

Imperial College London

Department of Chemical Engineering

**Fundamentals of Peptide Adsorption
in Reversed-Phase Liquid Chromatography**

Thesis presented by

Oscar Mario Mercado Valenzo

in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of

Dr Daryl R Williams

24th October 2024

Statement of Originality

I declare that the work presented in this document is my own. Material derived from the work of others has been appropriately acknowledged.

Copyright declaration

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

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 and do not use it, or any derivative works, for a commercial purpose.

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.

Abstract

This Thesis investigates the fundamentals of peptide adsorption in reversed-phase liquid chromatography (RPLC), focusing on improving peptide purification strategies, which are crucial for research and industrial applications. Peptides, short chains of amino acids, are increasingly significant in pharmaceutical research and drug development due to their therapeutic potential. The main challenge in peptide purification is achieving high purity, which is essential for the efficacy and safety of peptide-based drugs.

The research explores the key parameters influencing peptide retention and separation in RPLC, including hydrophobicity, solvent composition, column characteristics, and mobile phase interactions. It also introduces the development of the SMART (Systematic Multi-parametric Adaptive Responsive and Tailored) algorithm, which optimises elution profiles to enhance separation efficiency, reduce solvent usage, and increase throughput. This algorithm was tested on various peptides, including glucagon, demonstrating its potential for analytical and industrial applications.

A comprehensive study of different RPLC column packing materials was conducted, with silica-based columns being the focus. The Thesis explores the limitations of current RPLC methods, particularly the difficulty in separating closely related peptide impurities. To address this, it proposes alternative column selection strategies and optimised elution conditions to improve the resolution of these impurities.

Overall, this Thesis contributes to advancing peptide purification techniques by improving theoretical understanding and offering practical methodologies. The results have direct implications for the pharmaceutical industry, where efficient, scalable, and sustainable purification processes are critical for producing high-purity peptide therapeutics. The research lays a foundation for future innovations, particularly in optimising chromatographic methods for more complex peptide systems and improving sustainability in large-scale manufacturing processes.

Acknowledgements

I want to express my sincere gratitude to my supervisor, Professor Daryl Williams, for his guidance and support in both professional and personal aspects throughout my PhD experience. His feedback and contributions to this project were remarkable in completing this Thesis.

I wish to acknowledge the financial support from the Engineering and Physical Sciences Research Council (EPSRC) and Eli Lilly and Company through the Pharma-SEL Prosperity Partnership (grant numbers EP/T005556/1 and EP/T518207/1). Without their funding, this research project would not have been possible. Similarly, I would like to give major thanks to the Department of Chemical Engineering for providing me with a Departmental Scholarship. Without their support, I would not have had this opportunity.

In addition, I want to express my warmest appreciation to the invaluable people I met during this journey, especially those who supported me emotionally and mentally during my PhD experience. Thank you, Ana Lopez, Ben Kistnah, Mum and Dad.

List of publications

- Al Musaimi, O.; Mercado Valenzo, O. M.; Williams, D. *Prediction of Peptides Retention Behaviour in Reversed-Phase Liquid Chromatography Based on their Hydrophobicity*. Journal of Separation Science 2022, 46. DOI: 10.1002/jssc.202200743.
- Al Musaimi, O.; Mercado Valenzo, O. M.; Williams, D. *Factors Influencing the Prediction Accuracy of Model Peptides in Reversed-Phase Liquid Chromatography*. Separations 2023, 10. DOI: 10.3390/separations10020081.
- Al Musaimi, O.; Ng, K. W.; Gavva, V.; Mercado Valenzo, O. M.; Banu Haroon, H.; Williams, D. *Elastin-Derived Peptide-Based Hydrogel: A Potential for Drug Delivery Systems*. Gels 2024, 10. DOI: 10.3390/gels10080531.

Manuscripts in preparation

- Mercado Valenzo, O.M.; Williams, D. *Chromatographic Statistical Analysis of Parameters Influencing Peptide Retention in Reversed-Phase High-Performance Liquid Chromatography*.
- Mercado-Valenzo, O.M.; Peng, J.; Sun, Y.; Williams, D. *Characterisation of Peptide Adsorption Mechanisms in Reversed-Phase High-Performance Liquid Chromatography*.
- Mercado-Valenzo, O.M.; Peng, J.; Sun, Y.; Williams, D. *Comparison of Adsorption Isotherm Tools for Small Molecules*.
- Mercado-Valenzo, O.M.; Williams, D. *Use of Pre- and Post-Column Detection to Improve Frontal Analysis Measurements for Adsorption Isotherm Determination*.
- Mercado Valenzo, O.M.; Williams, D. *Method Optimisation Using a SMART Algorithm for Multi-isocratic Elution for RPLC Peptide Purification*.
- Mercado Valenzo, O.M.; Williams, D. *Comparison of Surface Excess Isotherm of Commercial RPLC Columns*.
- Mercado-Valenzo, O.M.; Thangaraj, S.; Racaitis, I.; Williams, D. *Characterisation of Polyethylene Fibres as a Reference Hydrophobic Material for Chromatographic Separation of Peptides*.

List of Oral and Poster Presentations

- Mercado-Valenzo, O.M.; Thangaraj, S.; Racaitis, I.; Williams, D. *Characterisation of Polyethylene Fibres as a Reference Hydrophobic Material for Chromatographic Purification of Peptides*. 2022 – International Symposium of Chromatography. Budapest, Hungary. 2024 – TIDES USA: Oligonucleotide and Peptide Therapeutics. Boston, USA.
- Mercado-Valenzo, O.M.; Peng, J.; Sun, Y.; Williams, D. *Measurement of Peptide Adsorption Isotherms on Reversed-Phase Columns*. 2023 – HPLC. Dusseldorf, Germany.
- Mercado-Valenzo, O.M.; Peng, J.; Sun, Y.; Williams, D. *Adsorption Mechanism of Biopharmaceuticals (Peptides) onto Commercial RPLC Columns*. 2023 – American Institute of Chemical Engineers (AIChE) Meeting. Orlando, USA.

Table of Contents

Chapter 1 - Introduction.....	21
1.1 Importance of peptides.....	21
1.1.1 Current use of Peptides	21
1.1.2 Economics.....	22
1.2 Overview of peptide purification	23
1.2.1 Advantages of RPLC for peptide purification	23
1.2.2 Limitations of RPLC	25
1.2.3 Environmental footprint of RPLC.....	26
1.3 Motivation for the study	28
1.4 Research aims and objectives	28
1.4.1 Specific objectives	29
1.5 Structure of the Thesis.....	30
Chapter 2 - Peptides.....	32
2.1 Properties.....	32
2.1.1 Chain Length.....	32
2.1.2 Structure	34
2.1.3 Hydrophobicity.....	35
2.2 Peptide functions	37
2.3 Formation.....	38
2.3.1 Natural synthesis	38
2.3.2 Chemical synthesis.....	38
2.4 Peptides on RPLC separation	41
Chapter 3 - RPLC resins	43
3.1 Silica architecture	43

3.1.1	Classification.....	45
3.2	Silica resin properties.....	46
3.2.1	Microscale properties	47
3.2.2	Macroscale properties of silica resins	49
3.2.3	Silanol activity.....	50
3.2.4	Column hydrophobicity	53
3.3	Column Selection Systems.....	54
3.3.1	Quantitative Structure-Retention Relationships in Column Selection Systems ...	57
3.4	Using peptides as column descriptors.....	58
3.4.1	Peptides in CSS	59
Chapter 4 - Fundamentals of Liquid Chromatography		61
4.1	LC system anatomy.....	61
4.1.1	Volumes in the LC system.....	63
4.2	Mobile Phase Elution Profiles.....	73
4.2.1	Isocratic.....	73
4.2.2	Gradient Elution Profiles	74
4.2.3	Complex.....	75
4.3	Retention parameters.....	76
4.3.1	Retention factor k.....	76
4.3.2	Selectivity factor	77
4.3.3	Number of theoretical plates	78
4.3.4	Resolution.....	79
4.3.5	Symmetry factor.....	80
4.4	Modes/Types of Liquid Chromatography	81
4.5	Adsorption and Partition Processes on RPLC.....	82
4.5.1	Isotherms.....	82
4.5.2	Inverse chromatographic methods to measure isotherms	87

4.6	Other interactions in RPLC.....	95
4.6.1	Excess isotherms	97
Chapter 5 - Chromatography at the manufacturing scale		103
5.1	Definitions.....	103
5.1.1	Purity, yield and throughput.....	103
5.1.2	Analytical and preparative scale	104
5.1.3	Linear and nonlinear chromatography.....	105
5.1.4	Ideal and nonideal chromatography.....	106
5.2	Scaling Up RPLC for Peptide Purification	107
5.2.1	Differences Between Lab-Scale and Industrial-Scale RPLC	108
5.2.2	Theoretical Framework for Scaling.....	108
5.2.3	Impact on Gradient Optimization for Large-Scale Systems	109
5.2.4	Pressure and Temperature Considerations in Scale-Up	110
5.2.5	Importance of particle size considerations	110
5.3	Peak Shaving.....	111
Chapter 6 - Materials and Methods		113
6.1	Materials.....	113
6.1.1	General.....	113
6.1.2	Peptides.....	113
6.2	LC instruments.....	117
6.2.1	Analytical scale.....	117
6.2.2	Preparative scale	119
6.2.3	Columns.....	119
6.3	Methods	121
6.3.1	Tanaka method.....	121
6.3.2	Design of experiments (DoE)	124
6.3.3	Frontal analysis for small molecules.....	125

6.3.4	Peak maxima for peptides and small molecules.....	126
6.3.5	Excess isotherm and VM measurement using MDM.....	128
6.3.6	SMART method for peptide separation	129
Chapter 7 - Initial experiments.....		131
7.1	Column suitability and selection.....	131
7.2	Mapping of chromatographic behaviour of model peptides at different conditions.	137
7.2.1	Effect of peptide sequence and organic content.....	137
7.2.2	Effect of flow rate.....	138
7.2.3	Effect of ion-pairing agent.....	139
7.2.4	Effect of temperature	140
7.2.5	Principal Component Analysis	141
7.3	Conclusions.....	143
Chapter 8 - RPLC Adsorption of Small Molecules		145
8.1	Determination of V_M and τ	145
8.2	Breakthrough curve and adsorption isotherm by Frontal Analysis.....	148
8.3	Adsorption isotherm by Peak Maximum	153
8.4	Comparison between isotherm estimation methods.....	157
8.5	Conclusions.....	159
Chapter 9 - On Peptide Adsorption.....		160
9.1	Peptide chromatographic retention during isocratic elution	160
9.2	Isotherm behaviour under adsorption conditions.....	163
9.2.1	Adsorption behaviour and modelling.....	169
9.2.2	Proposed peptide adsorption behaviour	174
9.3	Isotherm behaviour under elution conditions.....	176
9.4	Conclusion.....	178
Chapter 10 - SMART algorithm		180
10.1	Retention theory.....	180

10.1.1	Isocratic elution.....	181
10.1.2	Gradient elution	182
10.2	Standard peptides – proof of concept	183
10.3	Glucagon – analytical scale RPLC separation	188
10.4	Glucagon – scaling up	197
10.5	Conclusions.....	202
Chapter 11	- Liquid-solid surface adsorption in RPLC	204
11.1	Effect of adsorbent surface ligand chemistry	204
11.2	Effect of particle size	212
11.3	Effect of organic solvent	217
11.4	Effect of column length	231
11.5	Effect of ion-pairing modifier	232
11.6	Use of excess isotherm as a column characterisation system	232
11.7	Conclusions.....	233
Chapter 12	- Final remarks and future work.....	236
12.1	Final remarks.....	236
12.2	Future Work.....	238
Chapter 13	- References	240
Appendix I	- Peptide hydrophobicity estimation: methods and scales	251
Appendix II	- Derivation of the Peak Maxima (PM) method.....	260
Appendix III	- Derivation of excess isotherm equations.....	262
Appendix IV	- Peptides under elution conditions	266
Appendix V	- Characterisation of Polyethylene Fibres as a Reference Hydrophobic Material for Chromatographic Separation of Peptides	268

