *Pharm Res,* vol. 35, no. 9, p. 168, Jun 26 2018, doi: 10.1007/s11095-018-2446-x.
[379] V. Marassi *et al.*, "Application of Af4-Multidetection to Liraglutide in Its Formulation: Preserving and Representing Native Aggregation," *Molecules,* vol. 27, no. 17, Aug 26 2022, doi: 10.3390/molecules27175485.
[380] I. Dimitrov, F. Hodzhaoglu, and I. Ivanov, "In Vitro Dissolution of Insulin Crystal Polymorphs at Model Conditions Relevant to In Vivo Environment," *Dissolution Technologies,* vol. 20, no. 4, pp. 11-16, 2013, doi: 10.14227/dt200413p11.
[381] S. Havelund, U. Ribel, F. Hubalek, T. Hoeg-Jensen, P. O. Wahlund, and I. Jonassen, "Investigation of the Physico-Chemical Properties that Enable Co-Formulation of Basal Insulin Degludec with Fast-Acting Insulin Aspart," *Pharm Res,* vol. 32, no. 7, pp. 2250-8, Jul 2015, doi: 10.1007/s11095-014-1614-x.
[382] R. K. Chitta, D. L. Rempel, M. A. Grayson, E. E. Remsen, and M. L. Gross, "Application of SIMSTEX to oligomerization of insulin analogs and mutants," *J Am Soc Mass Spectrom,* vol. 17, no. 11, pp. 1526-1534, Nov 2006, doi: 10.1016/j.jasms.2006.08.004.
[383] L. Bromberg, J. Rashba-Step, and T. Scott, "Insulin particle formation in supersaturated aqueous solutions of poly(ethylene glycol)," *Biophys J,* vol. 89, no. 5, pp. 3424-33, Nov 2005, doi: 10.1529/biophysj.105.062802.

[384] S.-H. Wang, X.-Y. Dong, and Y. Sun, "Effect of (−)-epigallocatechin-3-gallate on human insulin fibrillation/aggregation kinetics," *Biochemical Engineering Journal,* vol. 63, pp. 38-49, 2012, doi: 10.1016/j.bej.2012.02.002.
[385] C. Ren *et al.*, "Bioinspired pH-Responsive Microalgal Hydrogels for Oral Insulin Delivery with Both Hypoglycemic and Insulin Sensitizing Effects," *ACS Nano,* vol. 17, no. 14, pp. 14161-14175, Jul 25 2023, doi: 10.1021/acsnano.3c04897.
[386] Y. Zhang, Y. Wang, Y. Tang, R. Li, and Y. Ji, "An online immobilized pepsin microreactor based on polymer monoliths for screening inhibitors from natural products," *Analytical Methods,* vol. 11, no. 18, pp. 2465-2472, 2019, doi: 10.1039/c9ay00343f.
[387] Y. Long and T. D. Wood, "Immobilized pepsin microreactor for rapid peptide mapping with nanoelectrospray ionization mass spectrometry," *J Am Soc Mass Spectrom,* vol. 26, no. 1, pp. 194-7, Jan 2015, doi: 10.1007/s13361-014-1015-8.
[388] J. Ahn, M. C. Jung, K. Wyndham, Y. Q. Yu, and J. R. Engen, "Pepsin immobilized on high-strength hybrid particles for continuous flow online digestion at 10,000 psi," *Anal Chem,* vol. 84, no. 16, pp. 7256-62, Aug 21 2012, doi: 10.1021/ac301749h.
[389] S. Hudson, E. Magner, J. Cooney, and B. K. Hodnett, "Methodology for the immobilization of enzymes onto mesoporous materials," *J Phys Chem B,* vol. 109, no. 41, pp. 19496-506, Oct 20 2005, doi: 10.1021/jp052102n.
[390] R. Chang *et al.*, "Sensory-Guided Isolation, Identification, and Active Site Calculation of Novel Umami Peptides from Ethanol Precipitation Fractions of Fermented Grain Wine (Huangjiu)," *Foods,* vol. 12, no. 18, Sep 11 2023, doi: 10.3390/foods12183398.
[391] M. G. Holyavka and V. G. Artyukhov, "Prospects and problems in enzyme immobilization methodology: comprehensive review," *Biophys Rev,* vol. 17, no. 5, pp. 1471-1509, Oct 2025, doi: 10.1007/s12551-025-01360-4.
[392] P. V. Iyer and L. Ananthanarayan, "Enzyme stability and stabilization—Aqueous and non-aqueous environment," *Process Biochemistry,* vol. 43, no. 10, pp. 1019-1032, 2008, doi: 10.1016/j.procbio.2008.06.004.
[393] M. Kolinski, S. Kmiecik, R. Dec, M. Piejko, P. Mak, and W. Dzwolak, "Docking interactions determine early cleavage events in insulin proteolysis by pepsin: Experiment and simulation," *Int J Biol Macromol,* vol. 149, pp. 1151-1160, Apr 15 2020, doi: 10.1016/j.ijbiomac.2020.01.253.
[394] C. Vorauer *et al.*, "Rapid Assessment of Pepsin Column Activity for Reliable HDX-MS Studies," *J Am Soc Mass Spectrom,* vol. 32, no. 9, pp. 2386-2390, Sep 1 2021, doi: 10.1021/jasms.1c00080.

12. SUPPLEMENTARY MATERIALS

12.1. Calibration curves

12.1.1. Peptide immobilisation solutions

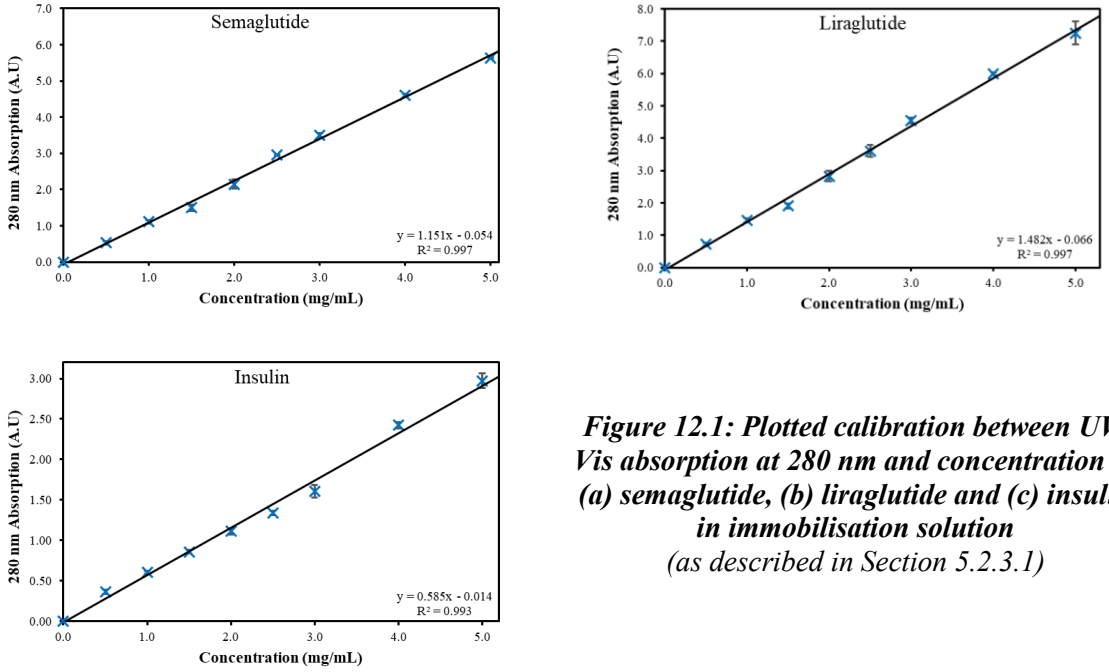


Figure 12.1: Plotted calibration between UV-Vis absorption at 280 nm and concentration of (a) semaglutide, (b) liraglutide and (c) insulin in immobilisation solution
(as described in Section 5.2.3.1)

12.1.2. Pepsin immobilisation solutions

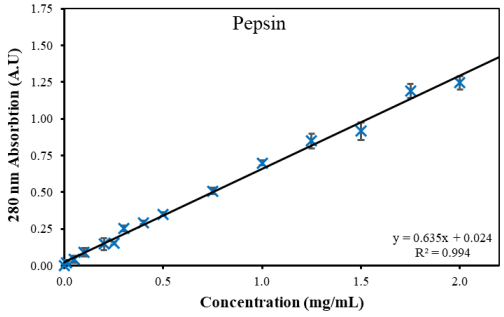


Figure 12.2: Plotted calibration between UV-Vis absorption at 280 nm and concentration of pepsin *in 50 mM pH 4.4 sodium citrate (as described in Section 5.2.3.3)*

12.1.3. Peptide self-interaction chromatography concentrations

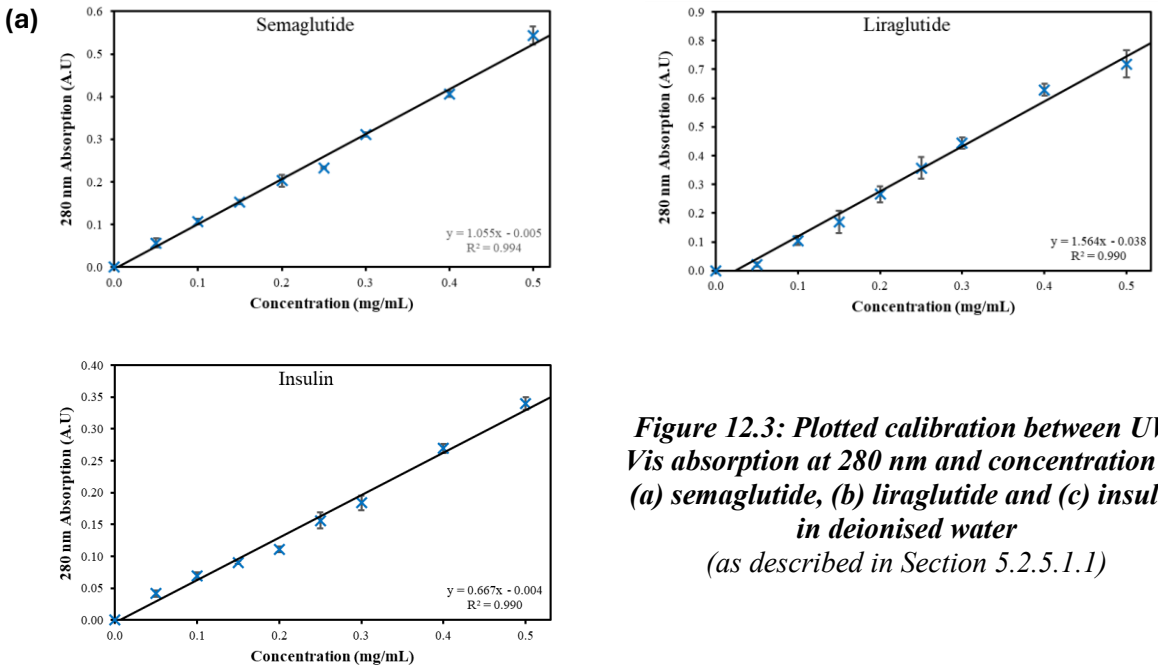


Figure 12.3: Plotted calibration between UV-Vis absorption at 280 nm and concentration of (a) semaglutide, (b) liraglutide and (c) insulin in deionised water
(as described in Section 5.2.5.1.1)

12.2. Example elution curves

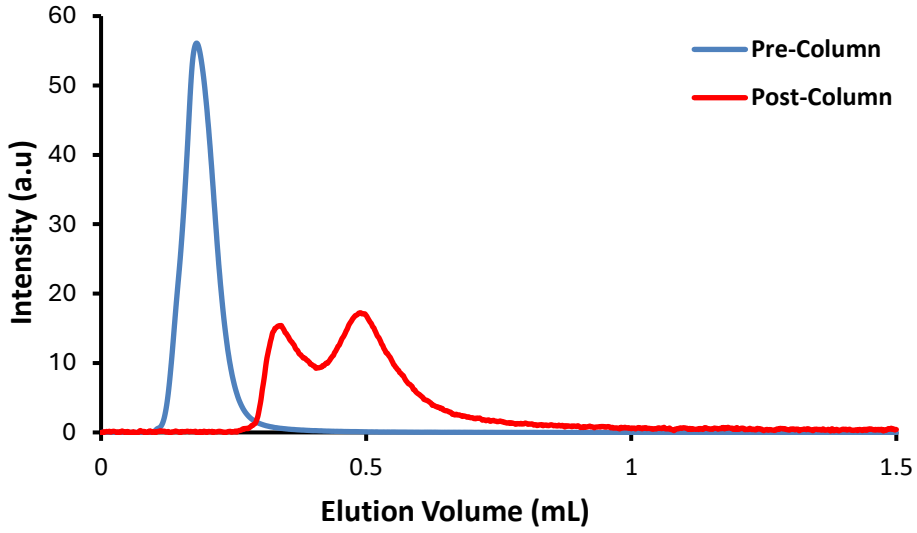


Figure 12.4: Example elution profile of 0.1 mg/mL semaglutide in pH 5.0 mobile phase *through a tris-functionalised HW-65 AF-Tresyl-650M stationary phase at 0.05 mL/min in 20 mM sodium citrate (see Section 6.3)*

12.3. Size-exclusion chromatography calibration curve

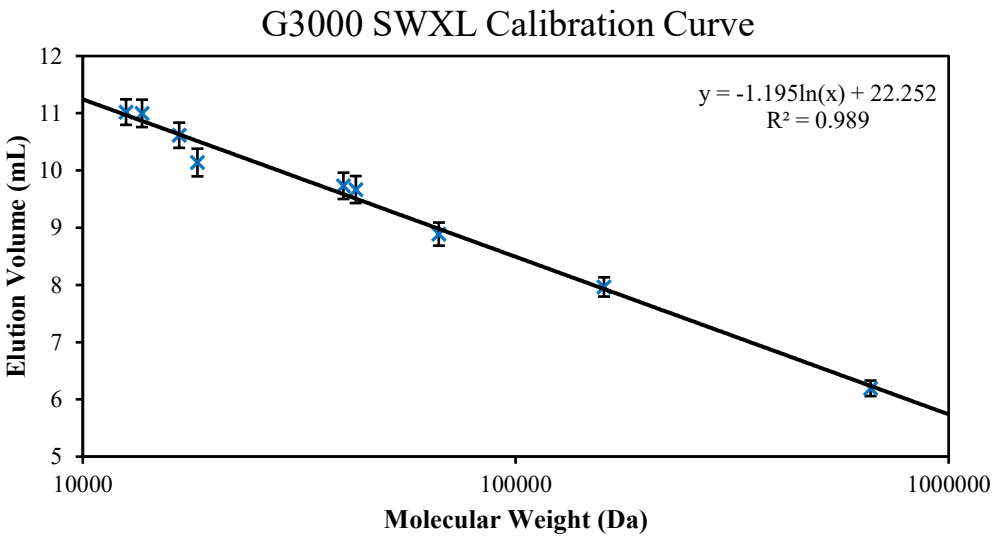


Figure 12.5: Plotted calibration between protein standard mixture molecular weights and retention time in a TSKgel G3000SW$_{XL}$ column