List of Figures

Figure 1.1 Components of a Reversed-Phase Liquid Chromatography mobile phase.....	24
Figure 2.1 Euler diagram of natural amino acids.....	33
Figure 2.2 The peptide bond.	34
Figure 2.3 The hydrophobic effect.....	36
Figure 2.4 Hydrophobic cores in globular proteins	36
Figure 2.5 Bio-specificity and manufacturing costs relationship of pharmaceutical products.	37
Figure 2.6 General purity requirements for peptides.	40
Figure 3.1 Architecture of silicon dioxide (silica) materials.....	43
Figure 3.2 Heterogenous surface of a silica resin	44
Figure 3.3 Commonly used RPLC ligands/phases	45
Figure 3.4 Morphologies of chromatographic silica resins	46
Figure 3.5 Surface silanol groups present in RPLC resins	51
Figure 3.6 Classical techniques for RPLC resin characterisation.....	53
Figure 3.7 Graphical results of the Tanaka method	56
Figure 4.1 General schematic diagram of a modular LC system	63
Figure 4.2 Graphical summary of volumes in an LC system	64
Figure 4.3 Graphical calculation of dwell volume.....	66
Figure 4.4 Mobile phase elution profiles used in chromatography	73
Figure 4.5 Graphical summary of chromatographic retention parameters.....	78
Figure 4.6 Effect of the adsorption isotherm shape on chromatographic peak shape.....	84
Figure 4.7 IUPAC classification of gas phase adsorption isotherms	85
Figure 4.8 Experimental set-up for frontal analysis	89
Figure 4.9 Detector response to frontal analysis experiments.....	90
Figure 4.10 Representation of an infinitesimal chromatographic band migrating through a column.....	91
Figure 4.11 Experimental set-up for peak maxima	92
Figure 4.12 Graphical summary of isotherm determination through peak maxima experiments	93
Figure 4.13 Multifaceted interactions during peptide separation through RPLC.	97
Figure 4.14 Gibbs excess adsorption theory.....	98

Figure 4.15 Effect of small solvent injection into a fully equilibrated chromatographic column.....	99
Figure 4.16 Retention volume of perturbation peaks for different volume percentages of organic solvent in different RPLC columns.....	100
Figure 4.17 Excess adsorption isotherms of an organic modifier onto RPLC columns.....	102
Figure 5.1 Ternary diagram of chromatographic separation performance.....	104
Figure 5.2 Analytical and preparative scale method comparison	112
Figure 6.1 Full structure of Krokhin peptide standards.....	115
Figure 6.2 Secondary structure of glucagon(1-29)	116
Figure 6.3 Schematic of analytical HPLC used for experimentation.....	117
Figure 6.4 Back pressure test of pure DI water into columns in Table 6.4.	121
Figure 6.5 Chemical structure of the probes used in the Tanaka method.....	123
Figure 6.6 Diagram of the SMART algorithm. See text for description.....	129
Figure 7.1 Visualisation of the Tanaka method results for RPLC columns	133
Figure 7.2 Chromatograms from Run 1 of the Tanaka method.....	134
Figure 7.3 Chromatograms from Run 2 of the Tanaka method.....	134
Figure 7.4 Chromatograms from Run 3 of the Tanaka method.....	135
Figure 7.5 Chromatograms from Run 4 of the Tanaka method.....	135
Figure 7.6 Expansion to the Tanaka method for α_{CH_2} values.....	136
Figure 7.7 Effect of buffer concentration on retention.	139
Figure 7.8 UV signal masking effect caused by salt buffers.....	140
Figure 7.9 Principal Component Analysis of chromatographic response under DOE	143
Figure 8.1 Estimation of V_M values in ACN/H ₂ O mixtures for 150-mm length RPLC columns.....	147
Figure 8.2 Excess isotherm of ACN/H ₂ O mobile phase onto 150-mm length RPLC columns.....	148
Figure 8.3 Breakthrough curve of caffeine for C18 column	149
Figure 8.4 Breakthrough curve of caffeine for C8 column.....	150
Figure 8.5 Breakthrough curve of caffeine for Phenyl-Hexyl column.....	150
Figure 8.6 Adsorption isotherms of caffeine for RPLC column by FA method	151
Figure 8.7 Example of FA data adjustment based on pre- and post- column detection	152
Figure 8.8 Peak signal response for different injection volumes.....	154
Figure 8.9 Peak signal response for different sample concentrations.....	155

Figure 8.10 Peak signal response of constant injection mass at different volumes.	155
Figure 8.11 Adsorption isotherms computed using ECP for Luna C18 at different injection volumes.....	156
Figure 8.12 Comparison of caffeine adsorption isotherms calculated using PM and FA	157
Figure 9.1 Retention volumes of peptides for Luna C18(2) column.....	161
Figure 9.2 Minor disturbance/perturbation method for 150-mm RPLC columns	163
Figure 9.3 Adsorption isotherm of peptide P1 for RPLC columns	164
Figure 9.4 Adsorption isotherm of peptide P2 for RPLC columns	165
Figure 9.5 Adsorption isotherm of peptide P3 for RPLC columns.	165
Figure 9.6 Adsorption isotherm of peptide P4 for RPLC columns.	166
Figure 9.7 Adsorption isotherm of peptide P5 for RPLC columns	166
Figure 9.8 Adsorption isotherm of peptide P6 for RPLC columns	167
Figure 9.9 Peak maxima (PM) method chromatograms.....	170
Figure 9.10 Excess adsorption isotherm of ACN/H ₂ O + 0.1%TFA onto RPLC columns.	172
Figure 9.11 Total mass of ACN/H ₂ O + 0.1%TFA eluent adsorbed onto RPLC with interfaces of theoretical finite thickness τ	173
Figure 9.12 Summary of adsorption behaviour of standard peptides on RPLC columns	175
Figure 9.13 Chromatograms of standard peptides at eluting conditions.....	177
Figure 10.1 Gradient elution order of P2, P3, and P4.....	184
Figure 10.2 Isocratic elution for P2, P3, and P4.....	185
Figure 10.3 Multi-isocratic run for P2, P3, and P4.....	187
Figure 10.4 Gradient elution of crude glucagon	189
Figure 10.5 Isocratic experiments for glucagon	190
Figure 10.6 LSS model applied to main peak and critical peaks in a Luna C18 column.....	192
Figure 10.7 LSS model applied to main peak and critical peaks in a Luna C8 column.	193
Figure 10.8 LSS model applied to main peak and critical peaks in a Luna Phenyl-Hexyl column.	194
Figure 10.9 Final method design for glucagon in three analytical columns.....	195
Figure 10.10 Comparative chromatograms of crude glucagon gradient purification	196
Figure 10.11 Comparative chromatograms of crude glucagon SMART purification	197
Figure 10.12 Analytical versus semi-prep columns in analytical LC system	198

Figure 10.13 Analytical versus semi-prep columns in analytical LC system	199
Figure 10.14 Semi-preparative and analytical LCs comparison.....	201
Figure 11.1 Migration of acetonitrile and water molecules in an RPLC system	205
Figure 11.2 Migration of pure acetonitrile and water molecules in an RPLC system.....	207
Figure 11.3 Retention volume using the minor disturbance method in 5 μm columns using acetonitrile (ACN)	207
Figure 11.4 Excess adsorption isotherms of acetonitrile (ACN) in water on 5 μm columns.....	209
Figure 11.5 Total water/acetonitrile eluent mass adsorbed onto the 5 μm columns	211
Figure 11.6 Determination of V_M using the minor disturbance method in 10- μm columns using acetonitrile (ACN)	213
Figure 11.7 Determination of V_M using the minor disturbance method in 10 μm PREP columns using acetonitrile (ACN)	213
Figure 11.8 Excess adsorption isotherms of acetonitrile (ACN) in water on 10 μm columns.....	214
Figure 11.9 Excess adsorption isotherms of acetonitrile (ACN) in water on 10 μm PREP columns.....	214
Figure 11.10 Total water/acetonitrile eluent mass adsorbed onto the 10 μm columns.....	216
Figure 11.11 Total water/acetonitrile eluent mass adsorbed onto the 10 μm PREP columns.	217
Figure 11.12 Determination of V_M using the minor disturbance method in 5 μm columns using methanol (MeOH)	222
Figure 11.13 Determination of V_M using the minor disturbance method in 10 μm columns using methanol (MeOH)	223
Figure 11.14 Determination of V_M using the minor disturbance method in 10 μm PREP columns using methanol (MeOH)	223
Figure 11.15 Determination of V_M using the minor disturbance method in 5 μm columns using ethanol (EtOH)	224
Figure 11.16 Determination of V_M using the minor disturbance method in 10 μm columns using ethanol (EtOH)	224
Figure 11.17 Determination of V_M using the minor disturbance method in 10 μm PREP columns using ethanol (EtOH)	225

Figure 11.18 Excess adsorption isotherms of methanol (MeOH) in water on 5 μm columns.....	225
Figure 11.19 Excess adsorption isotherms of methanol (MeOH) in water on 10 μm columns.....	226
Figure 11.20 Excess adsorption isotherms of methanol (MeOH) in water on 10 μm PREP columns.....	226
Figure 11.21 Excess adsorption isotherms of ethanol (EtOH) in water on 5 μm columns.....	227
Figure 11.22 Excess adsorption isotherms of ethanol (EtOH) in water on 10 μm columns.....	227
Figure 11.23 Excess adsorption isotherms of ethanol (EtOH) in water on 10 μm columns.....	228
Figure 11.24 Total water/methanol eluent mass adsorbed onto the 5 μm columns.....	228
Figure 11.25 Total water/methanol eluent mass adsorbed onto the 10 μm columns.....	229
Figure 11.26 Total water/methanol eluent mass adsorbed onto the 10 μm PREP columns.....	229
Figure 11.27 Total water/ethanol eluent mass adsorbed onto the 5 μm columns.....	230
Figure 11.28 Total water/ethanol eluent mass adsorbed onto the 10 μm columns.....	230
Figure 11.29 Total water/ethanol eluent mass adsorbed onto the 10 μm PREP columns.....	231

List of Tables

Table 2.1 Impurity classes from SPPS procedures	40
Table 2.2 Effect of peptide properties in RPLC separation	42
Table 4.1 Proposed unretained molecules for liquid chromatography.....	71
Table 4.2 Comparative of gradient and isocratic elution methods	74
Table 4.3 Summary of types of liquid chromatography	81
Table 4.4 Summary of adsorption isotherm models.....	86
Table 4.5 Parameters involved in peptide separation through RPLC.....	96
Table 5.1 Summary of differences between chromatographic scales.....	107
Table 6.1 Characteristics of peptide standards.....	114
Table 6.2 Extra-column and dwell volumes of analytical LC.....	118
Table 6.3 Properties of columns used in Chapters 7-10.....	120
Table 6.4 Properties of columns used in Chapter 11.....	120
Table 6.5 Tanaka test conditions.....	122
Table 6.6 Factors used for the design of experiments in Chapter 7.....	124
Table 7.1 MANOVA analysis of chromatographic response under DoE.....	137
Table 8.1 Langmuir isotherm parameters for caffeine adsorption onto RPLC columns.	158
Table 9.1 Langmuir parameters for standard peptide isotherms.	171
Table 10.1 Gradient analysis for P1, P2, and P3	186
Table 10.2 Purification analysis of crude glucagon reaction mixture in C18 analytical column.	196
Table 11.1 ACN/H ₂ O excess isotherm characteristics for 5 µm RPLC columns.....	210
Table 11.2 Surface properties of 5 µm RPLC columns with an ACN/H ₂ O adsorption interface layer.....	211
Table 11.3 ACN/H ₂ O excess isotherm characteristics for 10-µm RPLC columns.....	215
Table 11.4 Surface properties of 10-µm RPLC columns with an ACN/H ₂ O adsorption interface layer.	216
Table 11.5 Organic/H ₂ O excess isotherm characteristics for RPLC columns.	219
Table 11.6 Surface properties of RPLC columns with an organic/H ₂ O adsorption interface layer.	220

Nomenclature

%B or B% - Percentage of organic solvent/modifier in mobile phase [percentage]

C_{ads} - Maximum concentration of pure solvent B in the interphase layer [mg/mL]

C_e - Concentration of the analyte in the mobile phase [mg/L]

C_{max} - Maximum analyte concentration in frontal analysis [mg/L]

C_s - Solubility of adsorbate in the mobile phase [mg/L]

H_{e,2} - Height of the smaller peak/plateau [dimensionless]

h_{max} - Height of the peak in chromatographic data [dimensionless]

K_L - Langmuir adsorption constant [mL/μmol]

k_w - Theoretical retention factor in 100% aqueous phase [dimensionless]

m_{adsorbent} - Mass of stationary phase used in adsorption experiments [g]

q_e - Equilibrium adsorption capacity [mg/g]

q_i - Instantaneous adsorption capacity under specific conditions [μmol/mL]

q_m - Maximum adsorption capacity [μmol/mL]

R_s - Resolution between chromatographic peaks [dimensionless]

S_A - Surface area per volume of adsorbent [m²/mL]

t₅₀ - Midpoint of the gradient curve time [minutes]

t_d - Dwell time, gradient delay [minutes]

t_g - Duration of gradient elution [minutes]

t_M - Mobile phase elution time [minutes]

t_r - Retention time of the analyte [minutes]

V₀ - Dead volume [mL]

V_{ads} - Volume of adsorbed layer per unit surface area [mL/m²]

V_d - Dwell volume, delay in solvent mixing [mL]

V_{EC} - Extra-column volume [mL]

V_{inter} - Interparticle volume in packing material [mL]