12.4. Example raw dynamic light scattering correlograms

12.4.1. Semaglutide

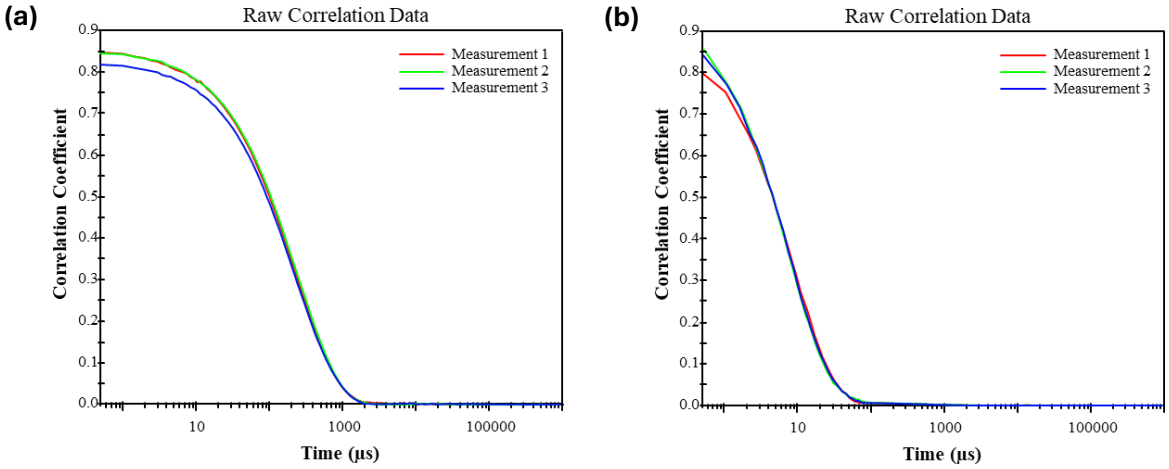


Figure 12.6 Correlograms from dynamic light scattering measurements of 0.5 mg/mL semaglutide at (a) pH 4.0 and (b) pH8.0 corresponding to Figure 7.1(a) and (b)

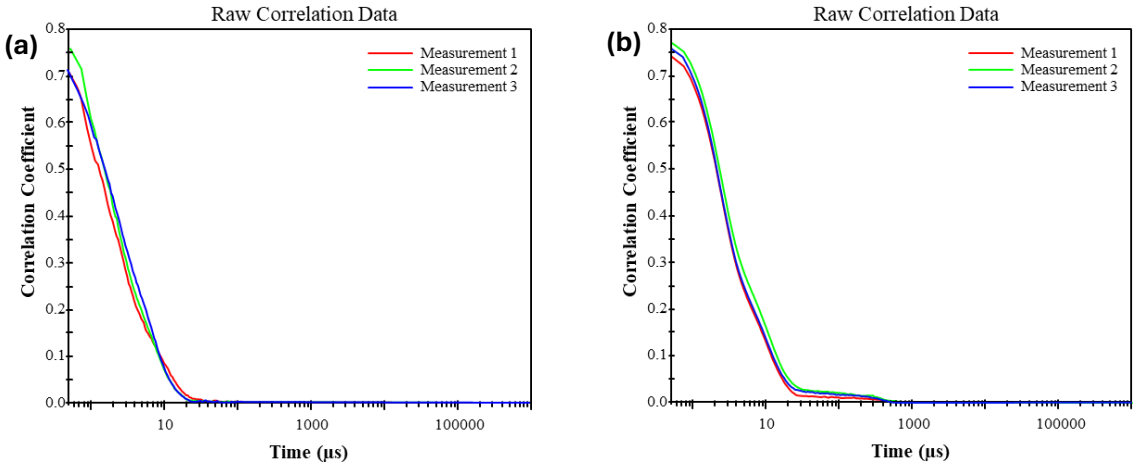


Figure 12.7: Correlograms from dynamic light scattering measurements of 0.1 mg/mL semaglutide *in water (a) immediately after dissolution and (b) after 24 hours corresponding to Figure 7.5*

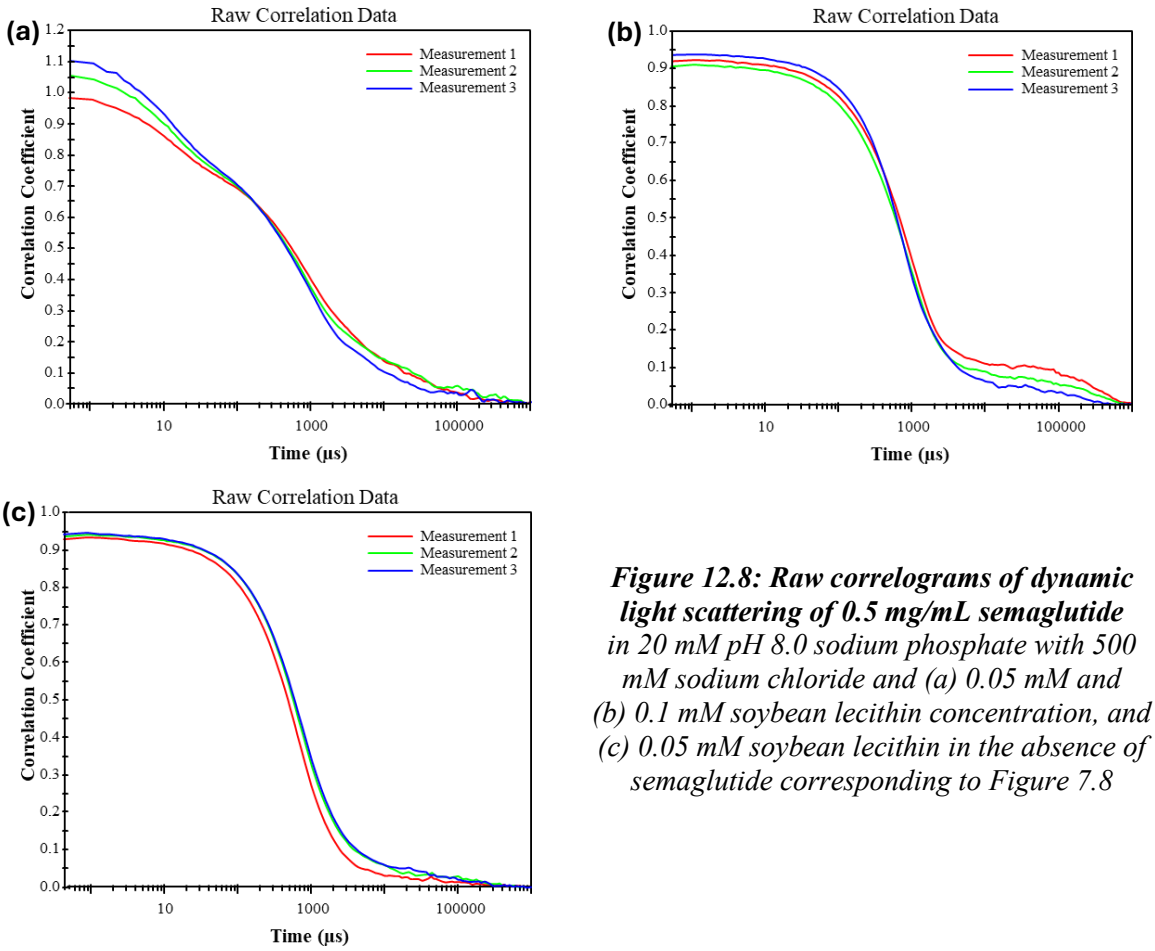


Figure 12.8: Raw correlograms of dynamic light scattering of 0.5 mg/mL semaglutide *in 20 mM pH 8.0 sodium phosphate with 500 mM sodium chloride and (a) 0.05 mM and (b) 0.1 mM soybean lecithin concentration, and (c) 0.05 mM soybean lecithin in the absence of semaglutide corresponding to Figure 7.8*