V_{intra} - Intraparticle volume within stationary phase particles [mL]

V_{ligand} - Volume occupied by bonded ligands on stationary phase [mL]

V_M - Hold-up or mobile phase volume in the column [mL]

V_n - Net retention volume, excluding stationary phase [mL]

V_r - Retention volume [mL]

V_R - Retention volume adjusted for adsorbent conditions [mL]

V_s or V_{sp} - Volume of the stationary phase [mL]

$W_{0.05H}$ - Width of the peak at 5% of peak height [minutes or mL]

$W_{0.5H}$ - Width of the peak at half peak height [minutes or mL]

η_B - Moles of solvent B adsorbed onto surface [mol/m²]

$\phi_{B,0}$ - Initial organic solvent volume fraction [dimensionless]

ϕ_B - Volume fraction of organic solvent in the mobile phase [dimensionless]

$\Delta\phi_{B,g}$ - Change in organic solvent fraction during gradient phase [dimensionless]

$\Delta\phi_B$ - Change in solvent fraction during gradient [dimensionless]

F - Bulk liquid mobile phase flow rate [mL/min]

HETP - Height Equivalent to a Theoretical Plate [mm]

k or k' - Capacity or retention factor for isocratic conditions, equilibrium-based constant [dimensionless]

N - Number of theoretical plates, indicating column efficiency [dimensionless]

q - Instantaneous adsorption quantity [mg/g]

RMSE - Root Mean Square Error of model fit [dimensionless]

S - Symmetry factor or tailing factor, describing peak symmetry [dimensionless]

T - Temperature of the chromatography system [°C]

α - Selectivity factor, the relative separation between two compounds [dimensionless]

τ - Layer thickness in adsorption isotherm calculations [nm]

Abbreviations

ACN - Acetonitrile

C18 - Octadecylsilane, a common RPLC column type

CSPS - Classical Solution Phase Synthesis

DI water - Deionised Water

DMSO - Dimethyl Sulfoxide

DoE - Design of Experiments

ECV - Extra-Column Volume

EMA - European Medicines Agency

EPSRC - Engineering and Physical Sciences Research Council

EtOH - Ethanol

FDA - Food and Drug Administration

GLP-1 RAs - Glucagon-Like Peptide-1 Receptor Agonists

HETP - Height Equivalent to a Theoretical Plate

HPLC - High-Performance Liquid Chromatography

LC-MS - Liquid Chromatography-Mass Spectrometry

LPPS - Liquid-Phase Peptide Synthesis

MeOH - Methanol

MS/MS - Tandem Mass Spectrometry

PCA - Principal Component Analysis

PDA - Photodiode Array

QSRR - Quantitative Structure-Retention Relationship

RPLC - Reversed-Phase Liquid Chromatography

RT - Retention Time

SMART Algorithm - Systematic Multi-parametric Adaptive Responsive and Tailored Algorithm

SPPS - Solid-Phase Peptide Synthesis

TFA - Trifluoroacetic Acid

USP - United States Pharmacopeia

UV-Vis - Ultraviolet-Visible Spectroscopy

Chapter 1

Introduction

1.1 Importance of peptides

Peptides are short chains of amino acids linked by peptide bonds. They typically comprise 2 to 50 amino acids, whereas proteins are longer chains with typically 50 to 200 amino acids. Peptides are crucial in biological processes, acting as hormones, enzymes, and signalling molecules. They can be classified based on their origin (natural or synthetic), function (e.g., antimicrobial, signalling), or structure (linear or cyclic)[1].

Peptides have emerged as crucial components in the development of drugs for major medical conditions including diabetes, cancer, and infectious diseases. Their ability to interact with specific molecular targets with high affinity and selectivity underpins their therapeutic efficacy[2]. For instance, peptide-based drugs like insulin analogues and GLP-1 receptor agonists have revolutionised diabetes and obesity management, improving glycaemic control and patient outcomes.

1.1.1 Current use of Peptides

Peptides are fundamental in biomedical research as tools for both disease diagnostics and as therapeutic agents. They are also employed in the targeted delivery of nucleic acids, drugs, and imaging agents to specific cells and tissues, enhancing the precision and efficacy of treatments[3, 4]. These biomolecules are also used in developing vaccines and as biomarkers for disease detection, contributing to advancements in personalised medicine and early clinical diagnosis.

The growing importance of peptides in pharmaceutical and biological research is driven by their therapeutic potential, technological advancements, regulatory successes, sustainability, and diverse applications in biomedical research. Peptides are expected to play an increasingly critical role in addressing complex medical challenges and improving patient care. Between 2016 and

2022, 26 peptide drugs were approved by the FDA, reflecting the robust pipeline of peptide therapeutics in clinical and preclinical development[3]. This trend underscores the successful translation of peptide research into viable therapeutic options. Nevertheless, peptide-based drugs suffer from the inability to be delivered as oral therapeutics due to their vulnerability to enzymic degradation in the stomach as well as their poor solubility in the lipid layers in the gut. These limiting factors reduce their potential presence in the current pharmaceutical market.

Since the introduction of biopharmaceuticals in the early 1980s, their ability to target molecules with high specificity has made them essential for treating conditions such as metabolic, oncologic, and cardiovascular diseases [5]. As signalling molecules, peptides offer the potential to target complex physiological pathways with fewer side effects than traditional small-molecule drugs.

Technological innovations, such as solid-phase peptide synthesis (SPPS) and recombinant DNA technology, have revolutionised peptide production, creating complex and highly pure peptides. Novel delivery systems, including nanoparticle carriers, have improved the pharmacokinetic profiles of peptide drugs, facilitating their clinical application.

1.1.2 Economics

As the global population ages, a higher demand for treatments related to age-associated diseases exists. Peptides are increasingly used in therapies for conditions like osteoporosis and cardiovascular diseases[6, 7]. North America has the largest market value, followed by Europe and Asia-Pacific for these diseases. The Asia-Pacific region is expected to grow fastest due to increasing healthcare investments and increasing disease prevalence[8].

The global market for peptide therapeutics is valued at approximately \$43.5 billion for 2024 with projections suggesting it could reach \$68.7 billion by 2030[9-11]. The rapid take up of these peptide GLP-1 drugs for obesity control has led to predictions that peptide drug revenue could boost to the \$150 billion by the 2030s.

Peptide-based drugs are integral to the pharmaceutical industry. In 2019, 11 of the top 20 drugs by global sales were peptide-based, including peptides, proteins, antibodies, and fragment antigens[12]. It is important to mention that currently available market reports only estimate the

growth based on the market trend during 2018-2021. However, demand is expected to skyrocket due to newer and more popular uses of peptides and the release of intellectual property of the best-selling pharmaceutical peptides, opening the market for peptide generics.

A large surge of demand for peptides is expected in the upcoming years due to the use of peptides for weight loss, such as GLP-1 receptor agonists (GLP-1 RAs) (e.g., semaglutide and liraglutide), which have significantly expanded the market. These drugs are highly effective in reducing appetite and promoting weight loss. For instance, citizens of the United States of America spent over \$50 billion on weight loss drugs in 2023[13]. This surge in spending is influencing various sectors, including healthcare and pharmaceuticals. Additionally, the advent of generic peptides, such as Victoza [14], is set to increase the production of peptides, particularly GLP-1 RAs, to levels that have not yet been seen in the pharmaceutical industry.

1.2 Overview of peptide purification

Peptide purification is a critical step in both research and industrial contexts to ensure that peptides meet purity standards for their intended applications. The objectives of peptide purification include achieving high purity, maintaining biological activity, ensuring scalability, elimination of related chemical impurities and promoting sustainability.

High purity is crucial, as impurities can affect a peptides' biological activity and toxicity. In pharmaceutical applications, regulatory bodies like the FDA enforce stringent purity standards to ensure the safety and efficacy of peptide-based drugs[15]. Techniques such as high-performance liquid chromatography (HPLC) are commonly used to rigorously establish these purity levels.

1.2.1 Advantages of RPLC for peptide purification

The most widely used techniques for peptide purification is Reverse-Phase Liquid Chromatography (RPLC). RPLC is favoured for its ability to separate peptides based on their hydrophobicity, which is a key feature of many peptide sequences. The stationary phase in RPLC typically consists of hydrophobic alkyl chains, most commonly C4, C8, and C18 chain lengths, chemically bonded onto porous silica particles. Additional features can be added to these alkyl ligand

species to increase the interactions between the stationary phase and the analyte in the mobile phase to be separated. This phenomenon is explained further in Chapter 3. In RPLC, the mobile phase is composed of a gradient mixture of water and an organic solvent like acetonitrile or methanol, often including an acid modifier/ion pairing agent such as trifluoroacetic acid (TFA). Some typical examples of components of the mobile phase are shown in Figure 1.1. Peptides adsorb on the hydrophobic stationary phase, and their elution is controlled by increasing the organic solvent concentration, allowing peptides with varying hydrophobicity to be separated[16]. This method is highly effective in resolving peptides with only slight differences in their structure and length[17, 18].

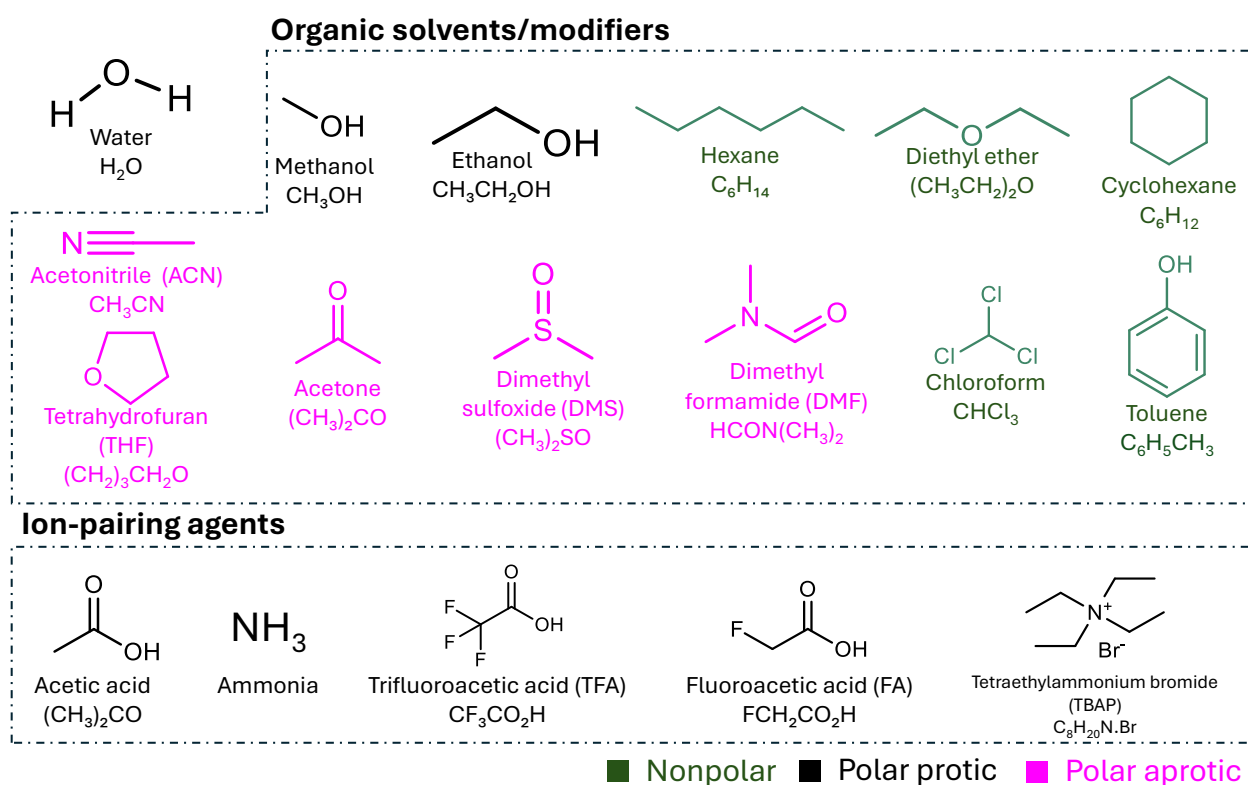


Figure 1.1 | Components of a Reversed-Phase Liquid Chromatography mobile phase.

The scalability of RPLC processes is a major advantage, allowing RPLC to be used across various stages of peptide production, from analytical research to large-scale industrial purification. The technique is also highly reproducible, ensuring that it can consistently deliver peptides of the required purity for both research and pharmaceutical applications[16]. Additionally, RPLC's

compatibility with mass spectrometry (LC-MS) enhances its utility by enabling real-time peptide purity monitoring, which is crucial in both development and regulatory settings.

In pharmaceutical contexts, regulatory bodies such as the FDA and EMA require comprehensive characterisation of peptide drugs, including their purity profiles[19]. RPLC, with its high resolution and ability to remove critical impurities, plays a central role in meeting these regulatory standards. It is especially essential for ensuring the removal of unwanted reaction byproducts like truncated sequences, racemised isomers, and other impurities that may affect the safety or efficacy of peptide drugs.

1.2.2 Limitations of RPLC

Although RPLC has been used for over five decades and is the technique of first choice for both analytical and industrial peptide purification, it faces notable limitations when dealing with closely related impurities arising from peptide synthesis, especially those generated during SPPS. These impurities, which may include truncated sequences, deletion sequences, or side products from incomplete reactions, often possess similar hydrophobic characteristics to the target peptide, making their separation via RPLC particularly challenging. This issue leads to compromised purity and can necessitate multiple RPLC purification stages, increasing both manufacturing cost and time.

Another time-consuming task is the selection of a suitable column for each RPLC separation. Column selection is a lengthy and complex process due to the vast array of commercially available columns, each with different properties and performance characteristics. To put this point into perspective, according to the United States Pharmacopoeia, by October 2024, there are around 1138 different octadecyl (C18) silane chemically bonded columns (USP code L1) commercially available in the market[20]. Therefore, RPLC column screening process involves evaluating multiple columns to identify the one that provides the best separation for a specific peptide mixture. Column material, particle size, pore size, and surface chemistry are crucial factors for separation efficiency and resolution.

Thus, extensive screening processes are necessary to ensure optimal performance but can be time-consuming and resource-intensive. Each column must be tested under a wide range of

processing conditions to determine its suitability, which involves running numerous experiments and analysing the results. This process takes significant time and requires substantial amounts of solvents and reagents, further increasing costs.

To streamline this process, some manufacturers use high-throughput screening methods and advanced analytical techniques to evaluate multiple columns quickly. Additionally, leveraging computational tools and predictive models can help narrow the column choices before physical testing occurs, saving time and resources.

The scale-up process for RPLC separation can also be both time-consuming and costly. The process requires meticulous separation of each component in the peptide mixture, which becomes increasingly complex with larger and more diverse reaction mixtures. This extended processing time can slow down overall production rates and timeframes. Additionally, the high volume of organic solvents needed for large-scale operations significantly increases costs. These solvents are expensive and require efficient recovery and recycling systems to mitigate environmental impact. Furthermore, the need for high-purity reagents and advanced monitoring systems to maintain consistent processing conditions adds to the financial burden. Consequently, while RPLC offers high resolution and reproducibility, its implementation at a manufacturing scale demands substantial time and financial investment.

1.2.3 Environmental footprint of RPLC

Sustainability is becoming increasingly important in peptide purification. Recent developments in RPLC techniques focus on green chemistry principles, such as reducing solvent consumption and waste production, making peptide purification more environmentally friendly. More challenging approaches include using environmentally friendly solvents and optimising energy consumption [22]. These sustainable practices benefit the environment and meet regulatory requirements and public expectations.

Organic solvents, such as acetonitrile and methanol, are widely used in the mobile phase to facilitate peptide separation. However, these solvents are classified as volatile organic compounds (VOCs), contributing to air pollution if not properly contained. The production and disposal of solvents also carry a considerable carbon footprint, and their improper disposal can lead to soil

and water contamination. Acetonitrile, for instance, poses a particular risk due to its toxicity to aquatic ecosystems.

Although the quantitative environmental impact of RPLC is generally not reported. The following can be estimated based on the experimental work of this Thesis:

- An analytical/semi-preparative LC system generally requires around 2-5 kWh for a period of 24 hours, depending on the system capacity, i.e., analytical with a flow rate range of 0.1-10 mL/min and semi-preparative at 0.5-20 mL/min.
- For an analytical instrument, the most used scale in academic and some quality control labs, 1-4 samples can be processed each hour for 'standard' gradient peptide purification protocols of gradient slope of 4%B per column volume (CV).
- Based on the same gradient method, for a column with CV equivalent to around 2.5 mL (4.6 mm internal diameter and 150 mm length) for 30 min, around 17.4 and 49.7 mL per hour of organic solvent and water, respectively, is used per injection.
- If the initial experiments in Section 7.2 were to be conducted monthly during the four years of this PhD project. I would require around 0.7 and 3 metric tons of solvent and water, respectively.
- During my PhD, approximately 307.5 L of organic solvent were used: acetonitrile (70%), methanol (17%), and ethanol (13%). Water consumption was not directly quantified; however, the mobile phase discarded was approximately 1400 L (30 L per month).

While these values are surprisingly high, it is important to emphasise that all experiments were conducted on a small scale. Consequently, the solvent usage is not as impactful. Under manufacturing conditions, these same experiments would be scaled up approximately 500 times. Yet, they still do not reflect the amount of solvent and water used in current pharmaceutical production.

In addition, incorporating green chemistry into peptide synthesis will help to make production more environmentally friendly. Innovations in solvent selection, reaction conditions, and purification methods reduce the environmental impact of peptide manufacturing. As the demand for peptide-based therapies continues to grow, sustainable practices will be vital to minimising the ecological footprint of peptide production.

1.3 Motivation for the study

Traditionally, developing RPLC methods involves meticulously iterating different columns and adjusting the mobile phase gradient steepness or slope. This approach aims to achieve better resolution of closely related impurities. However, in most cases, two chromatographic process steps are required to obtain purity standards for peptides at the manufacturing scale. Additionally, this process consumes a significant amount of time and leads to considerable resource expenditure, making it an inefficient method in many cases.

Researchers and industrial chromatographers continually explore alternative strategies to address these inefficiencies. These include optimising single-step chromatographic methods, integrating advanced detection techniques, and employing novel stationary phases that offer better selectivity and resolution. However, a complete understanding of the fundamentals of the chromatographic phenomena during peptide adsorption is still missing.

This knowledge gap often impedes the development of more efficient and cost-effective purification processes. To bridge this gap, ongoing research focuses on gaining deeper insights into the interactions between peptides and the chromatographic system. This includes studying the peptide adsorption and the impact of various chromatographic parameters on separation efficiency. By enhancing our understanding of these fundamental processes, the goal is to develop more robust and streamlined methods that can achieve high purity standards with greater efficiency and reduced resource consumption.

1.4 Research aims and objectives

The aims and objectives of this project are contained within one task of the Pharma-SEL Prosperity project; task 3.2.1. This project is a large EPSRC funded research programme based at Imperial College London and University College London, starting in 2020. For task 3.2.1, the primary aim is to experimentally study the fundamentals of peptide sorption onto RPLC columns, considering factors such as solvent composition, ionic strength, peptide structure, hydrophobicity, and column properties.

This research aims to develop and optimise advanced peptide purification methods using RPLC, addressing key challenges in separation efficiency, resolution, and scalability. This research

seeks to improve purification outcomes by systematically optimising column selection, solvent composition, and gradient elution profiles, with a particular focus on overcoming the difficulties associated with closely related peptide impurities and scaling purification for industrial applications.

1.4.1 Specific objectives

1. *Investigate the relationship between peptide properties and adsorption behaviour in RPLC*
 - 1.1. Examine how peptide structure (e.g., chain length, secondary structure) and hydrophobicity influence adsorption to RPLC stationary phases.
 - 1.2. Identify how variations in peptide properties affect retention times and resolution, focusing on understanding the complexities of separating closely related peptide impurities.
2. *Optimise column selection based on peptide structure and surface chemistry interactions*
 - 2.1. Explore the impact of different RPLC column properties (e.g., pore size, particle size, and surface chemistry) on peptide separation efficiency
3. *Optimise gradient elution profiles and solvent composition for peptide purification*
 - 3.1. Investigate how varying solvent composition (e.g., organic solvent concentrations, pH modifiers) influences peptide retention, peak resolution, and run times.
 - 3.2. Develop optimised gradient elution strategies that maximise resolution for diverse peptide mixtures while balancing solvent use and efficiency.
4. *Scale peptide-purification methods for industrial applications, focusing on throughput and efficiency*
 - 4.1. Develop strategies to improve throughput, minimise run times, and reduce solvent consumption while maintaining product purity and yield at an industrial scale.
 - 4.2. Explore the cost-saving potential of optimised purification strategies for large-scale peptide production, focusing on reducing solvent and energy use without compromising separation quality.

By addressing these objectives, the project aimed to develop a deeper understanding of the chromatographic phenomena involved in peptide adsorption, ultimately leading to more efficient and effective purification processes.

1.5 Structure of the Thesis

This Thesis begins by exploring peptides, focusing on their defining characteristics and roles in biological processes. Chapter 2 highlights how factors such as amino acid length and hydrophobicity influence peptide stability and functionality, setting the stage for the synthesis and purification methods discussed later. This understanding is essential for the application of peptides in pharmaceutical manufacturing.

Moving forward, Chapter 3 transitions into the design and properties of chromatographic columns, emphasising silica's importance due to its mechanical stability. The discussion on column geometry, particle size, and ligand types directly links to the foundational separation techniques for efficient chromatography, particularly in RPLC.

Building on this, Chapter 4 presents a detailed overview of liquid chromatography systems, focusing on operational factors critical for optimising separation. The Chapter contrasts isocratic and gradient elution methods, focusing on how these techniques impact analyte separation in RPLC, setting the framework for more advanced discussions on industrial-scale chromatography.

The Thesis then scales these concepts up in Chapter 5, which examines chromatography at the manufacturing level. This Chapter bridges the gap between laboratory techniques and industrial applications, highlighting the challenges and optimisations required for scaling purification processes.

Chapter 6 outlines the materials and methods used in the experiments that validate the research findings. These methods support the proposed experiments in Chapter 7, which focus on determining key parameters for peptide separation, including column and chromatographic parameter selection.

Chapter 8 shifts to a comparative adsorption study using caffeine as a small molecule model. This study lays the groundwork for deeper research of peptide adsorption in Chapter 9, where the interactions between peptides and the chromatographic system are analysed. Here, factors like peptide hydrophobicity and ligand interactions are explored, which are crucial for optimising purification techniques.

In Chapter 10, the introduction of the SMART algorithm demonstrates significant improvements in separation efficiency and scalability, outperforming traditional gradient methods, for peptide purification. This development is critical for industrial applications and reinforces the earlier discussions on method scalability.

Chapter 11 provides an in-depth examination of liquid-solid adsorption mechanisms in RPLC. Understanding how surface interactions influence separation is key to improving chromatographic performance, tying back to the practical applications of peptide purification discussed throughout the Thesis.

This Thesis concludes with Chapter 12, which summarises the key findings and their implications for the field of peptide purification. This final Chapter ties together the contributions of this research and outlines a clear path for further innovations in peptide purification. It reflects on the advances in understanding peptide characteristics, adsorption behaviours, and optimising RPLC methods. The Chapter highlights how the SMART algorithm has enhanced separation efficiency and scalability, offering significant potential for industrial applications.

The final Chapter also proposes future work that could expand the applicability of the developed methods to more complex peptide systems and larger-scale manufacturing. It suggests refining the SMART algorithm to improve its adaptability and long-term stability and exploring new chromatographic materials and techniques that may further enhance the purification process. Additionally, the importance of continuing to investigate peptide adsorption mechanisms is emphasised to ensure ongoing improvements in chromatographic performance and drug development processes.

Chapter 2

Peptides

This Chapter focuses on peptides, detailing their defining characteristics, significance in biological processes, and various synthesis methods. It begins by emphasising the critical role of amino acid length in determining their functionality and interactions. It explores the concept of hydrophobicity, which significantly influences peptide folding and stability. Moreover, this Chapter compares natural peptide synthesis with chemical methods, explaining the advantages and limitations of each and highlighting their relevance in pharmaceutical manufacturing.

2.1 Properties

2.1.1 Chain Length

Peptides are short chains of amino acids, typically composed of between 2 and 40-50 residues, linked by (alpha) peptide bonds. The upper limit of a peptide size varies depending on its application and the definition from the relevant regulatory agency. The United States Food and Drug Administration (FDA) considers any amino acid chain composed of 40 or fewer amino acids to be a peptide[21]; anything longer than 40 amino acids (AAs) but less than 100 AAs in size classifies as a 'chemically synthesised polypeptide' instead of 'peptide' or 'protein'[22]. The European Medicines Agency has a minimum size of 5 amino acids for a synthetic molecule to be considered a peptide, and no precise upper limit is suggested[19].

The properties of peptides are significantly influenced by their chain length, which is determined by the number and the type of amino acid residues in their sequence. Each amino acid will contribute differently to the peptide properties depending on the residue's properties (Figure 2.1).