12.4.2. Liraglutide

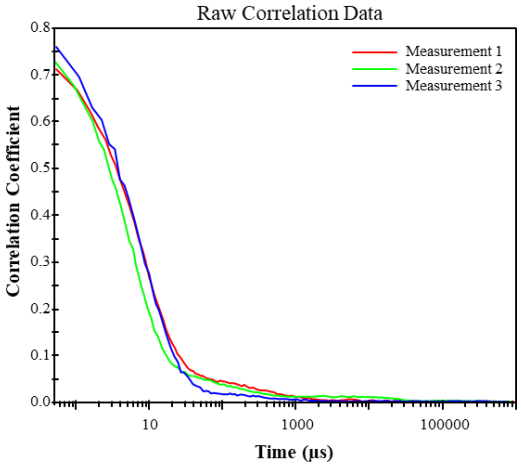


Figure 12.9: Correlogram of dynamic light scattering measurements of 0.5 mg/mL liraglutide *in pH 8.0 20 mM sodium phosphate buffer with 0 mM sodium chloride corresponding to Figure 8.1*

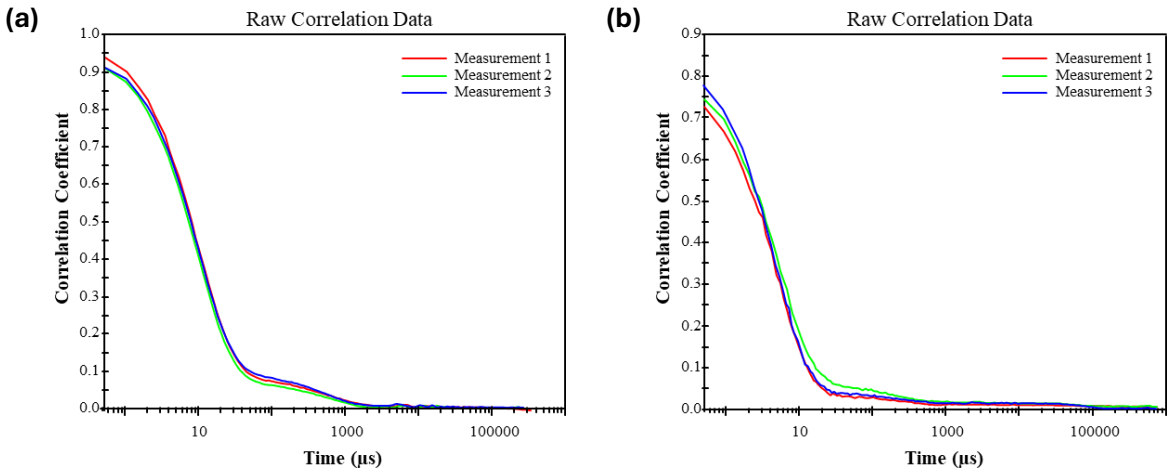


Figure 12.10: Correlograms of dynamic light scattering measurements of 0.5 mg/mL liraglutide after 24 hours incubation
in (a) pH 6.0 and (b) pH 8.0 20 mM sodium phosphate corresponding to Figure 8.3(a) and (b)

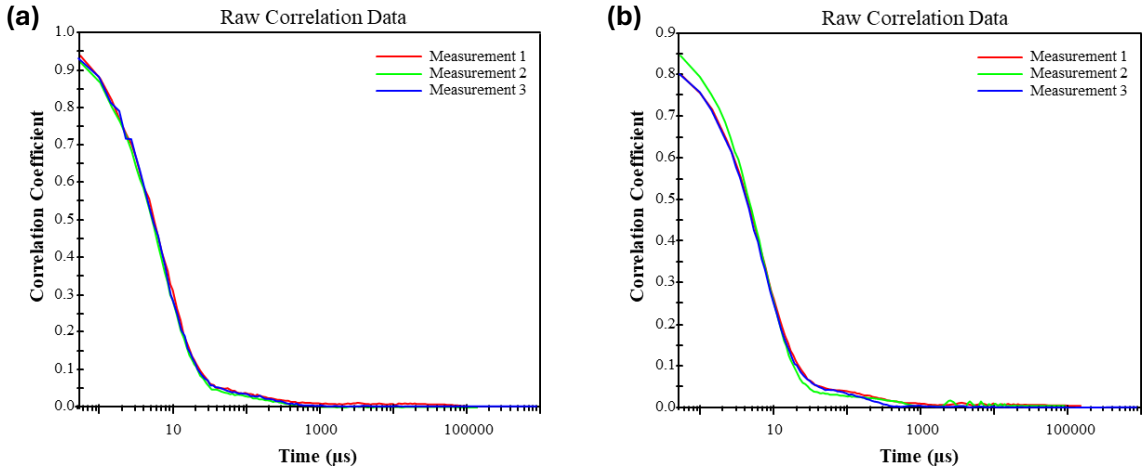


Figure 12.11: Correlograms from dynamic light scattering of 0.1 mg/mL liraglutide in water *self-interaction chromatography sample (a) immediately after dissolution and (b) after 6 hours corresponding to Figure 8.7*

12.4.3. Insulin

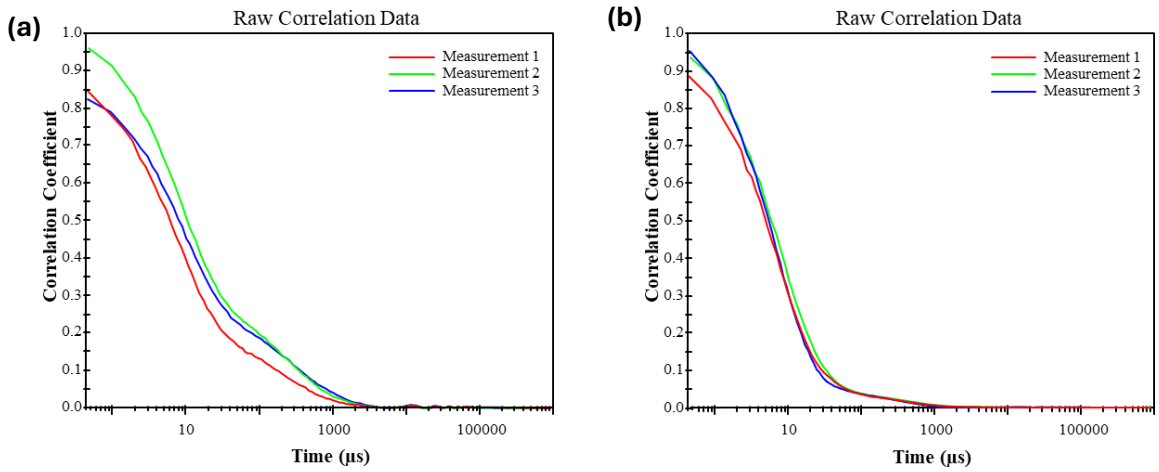


Figure 12.12: Correlograms of dynamic light scattering measurements of 0.5 mg/mL insulin *in pH 7.0 20 mM sodium phosphate with 0 mM sodium chloride (a) immediately and (b) 24 hours after dissolution corresponding to Figure 8.10(a) and (b)*