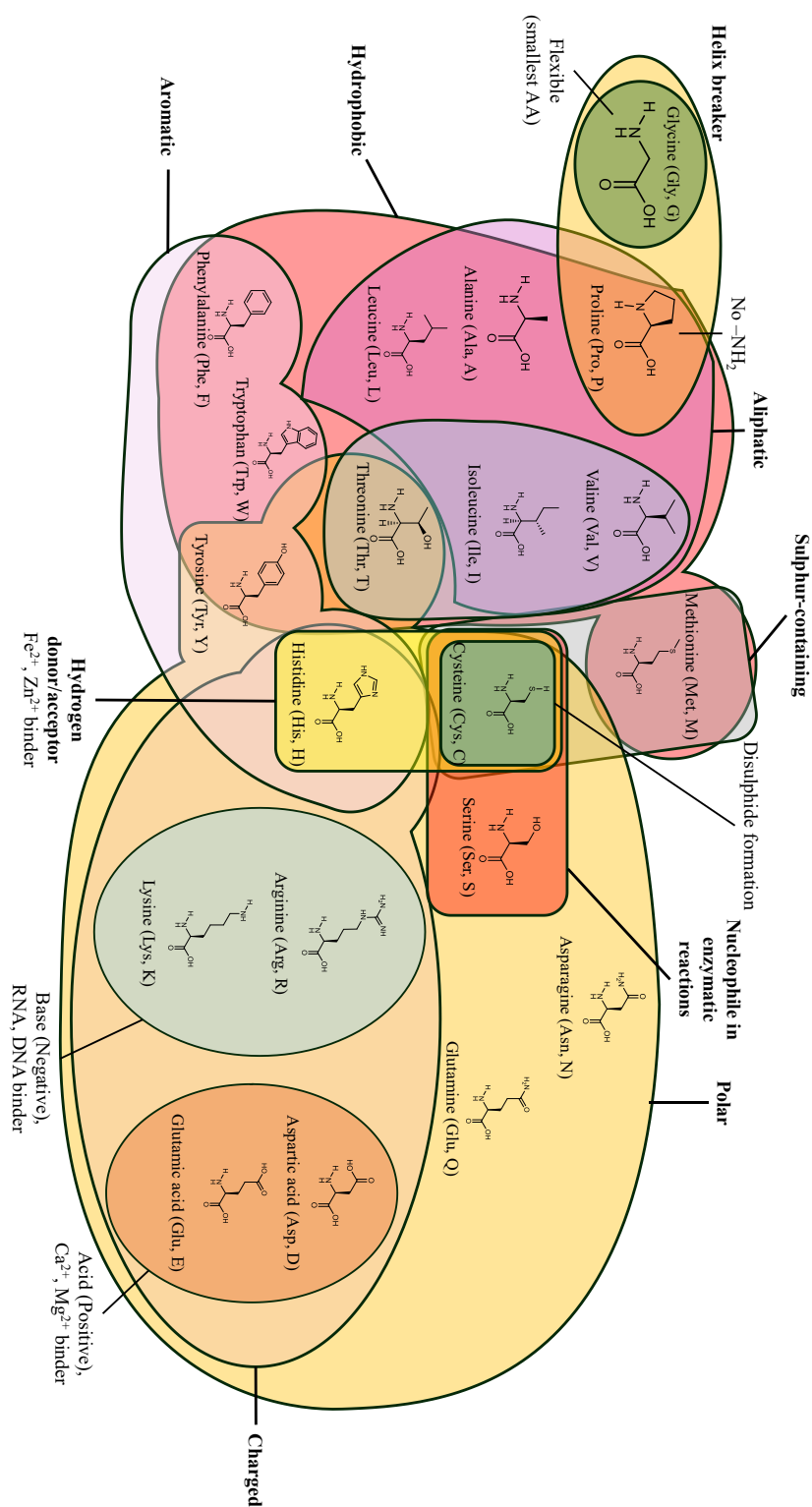


Figure 2.1 | Euler diagram of natural amino acids. Adapted from [23] with permission from the author, Professor Loren Williams.

The peptide bond is an amide linkage between the NH_2 and COOH groups of neighbouring amino acids that generally produces *trans* bonds (see Figure 2.2). Although the $\text{C}^\alpha\text{-C}$ bonds in amino acids are not rigid and can freely rotate, being only limited by the size and character of their R groups, the peptide bond is rigid and planar and not free to rotate as it has a partial double character[24].

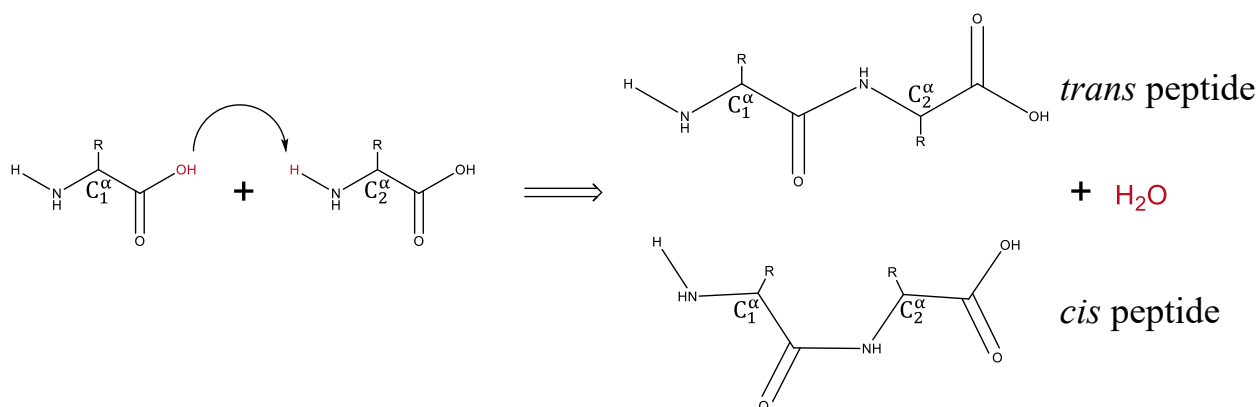


Figure 2.2 | The peptide bond.

2.1.2 Structure

As the peptide chain length increases, peptides can exhibit specific folding patterns critical for their functional roles in biological processes. Additionally, the length of a peptide can influence its interaction with receptors, enzymes, and other biomolecules, thereby modulating its physiological effects. As expected, a peptide's size and amino acid sequence play a crucial role in determining its physical, chemical, and biological properties, making the length a critical factor in peptide design and application.

Peptide structures are categorised into four levels: primary, secondary, tertiary, and quaternary. The primary structure is the specific sequence of amino acids in a peptide chain, linked by peptide bonds, which sets the foundation for the properties and functions of the peptide. Secondary structures, including α -helices and β -sheets, are stabilised by hydrogen bonds and contribute to the peptide's stability and shape[25]. α -Helices are right-handed coils formed by hydrogen bonds between amino acids. In contrast, β -sheets are formed by hydrogen bonds linking two or more segments of the peptide chain in parallel or antiparallel arrangements[26].

The tertiary structure is the overall three-dimensional shape of a single peptide chain, determined by interactions among the side chains of the amino acids, such as hydrophobic interactions, hydrogen bonds, ionic bonds, and disulfide bonds. This structure is crucial for the peptide's functional form, allowing specific interactions with other molecules. Quaternary structure arises when multiple peptide chains, or subunits, assemble into a functional protein complex, stabilised by similar interactions as the tertiary structure. This complex formation enables sophisticated functions and regulation, exemplified by many enzymes and structural proteins. Understanding these structural levels is particularly essential for comprehending how peptides perform their diverse biological roles and their applications in medicine and biotechnology. However, these structural properties also offer insight into potential adsorption mechanisms of peptides onto specific solid state adsorbents.

2.1.3 Hydrophobicity

Hydrophobicity is arguably the most investigated physicochemical property in peptide-based structures. It was noted early that the effect of this property is dominated by the initial chemical interactions during the folding of globular proteins, which is more important than other driving forces such as steric hindrance and electrostatic interactions. Thus, the peptide/protein folding process is commonly described through the hydrophobic effect, which describes that to achieve the most thermodynamically stable state in the system, water molecules around the peptide-based surface must be displaced (Figure 2.3)[27]. This enthalpic expulsion of disordered water results in the exclusion and compaction of hydrophobic residues into one (most peptides) or multiple cores (primarily proteins) within the global peptide-based structure. Examples of these hydrophobic cores in globular proteins can be seen in Figure 2.4. Next, the outside layer of water molecules forms a cage-like structure around decreasing the entropy and keeping the peptide-based in solution by stabilising the remaining non-polar residues at the surface.

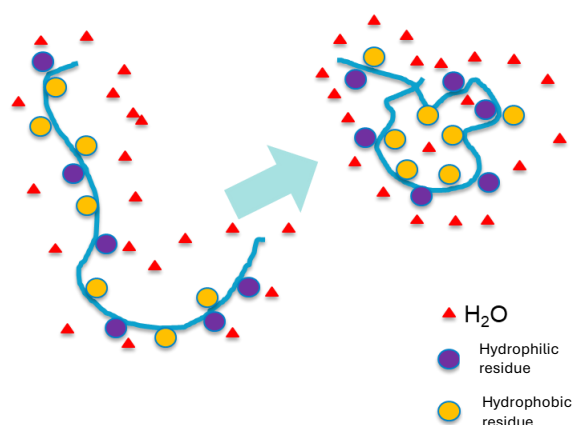


Figure 2.3 | The hydrophobic effect.

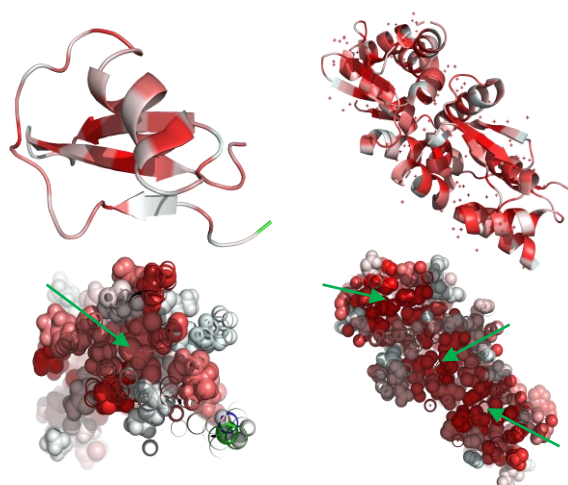


Figure 2.4 | Hydrophobic cores in globular proteins. Three-dimensional structures of a SWEET family transporter (left) and protein ZBTB7A (right). The red colour in both cartoons (top) and spheres (bottom) representations depicts regions of higher hydrophobicity according to the Eisenberg scale[28]. Hydrophobic cores are indicated with green arrows.

Peptide hydrophobicity is one of the key factors influencing separation efficiency in RPLC. Understanding the various methods and scales for estimating or describing hydrophobicity is fundamental for optimising RPLC conditions. However, no standard approach to peptide hydrophobicity determination is currently available. Instead, multiple methods/techniques have been proposed to determine hydrophobicity scales. Hydrophobicity scales are quantitative measures that rank amino acids or peptides based on their hydrophobicity or their tendency to avoid water and associate with non-polar environments.

Appendix I provides an in-depth historical perspective through selected examples of how hydrophobicity scales were developed, evolving from early experimental studies to modern computational models, which can be directly applied to improve predictions of peptide behaviour in RPLC. A more RPLC-focused literature review can be found in references [29] and [30].

2.2 Peptide functions

Peptide functions play crucial roles in various biological processes, acting as signalling molecules, hormones, and structural components of cells. Peptides can influence multiple physiological functions, including growth, metabolism, and immune responses[31]. Due to their specificity and versatility, peptides are extensively studied for their potential therapeutic applications in treating diseases and enhancing overall health. Their ability to interact precisely with biological targets makes them valuable medical research and drug development tools. In principle, peptides have proven therapeutic potential in medicine, but challenges, including low stability, poor membrane permeability and high patient costs, limit their applications[32]. From a manufacturing perspective, peptides are valuable as they sit in the middle regarding costs and bio-specificity, as the technologies required to produce peptides are similar as those for small molecules (Figure 2.5).

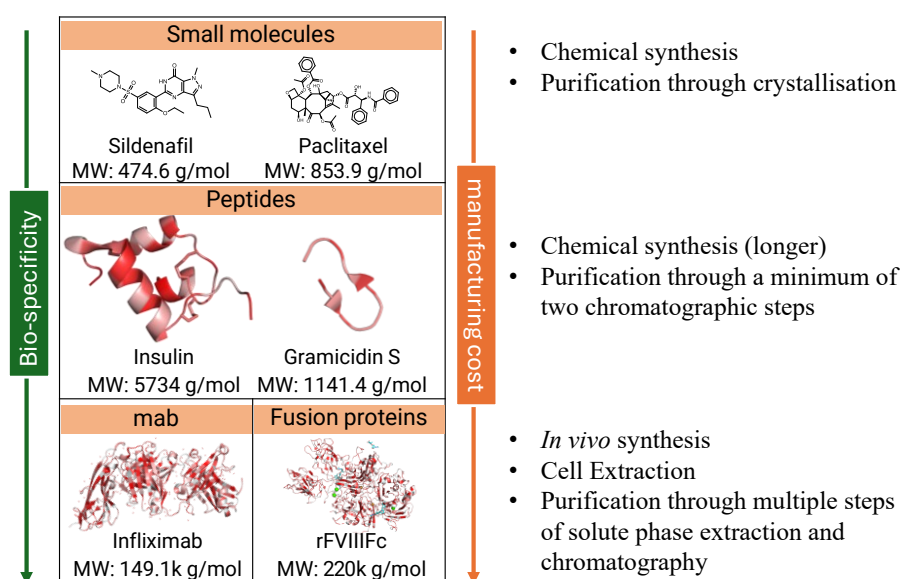


Figure 2.5 | Bio-specificity and manufacturing costs relationship of pharmaceutical products.