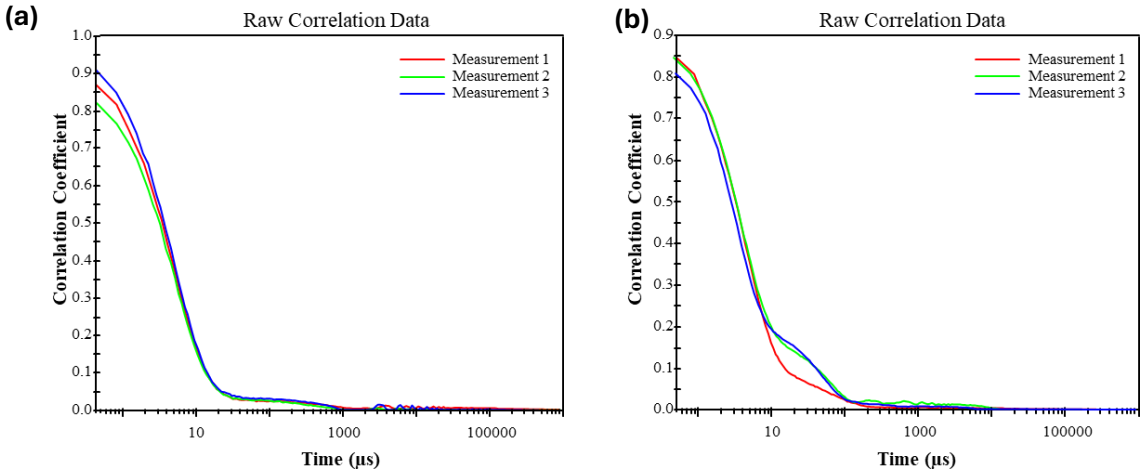


Figure 12.13: Correlograms of dynamic light scattering measurements of 0.1 mg/mL insulin *in water self-interaction chromatography sample (a) immediately after dissolution and (b) after 6 hours*

12.5. Exemplary Mass Spectra

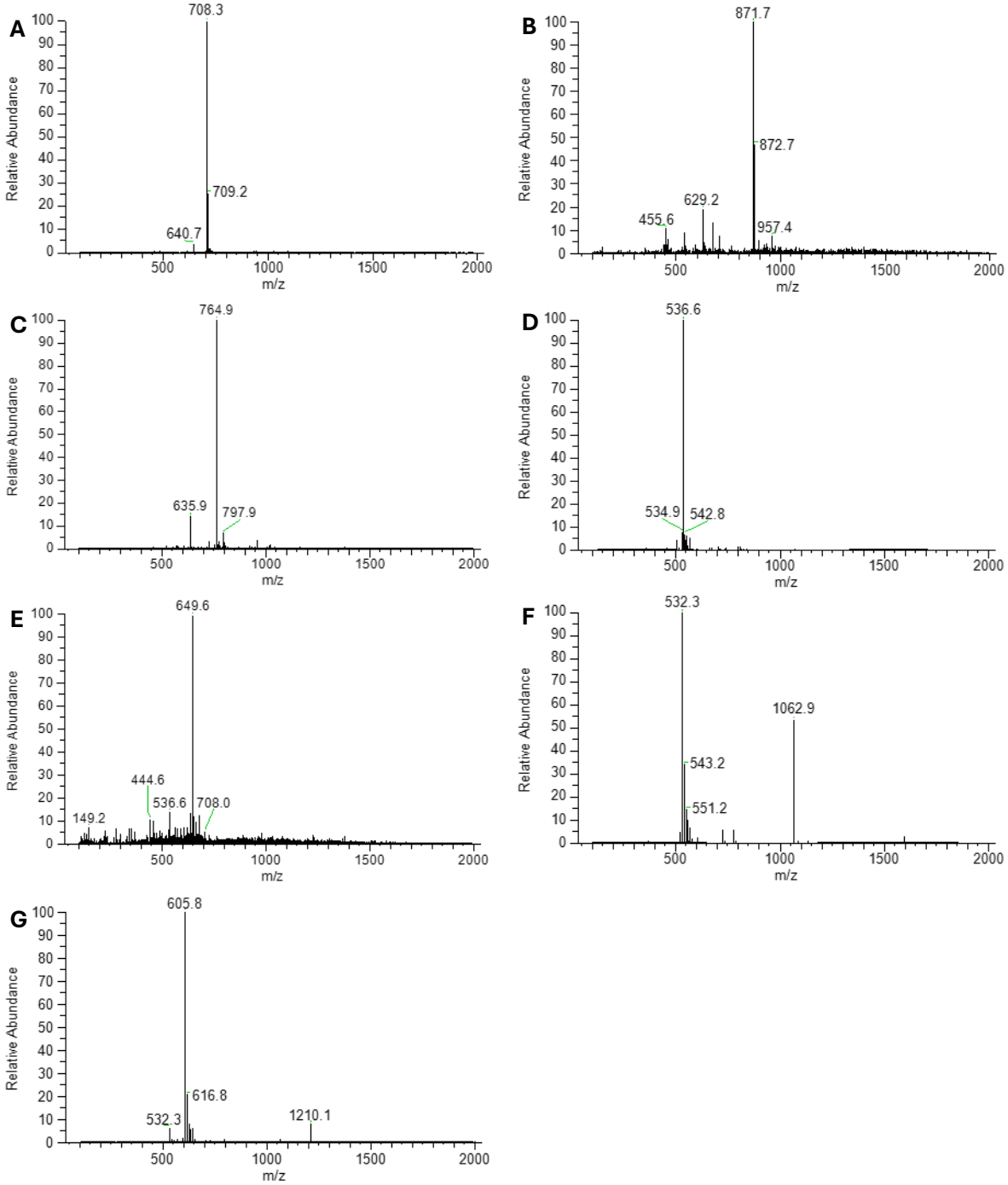


Figure 12.14: Mass spectra of the major peaks of semaglutide after elution through an immobilised pepsin reactor

(A-G) in order of retention time as outlined in Figure 9.15B

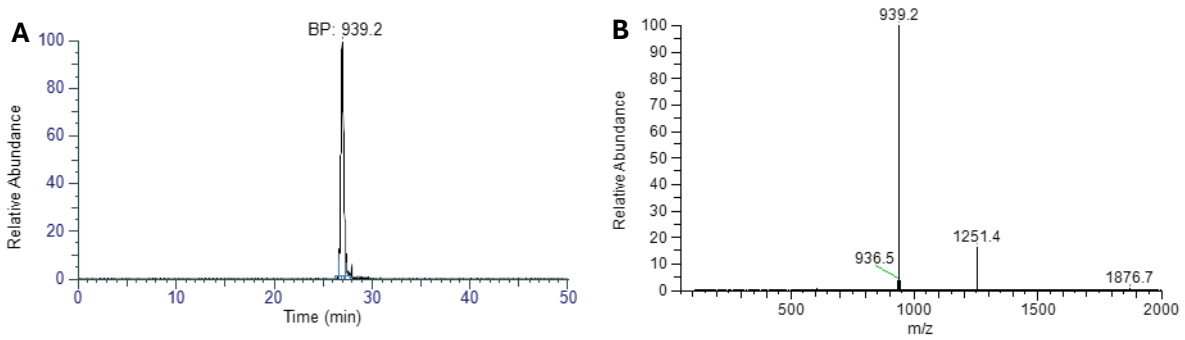


Figure 12.15: LC-MS (A) Base peak chromatogram and (B) primary peak spectrum of liraglutide after elution through an end-capped propylamino silica packed column

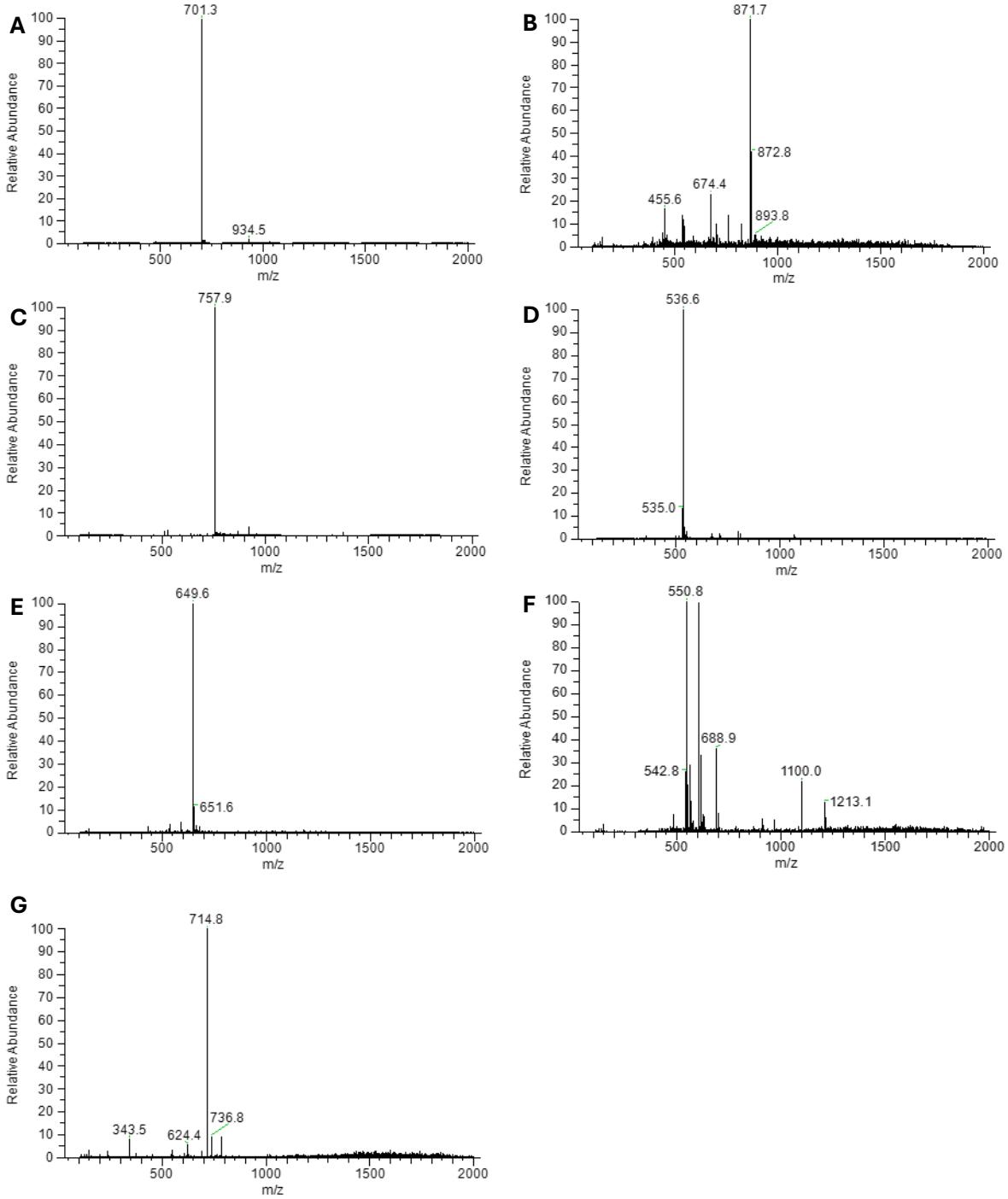


Figure 12.16: Mass spectra of the major peaks of liraglutide after elution through an immobilised pepsin reactor

(A-G) in order of retention time as outlined in Figure 9.16B

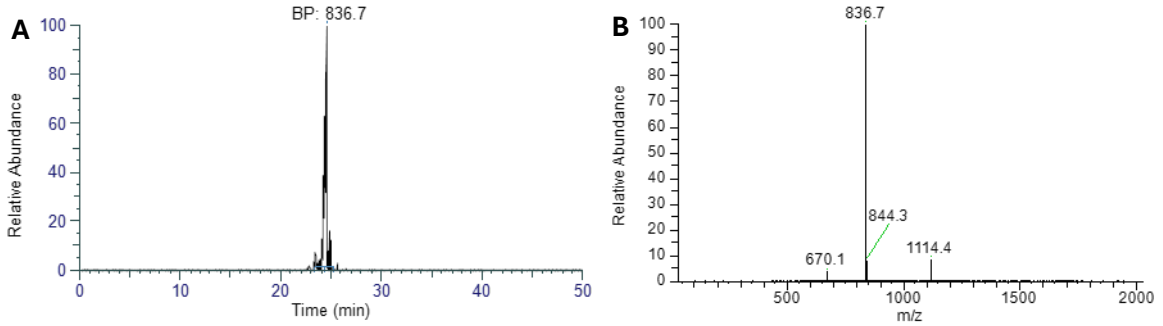


Figure 12.17: LC-MS (A) Base peak chromatogram and (B) primary peak spectrum of taspoglutide after elution through an end-capped propylamino silica packed column