2.3 Formation

2.3.1 Natural synthesis

Natural synthesis, commonly called protein biosynthesis, is the process by which a cell makes proteins[33]. A mature messenger ribonucleic acid (mRNA) sequence binds to a ribosome within the cytoplasm of a cell. There are 2 transfer RNA (tRNA) binding sites in the ribosome: the P (peptidyl) and A (aminoacyl) sites. The tRNA molecules are adaptor molecules that carry a specific amino acid. Once the first amino acid, methionine, is attached to the P site, the following amino acid is positioned in the A site. The peptidyl transferase enzyme complex creates a peptide bond between the two amino acids, and the newly formed dipeptide is translocated to site P to restart the process. The synthesis direction is N- to the C- terminus, where the free amino group (N-terminus) of the first amino acid is left intact, and the carboxyl group at the end (C-terminus) of a growing polypeptide chain and a free amino group on an incoming amino acid are bound together until completion, followed by dissociation of the ribosome complex.

In the case of larger biomolecules such as recombinant proteins and antibodies, natural synthesis is the preferred alternative, as current chemical synthesis technologies cannot produce such large molecules[34]. Although peptide synthesis through the natural pathway has been explored[35, 36] primarily through recombination technologies; this method of peptide production would not meet commercial demands as biological systems only produce low concentrations, making isolation and purification challenging and, in most cases, impractical. Thus, chemical synthesis is clearly preferred for industrial scale peptide manufacture.

2.3.2 Chemical synthesis

Peptides are chemically synthesised through classical solution peptide synthesis (CSPS), solid- (SPPS) or liquid- (LPPS) phase peptide synthesis. These synthesis methods are all stepwise procedures similar to natural synthesis, where each amino acid is sequentially incorporated into the synthesised sequence. The most crucial benefit of chemical synthesis is that it allows scientists to engineer peptides by incorporating unnatural or non-canonical amino acids to improve potency, linkers to improve solubility and other complex modifications such as cyclisation.

CSPS was the first method developed for peptide preparation. It is beneficial from a green and financial perspective, as it requires only a slight excess of reagents compared to newer techniques, but it is time-consuming. All intermediate peptides are isolated and characterised, which increases the level of control over the final sequence but compromises purity, time, and human effort[37].

LPPS combines the CSPS and SPPS technologies by carrying out all reactions in solution like CSPS, but the peptide is supported on a soluble tag. This tag gives unique properties that facilitate dissolution. These characteristics allow the technology to reduce the amount of solvent required while still being capable of obtaining high quantities of material[37]. Since this technology is still developing, no more details will be covered in this Thesis.

SPPS is the preferred technology for peptide synthesis as it is scalable and more effective for synthesising long peptide chains up to about 50 residues[38]. In this procedure, the peptide chain is grown on an insoluble solid phase resin. The resin, composed of millimetre-scale solid beads, is treated with short organic linkers. These linkers prevent bead aggregation. Unlike ribosome protein synthesis (which occurs from the N- to the C-terminus), peptide synthesis on solid support proceeds from the C- to the N-terminus. The first amino acid remains covalently coupled to the linker via its carboxylic acid terminus, protecting the C-terminus from side reactions. The peptide remains immobilised on the solid support until the sequence is complete, while liquid-phase reagents and byproducts are washed away. Depending on the linker, various C-terminal modifications are possible. Peptide elongation involves repeated cycles of deprotection of the N-terminus protecting group and coupling of the following amino acid derivative[39]. When the peptide chains are longer and attached to the resin, self-aggregation occurs more frequently on the resin surface. This reduces the efficiency of the amino acid condensation reaction[38]. Yield in SPPS is calculated using equation 2.1. The nature of SPPS facilitates automated peptide synthesis. It also facilitates post-translational backbone modifications and incorporation of non-natural amino acids.

$$\text{Yield}_{\text{SPPS}} = (\text{Yield}_{\text{step}})^{\text{number of AAs}} \quad (2.1)$$

The synthesis quality also depends on the number of peptide-related impurities in the crude mixture. Table 2.1 describes possible related impurities after SPPS. These impurities are

removed until the desired purity of the product is achieved. The level of purity will depend on the application (Figure 2.6). In terms of therapeutic impurities, these requirements are set by regulatory agencies, for example:

- According to the United States National Institute of Health, all peptides prepared through SPPS for toxicology and clinical studies should have a minimum purity of 93% for small quantities (5-100 mg) and 97% for gram-scale batches (1-50 g)[40].
- The FDA advises that all peptide-related impurities of 0.10%-0.50% of the drug substance should be fully characterised, including the structure[21].

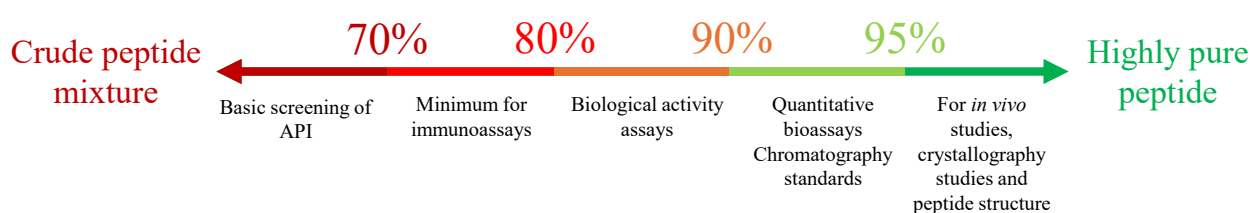


Figure 2.6 | General purity requirements for peptides.

Table 2.1 | Impurity classes from SPPS procedures. Information summarised from [41].

Impurity class	SPPS step	Description
Amino acid deletion	Inefficient coupling step	If the transient protecting group is not entirely removed or the incoming amino acid is not adequately activated during solid-phase peptide synthesis (SPPS), it can lead to inefficient coupling. Consequently, the resulting peptide chains may lack desired amino acid residues.
Amino acid insertion	Inefficient washing after coupling	An excess of Fmoc-protected amino acids is used to maximise coupling efficiency. However, if reagents aren't thoroughly washed away, an extra amino acid can be inserted into the desired peptide sequence. Additionally, when activated, the incoming amino acid may acylate an amide group on the resin-bound peptide chain. In the next coupling step, transient deprotection removes the Fmoc group from this acylated amino acid, exposing a reactive amine that reacts with the peptide bond, effectively inserting the amino acid into the sequence.
Covalently attached protecting groups	Removal of protecting groups	Protection groups are covalently attached to peptide sequences when the removal of permanent protecting groups is incomplete.
Oxidation/reduction	Prolonged exposure to light and air during storage	Methionine, histidine, lysine, cysteine, and tryptophan are prone to oxidation.

Diastereoisomerisation (racemisation)	Coupling step	Amino acid racemisation has been reported for classical and microwave-assisted Fmoc SPPS. Both direct enolisation and oxazolone (azalactone) formation are base-catalysed mechanisms.
Side- and end-chain impurities	Coupling	During coupling, some side chains can react and form a secondary product. For example, arginine can undergo aminations, which become a primary amine.
Aggregation	Any	Peptides have self-aggregation properties that can bind covalently or non-covalently two or more monomers together.

2.4 Peptides on RPLC separation

In RPLC, the separation of peptides is influenced by several intrinsic properties that determine their interaction with the hydrophobic stationary phase and their behaviour in the aqueous-based mobile phase. Key factors include hydrophobicity, which affects retention time based on the peptide's affinity for the hydrophobic stationary phase; amino acid composition and sequence, which influence hydrophobic interactions and peptide conformation; and peptide length, impacting the overall interaction with the stationary phase. Additionally, the charge and isoelectric point (pI) of the peptides, secondary and tertiary structures, and modifications play significant roles in their separation profiles. The solvent composition and operational temperature of the RPLC system also critically influence peptide retention and resolution. Understanding these properties is essential for optimising RPLC conditions to achieve efficient and precise peptide separation, which is crucial in proteomics, peptide mapping, and analytical biochemistry. A summary of peptide characteristics that can affect their RPLC separation is included in Table 2.2.

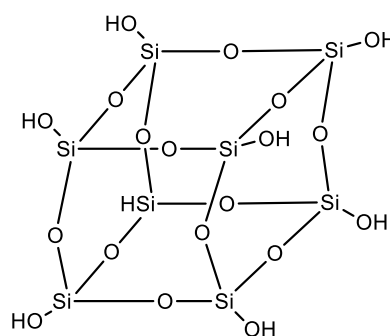
Table 2.2 | Effect of peptide properties in RPLC separation.

Peptide property	Definition	Effect on RPLC
Hydrophobicity	Peptide residues repel water and are associated with nonpolar environments.	Peptides with higher hydrophobicity tend to interact more strongly with the hydrophobic stationary phase of the RPLC column.
Amino acid composition	Specific types and order of amino acids in a peptide.	The presence of hydrophobic amino acids increases retention time. The sequence also affects peptide conformation, influencing how the peptide interacts with the stationary phase.
Peptide length	Number of amino acids in a peptide	Longer peptides generally have higher retention times because they have more hydrophobic interactions with the stationary phase. However, this can be influenced by the overall hydrophobicity and charge distribution.
Net charge	Influenced by the ionisable groups on the amino acids and the pH of the mobile phase	The charge of peptides can affect their interaction with the stationary phase and their solubility in the mobile phase. Adjusting the pH can change the ionisation state of the peptides, altering their retention times.
Secondary and tertiary structure	The folding and three-dimensional shape of the peptide.	Structured peptides may have different retention times compared to their denatured forms because the conformation affects how the peptide interacts with the stationary phase. Disulfide bonds and other intramolecular interactions can stabilise specific conformations.
Chemical and Post-translational Modifications (PTMs)	Chemical modifications to peptides include phosphorylation, glycosylation, and acetylation after their synthesis.	PTMs and chemical modifications can significantly alter the hydrophobicity, charge, and overall structure of peptides, affecting their retention and separation during RPLC. For example, phosphorylation usually increases hydrophilicity, reducing retention time.

RPLC resins

3.1 Silica architecture

The diagram illustrates a linear silicate polymer chain. It consists of a repeating unit of a silicate network. The backbone is formed by silicon atoms (Si) connected by oxygen atoms (O). Each silicon atom is bonded to two oxygen atoms within the chain and has two hydroxyl (OH) groups attached to it. The chain is enclosed in brackets with a subscript 'n'.



Page | 43

The silica surface is substantially covered with hydroxyl groups, known as silanols. The theoretical estimate of a maximum number of silanols involves the activation of all Si atoms available on the surface; this number is 4.6 Si atoms per nm² of surface area, and it is normally referred to as the Kiselev-Zhuravlev constant[42]. These Si-OH groups are the primary sites for chemical modification and thus play a significant role in the chromatographic behaviour of silica columns. Common chemical modifications involve incorporating a bonded phase or chemicals onto the surface using organosilanes, such as chlorosilanes or alkoxy silanes. The reaction commonly involves replacing surface hydroxyl groups with organic chains in a process called silanisation. Commonly used bonded phases include octadecylsilane (C18), octylsilane (C8), phenyl, cyano and amino groups. Examples of these common phases/ligands can be found in Figure 3.3. Each phase will give the surface a particular property, such as increased hydrophobicity, capacity to form π - π interactions and selectivity for polar compounds. Additional small silane molecules can be used to endcap any remaining active silanols, further reducing polar interactions. In practical chromatographic operations, this process improves peak shape and reproducibility. The intrinsic chemical nature of the silica and all these surface modifications produce a heterogenous surface that is depicted in Figure 3.2.