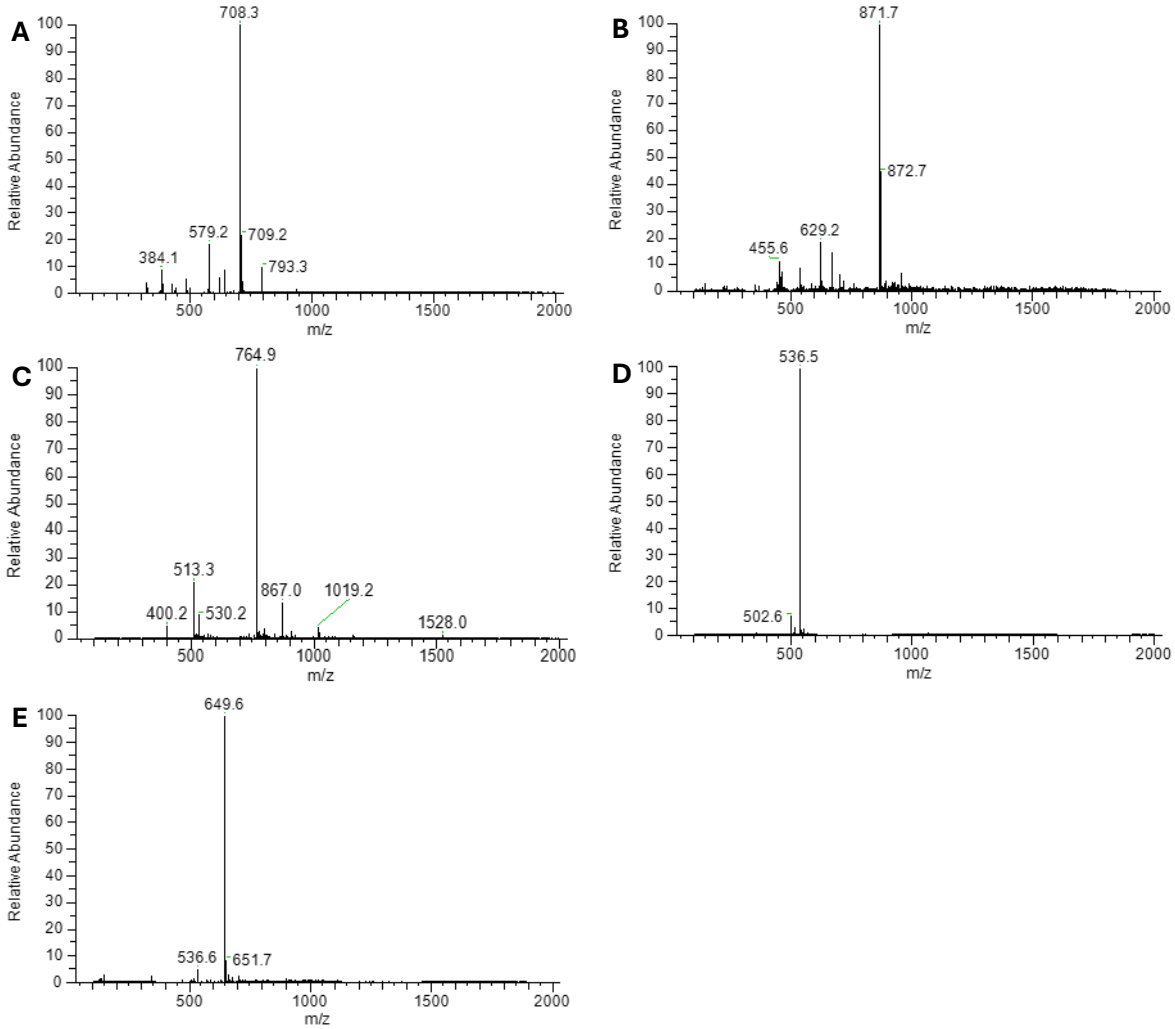


Figure 12.18: Mass spectra of the major peaks of taspoglutide after elution through an immobilised pepsin reactor

(A-E) in order of retention time as outlined in Figure 9.17B

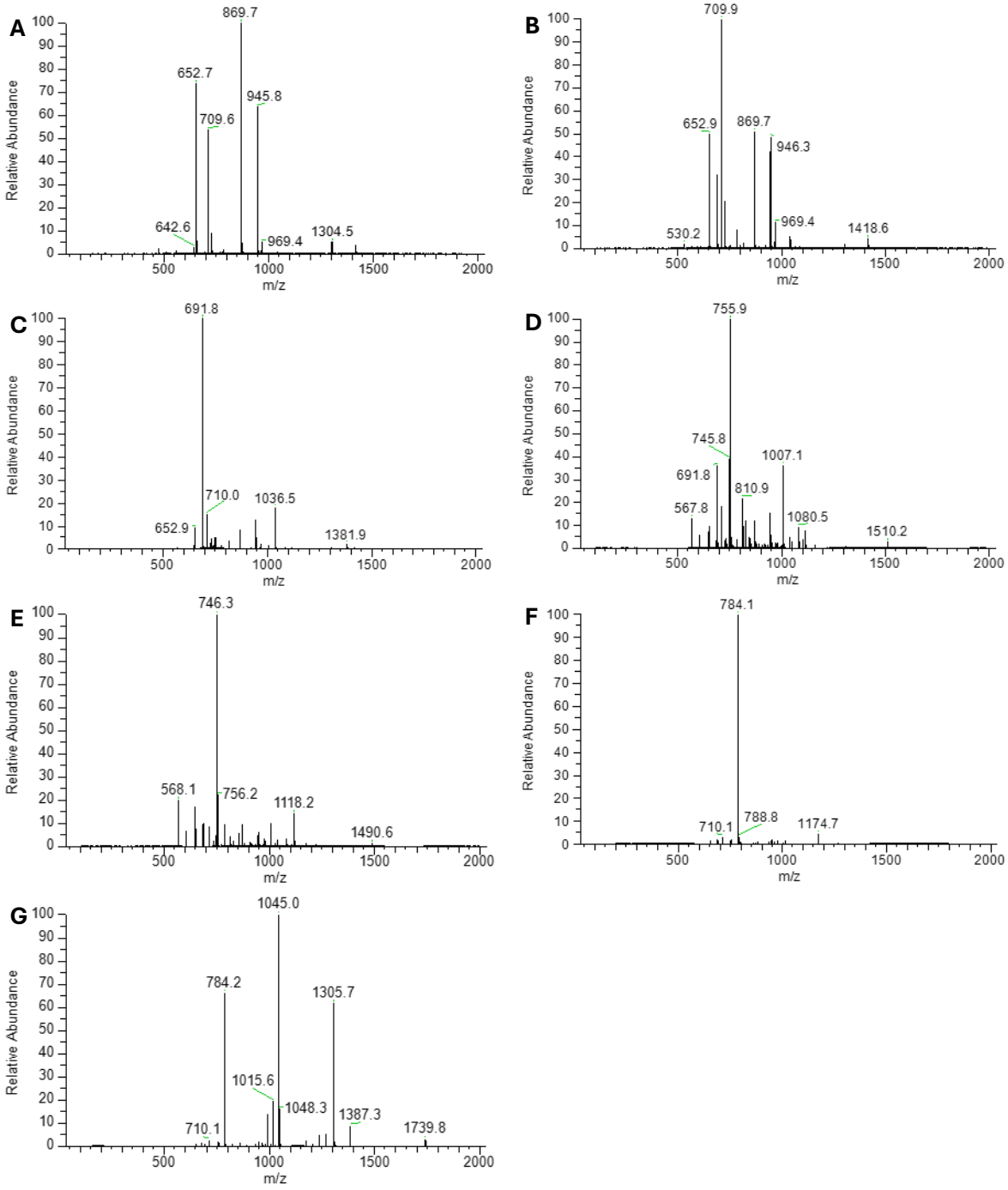


Figure 12.19: Mass spectra of the major peaks of insulin after elution through an immobilised pepsin reactor