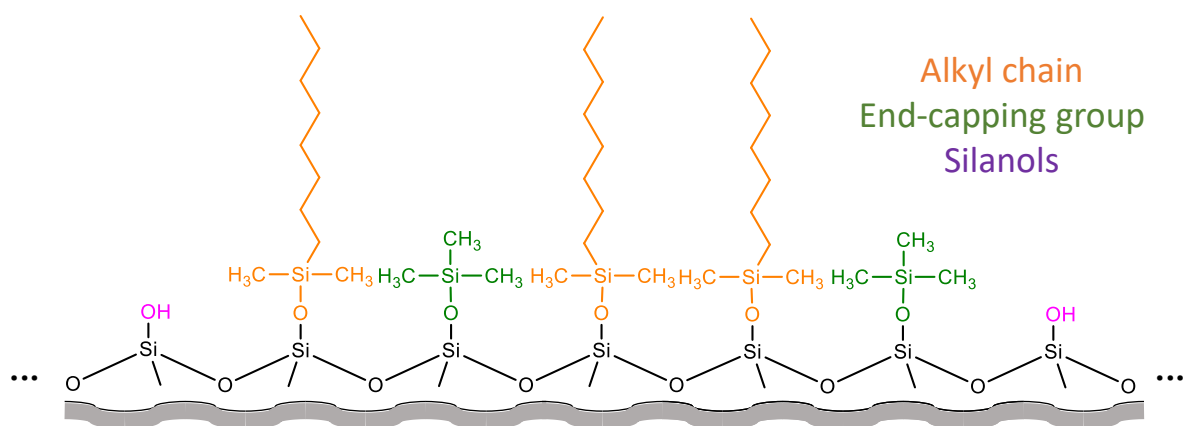


Figure 3.2 | Heterogenous surface of a silica resin.

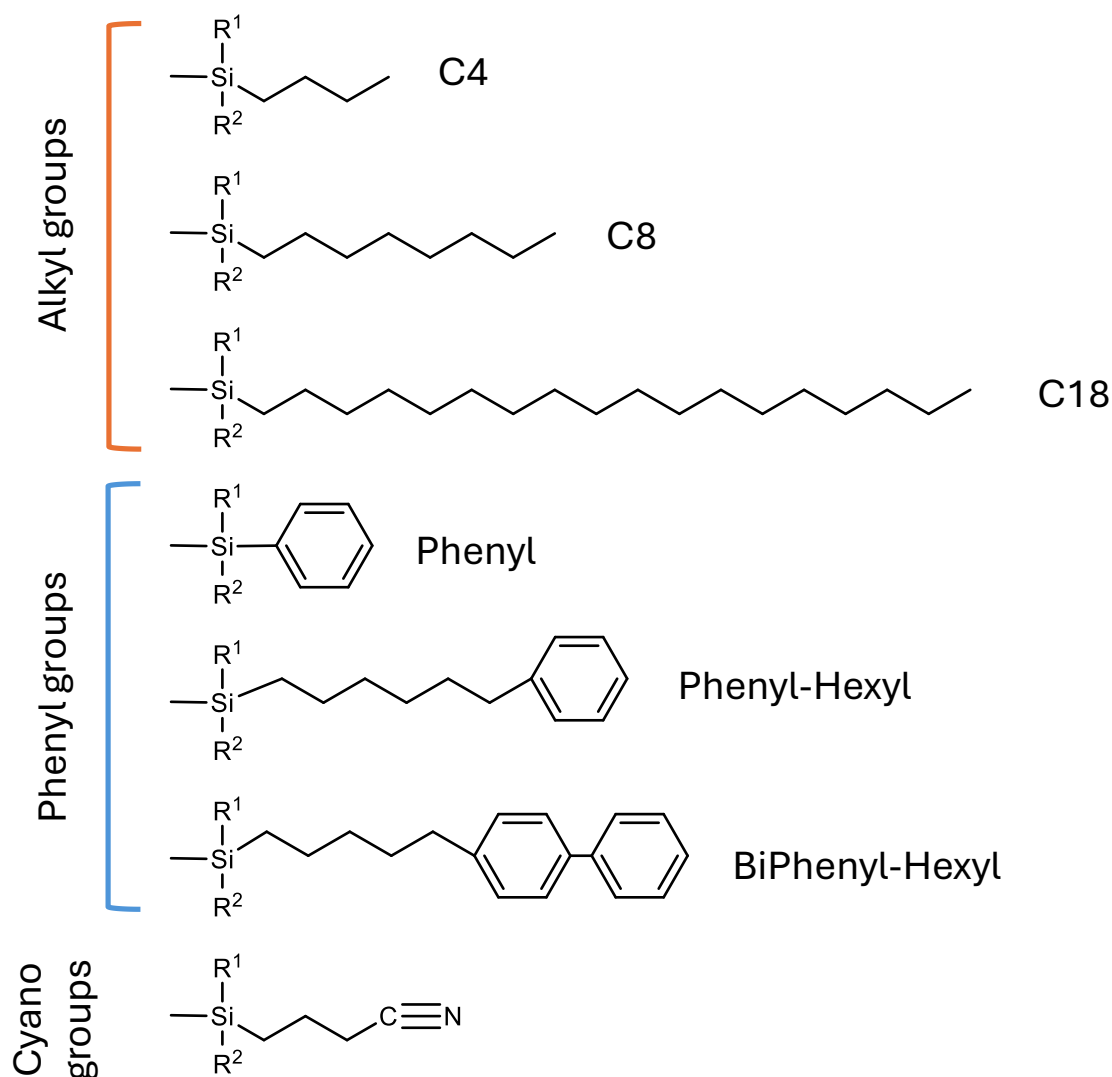


Figure 3.3 | Commonly used RPLC ligands/phases. Radicals R^1 and R^2 are replaced by shorter chains. The most commonly used groups are $-CH_3$.

3.1.1 Classification

Silica support materials are categorised into types A, B and C depending on the method employed to synthesise the material[43, 44]. The first type, type A, is prepared by precipitation from solutions of silicates and is characterised by higher acidity. This acidification is due to a high metal contaminant concentration, resulting in strong retention or peak shape deformation due to the chelation of solutes; thereby, this silica is generally avoided for chromatographic applications. Secondly, type B is synthesised through the aggregation phenomenon of an organosilane, such as tetraethoxysilane (TEOS), in the sol-gel process. This synthesis method gives the silica

advantages over its type A counterpart, including a higher population of surface silanol groups, more stability at low and intermediate pH ranges (up to 9.00), and an ultrapure structure with minimal metal contaminants. Finally, type C silica is produced from the silanisation/hydrosilylation process, thus also termed hydride silica. More than 95% of the silanol groups localised on the surface are replaced with silicon hydride groups during this preparation.

The nature of the different production methods for silica packings and the distinct quality of materials obtained necessitates a need to characterise the surface (types of silanols and siloxanes) and to determine the structural heterogeneity of the pores or channels present (cylindrical, funnel-shaped, and ink-bottled). Although these properties are highly influenced by the production method, surface properties can be characterised in all three types of silica at both micro- and macroscale.

3.2 Silica resin properties

Silica has a well-studied and controllable pore structure and morphology (fully porous, core-shell, or monolithic; Figure 3.4), which results in high reproducibility, rapid high transfer, and optimal chromatographic loadability. To stay in line with the objectives of the project, this Thesis will focus entirely on fully porous packings.

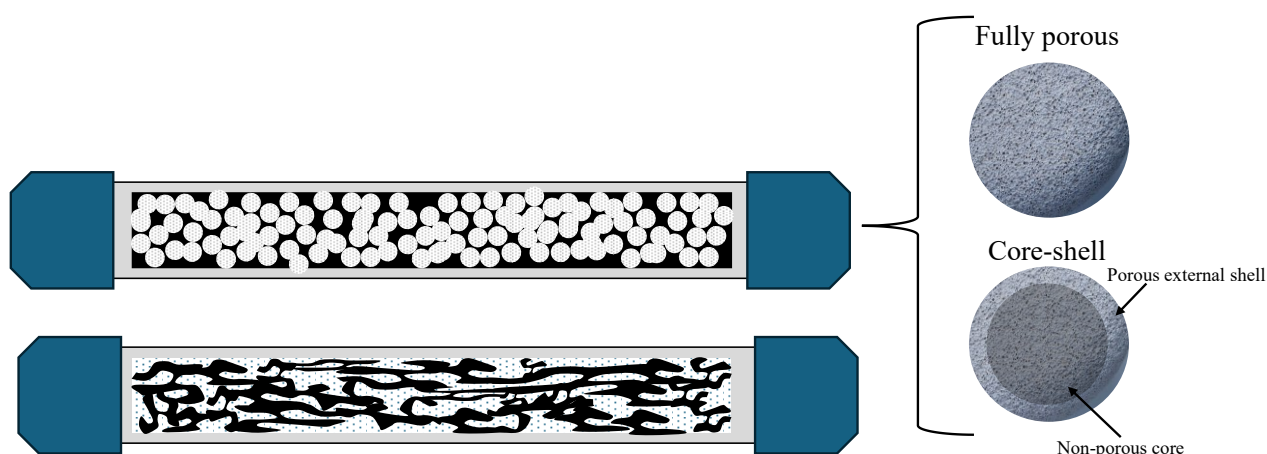


Figure 3.4 | Morphologies of chromatographic silica resins. (Top) Particulate resin can be divided into fully porous and core-shell. Channels are formed within the interstitial volume between the resin particles. (Bottom) Monolithic column. The resin is formed as a single polymerisation reaction that leads to a polymer network. Channels are formed by the through-pores between the resin's skeleton.

3.2.1 Microscale properties

The microscale properties of silica are determined by the characteristics of pores in the packing material, which define the surface topology. There are three types of pores present in silica: macropores (diameter >50 nm), mesopores (diameter 2-50 nm), and micropores (diameter <2 nm), where the distance is based on the diameter of the pore aperture[45]. Thus, a correct approximation of the surface topology compares descriptors such as surface area, mean pore diameter, specific pore volume, and pore size distribution.

All these pore-related properties have a significant impact on the features of the silica-based packings and the silica derivatisation for such materials. Small diameters are undesirable due to the steric hindrance of any solute molecules. The hindrance effect is evident even during the resin manufacturing process when incorporating the bonded phase becomes virtually impossible during the first end-capping step if the pore size is too small. This is particularly true for long hydrocarbon chains, such as the most common C18[46, 47]. In addition, if the pore is too narrow and cylindrical, silanisation at the edges of the pores induces an associative effect that 'blocks' the pore and prevents bonded phase population.

General visual information about porous materials is obtained from imaging techniques such as scanning electron microscopy (SEM) and transmission electron microscopy (TEM)[48]. The former provides a three-dimensional-like visual representation of the material surface, whereas the latter visualises organic matter comprised of light elements by staining materials with a heavy metal species containing solution. These techniques give a clear image of the surface of the materials, but they do not provide any quantitative chemical or thermodynamic data on the properties of chromatographic interest. Atomic force microscopy (AFM) has been suggested as a more reliable test because it is possible to obtain a sub-nanometre resolution of polymeric absorbents with dendrimer structures while being compared with data obtained from adsorption measurements[49]. Although AFM can potentially be the best imaging technique to analyse material surfaces, it still does not provide the level of chemical information required for complete characterisation.

Alternatively, surface chemical descriptors can be measured through the adsorption of selected molecules into the chromatographic resins of interest. For example, low-temperature nitrogen adsorption (LTNA) is the most popular method because the N₂ molecule can access all pore sizes,

and is guided by ASTM and ISO measurement standards. This standard technique offers a measurement of the surface area and measurement of pore volumes that can be normalised to the total mass of the silica being tested[50]. Thus, it provides a quick procedure for investigating new materials and conducting quality control during the production of currently commercialised packing materials.

N₂ adsorption measurements are then analysed and or modelled using equations from the Brunauer-Emmett-Teller (BET) theory, Barret-Joyner-Halenda (BJH) theory and density functional theory (DFT), which can provide more detailed information on the surface area, the pore size and structure. Although the BET and BJH are the most common models for characterising materials, they fail to describe pore size and pore structure morphology in silica. Therefore, DFT based analysis is frequently combined with LTNA data since its more complex adaptations, i.e., non-local (NLDFT) and quenched solid (QSDFT), have virtually overcome limitations of other models by also considering amorphous and microporous silica structures with the capacity of describing surface roughness, giving more realistic results[51, 52].

All of the N₂ adsorption descriptors measured by the previous techniques provided by these various analyses, are complemented with the calculation of the ligand coverage density of the packing material. This calculation is performed according to the Berendsen-de Galan equation for monofunctional silanes shown in eqn 3.2 [53, 54]:

$$\alpha_{RP}(\mu\text{mol m}^{-2}) = \frac{10^6 P_c}{1200 n_c - P_c(M - n_H)} \frac{1}{S_{BET}} \quad (3.2)$$

where α_{RP} is the coverage refers to the coverage density, P_c to the percentage of carbon content (%) calculated through elemental analysis (CHN), n_c is the number of carbon atoms in the ligand, M is the molecular weight of the ligand, n_H is the number of functional groups in the reactive group of the silane, accounting for the number of hydrogen atoms lost in the surface reaction and S_{BET} is the specific surface area ($\text{m}^2 \text{g}^{-1}$).