(A-E) in order of retention time as outlined in Figure 9.19

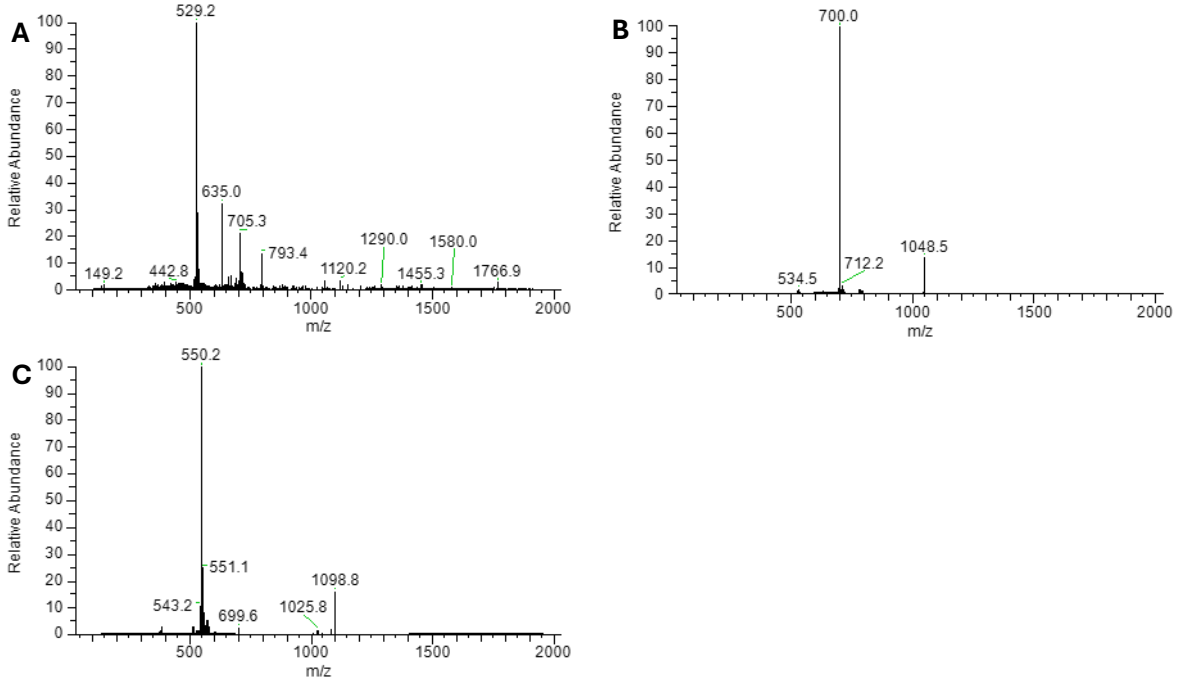


Figure 12.20: Exemplary mass spectra of the major peaks of vasopressin after elution through a column packed with ethanolamine end-capped propylamino silica resin (A-C)
in order of retention time as outlined in Figure 9.22B$_1$

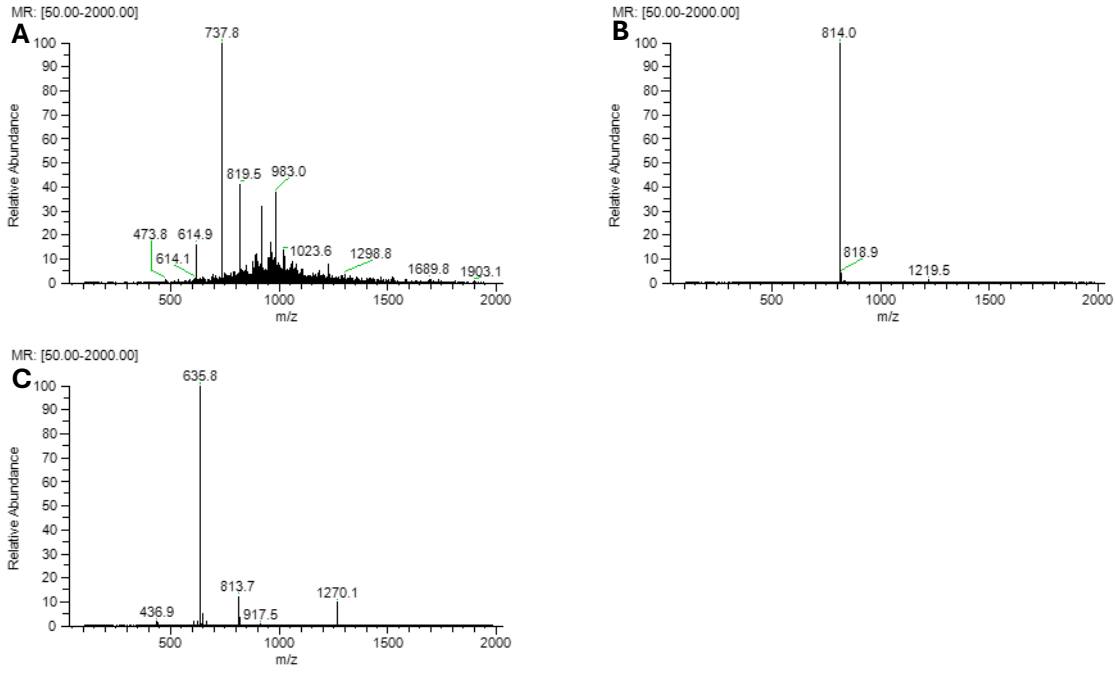


Figure 12.21: Mass spectra of the major peaks of terlipressin after elution through an immobilised pepsin reactor
(A-C) in order of retention time as outlined in Figure 9.22C$_2$

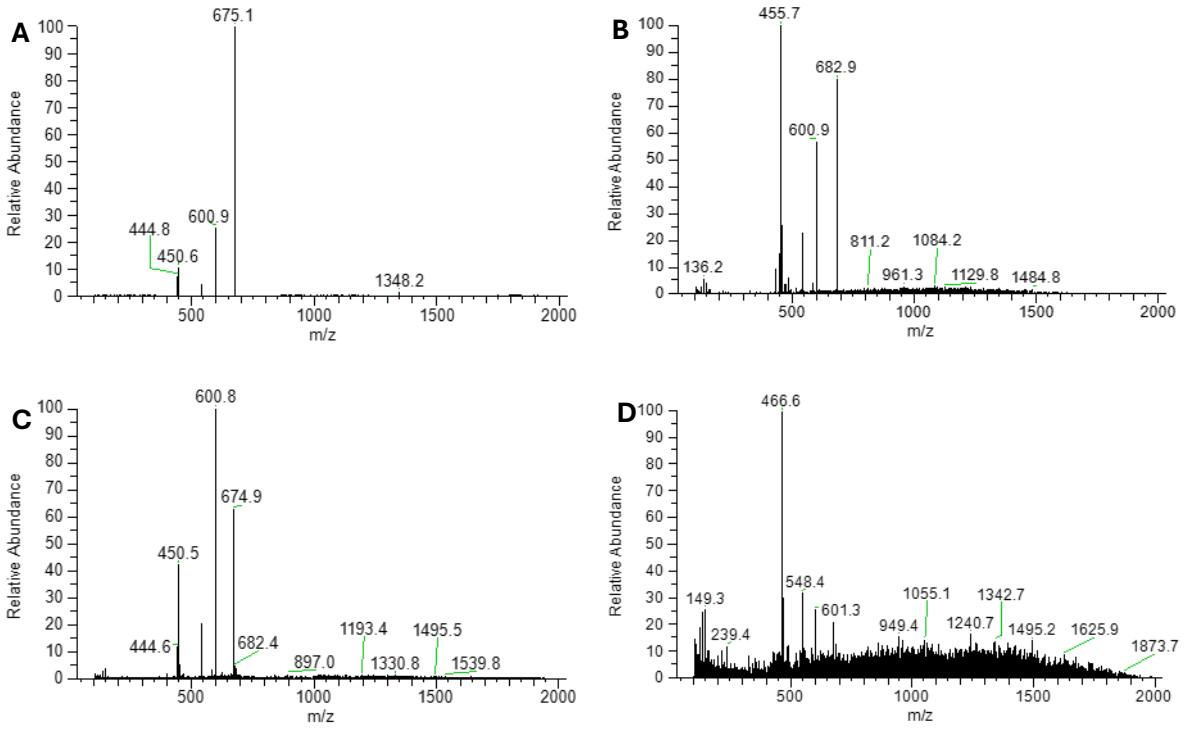


Figure 12.22: Mass spectra of the major peaks of substance P after elution through (A) a column packed with ethanolamine end-capped propylamino silica resin and (B-D) an immobilised pepsin reactor

in order of retention time as outlined in Figure 9.25