An alternative to the ligand coverage density method is thermogravimetric analysis (TGA) coupled with differential scanning (DSC). Heating the chromatographic silica packings to between 200 and 600 °C induces weight loss, usually due to the loss of organic groups attached to the

silica surface. Subsequently, DSC enables the measurement of the heat absorbed by the sample and the corresponding specific heat capacity.

The main problems of nitrogen adsorption methods, imaging, and elemental analysis are (1) they require the removal of the silica packing material from the column, (2) the analysis is performed at conditions that bear no relationship to the real operation conditions, such as in dried state or subjected to vacuum, (3) the sample material is not usable after the analysis, and (4) they provide little direct chemical or thermodynamic information. In operation, liquid chromatography columns are far from ideality; they always operate in wet conditions; the chemical interactions with solutes are not exclusively dependent on the resin; particles are not an exact copy of each other and are unevenly distributed along the column. Therefore, all current characterisation methods do not and will never provide enough information to characterise the intrinsic complexities of a RPLC chromatographic column or its adsorption mechanisms. Ideally, a suitable column characterisation should be performed via an *in-situ* approach.

A promising newer technique is the use of inverse-size exclusion chromatography (ISEC), which consists of injecting a series of known monodisperse species into a packed column, enabling the characterisation of the interstitial and pore volumes of porous packings[45]. This chromatographic method has successfully characterised a range of pores present in silica-based columns using mixtures of polyethylene glycols, oligosaccharides, methanol, and D₂O[55] and nanoparticles (NPs) of polystyrene. However, only the latter example provides enough data from 10 to 4000 Å to be adjusted to a probability density function that simulates the pore size distribution of the column. Other attempts with this approach have included using very spherical proteins such as ovalbumin and haemoglobin[56] as the probe species. This later strategy does not give reliable results because the probe molecules selected could potentially change due to misfolding or aggregation depending on the composition of the mobile phase. In addition, to maintain repeatability, it is important that the volume of the particles remain constant during the entire procedure; thus, non-swelling materials are preferred.

3.2.2 Macroscale properties of silica resins

Macroscale physical properties include particle size, shape, and size distribution. The first two can be easily characterised through the same microscopic imaging and spectroscopy techniques

described in the previous section. However, estimating the size distribution of particles is more challenging because the contents inside the chromatographic columns are not accessible without disrupting the structure of the column.

The interest in estimating the particle size distribution inside the column relies on how the resin is packed into the column, which in turn describes the mobile phase flow path, and, by extension, the interactions of the analytes with the surface of the packing material. The effect of the packing heterogeneity has been analysed computationally by fluid dynamics, where it has been confirmed that both the porosity and the uniformity of the distribution of the particles inside the column have a great influence on the flow path and chromatographic band broadening behaviour[57].

Over the years, the concepts of radial and axial heterogeneity (also called dispersion by some authors) have become of particular interest in estimating the distribution of particles inside chromatographic columns. Experimental estimations of these parameters have included using laser-dye set-ups[58], microcylinder electrodes as detectors[59] and on-column electrochemical micro-detectors[60]. Although these methods provide qualitative and semi-quantitative information about the packing material's heterogeneity, they are usually ignored in practice due to the impracticality of setting complex experimental layouts where the detectors have to be custom-built according to the column.

The most efficient and accurate method to measure these parameters would be to elute standard compounds with known adsorption behaviours. Gritti and Guiochon[61-63] determined the adsorption isotherms of phenol, caffeine, sodium 2-naphthalene sulfonate, and propranolol chloride as test probes to assess the chemical surface heterogeneity in RPLC materials. These compounds were selected because they all exhibit convex upward Langmuirian behaviour with no or minimal adsorbate-adsorbate interactions. Also, the difference in their molecular sizes allows the study of the influence of this variable on adsorption mechanisms.

3.2.3 Silanol activity

There are three types of silanol groups (Si-OH) present on the surface of silica packing materials: isolated (6-9% of the total content of all), vicinal (60-65%), and geminal (10-12%). The

difference between the last two is that vicinal silanols form hydrogen bonds with neighbouring silicon atoms, whereas geminal are two hydroxyl groups bonded to the same silicon atoms (Figure 3.5). The condensation of these two types of silanol particles produces siloxane groups comprising 20-35% of the total number of surface functional groups. The number of silanol groups could increase from the hydrolysis of functional groups in bi- or trifunctional silanes. In contrast, new silanol groups and siloxane moieties can be created from the bonding of organosilanes.

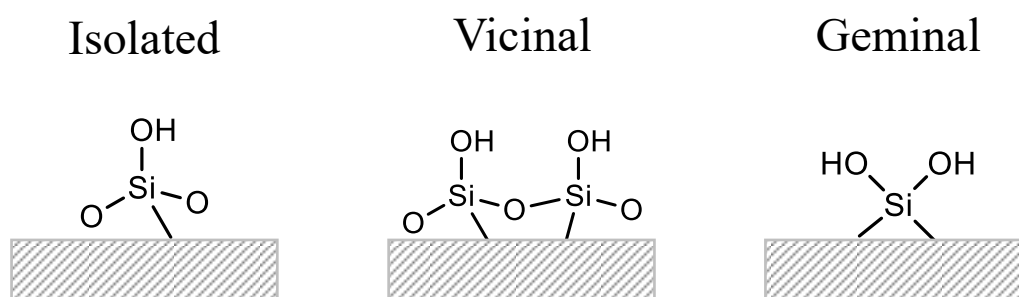


Figure 3.5 | Surface silanol groups present in RPLC resins.

Silanol groups can contribute to numerous intramolecular interactions such as hydrogen bonding, van der Waals, and London dispersion forces. It has been suggested that isolated and geminal silanols are the moieties with the most significant influence on the retention of solutes by forming polar active centres[64]. However, there has not been any definitive answer as to which silanol group dominates as the primary reaction and adsorption site. It should be emphasised that compared to silanols, siloxanes are hydrophobic entities unable to form donor-acceptor interactions; thus, it is believed that their effect on solute retention is neglectable[65]. Peptide analysis is generally performed at low pH to minimise the interactions of peptides with deprotonated residual silanols at the intermediate pH, which can cause detrimental band spreading and excessive tailing for ionised species.

Although the mere presence of residual silanols may not suffice to explain the origin of peak tailing in chromatography entirely, column manufacturers have attempted to reduce the density of these species on the stationary phase present, and thus their effect, by offering new chromatographic supports based on either wholly organic or hybridised solid supports. One approach involves partially substituting free silanols with methyl groups by developing spherical porous hybrid particles made by the co-condensation of methyltriethoxysilane with TEOS. These methyl hybrid materials are commercially available with an empirical formula of $\text{SiO}_2(\text{CH}_3\text{SiO}_{1.5})_{0.5}$

under the trade name of XTerra. Another similar approach is the ethylene-bridged hybrid (BEH) technology, which was introduced for the synthesis of inorganic-organic particles by the co-condensation of 1,2-bis(triethoxysilyl)ethane with TEOS producing an empirical formula of $\text{SiO}_2(\text{O}_{1.5}\text{SiCH}_2\text{CH}_2\text{SiO}_{1.5})_{0.25}$. This embedment of ethylene-bridged groups reduces by one-third the number of surface silanol groups present[66].

The most recent approach to reducing surface silanol groups is called Type C silica-hybrid materials. This material seems to be a very promising support, and it has been suggested that approximately 95% of the Si-OH groups are replaced with silicon-hydride groups (Si-H) moieties[43]. The large population of Si-H groups provides the particle with a lower polar surface, resulting in chromatographic analyses with better peak shape and resolution. However, the market for Type C silica is still in its early stages; therefore, there is not much definitive information available yet on the effectiveness of columns with this type of packing material.

Spectroscopic techniques offer more detailed information at an atomic level about the surface structure, conformation, and dynamics of column packings. The use of high-resolution ^{29}Si nuclear magnetic resonance coupled with cross-polarisation and magic-angle-spinning (^{29}Si CP/MAS NMR) is a powerful tool to determine the distribution of geminal and single silanol groups on heterogenous silica materials and thus the surface coverage with organic ligands[67, 68]. Chromatographic packings can also be characterised using Fourier-transform infrared (FTIR) spectroscopy, spectrofluorimetry, Raman spectroscopy, X-ray photoelectron spectroscopy (XPS), and static or time-of-flight secondary mass spectrometry (ToF-SIMS). These spectroscopy modes generally allow the chemical confirmation of organic ligand immobilisation onto the silica surface.

Since silanols are always partially ionised at acidic chromatographic conditions, the ζ potential can also be used to determine the charge distribution on the surface of particles; therefore, it is a valuable parameter for assessing the silanol activity of chromatography column packings.

A graphical summary of all techniques described in the last subsections can be found in Figure 3.6

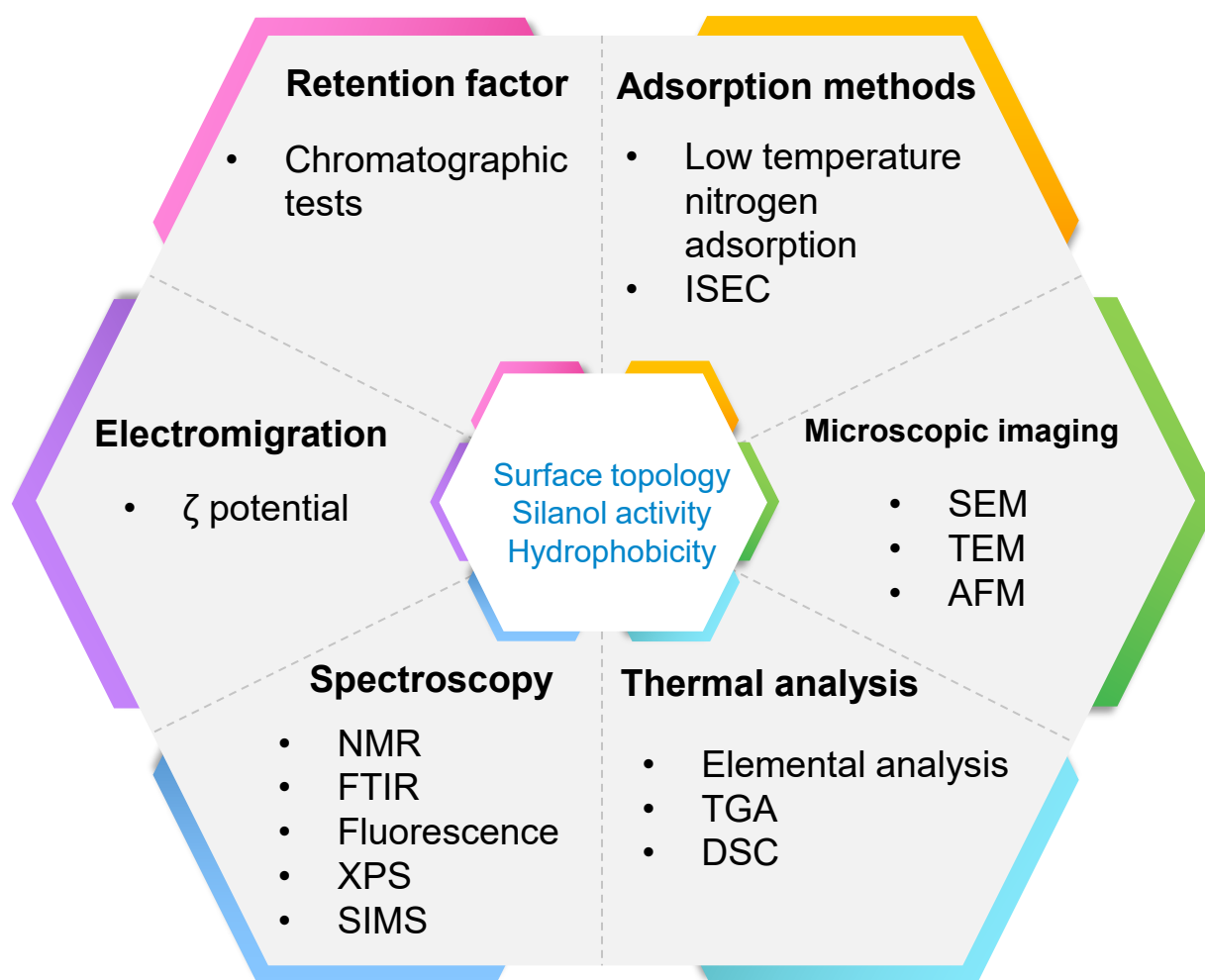


Figure 3.6 | Classical techniques for RPLC resin characterisation. Summary of common technologies to characterise silica-resin properties.

3.2.4 Column hydrophobicity

Current chromatographic methods for peptide purification modify the peptide adsorption selectivity of RPLC columns by changing the composition of the mobile phase. Although it is generally accepted in the literature and intuitively rational to think that column hydrophobicity has a crucial effect on the peptide adsorption mechanism, more research attention is needed to test or investigate this hypothesis rigorously. Perkins et al.[69] obtained data suggesting that the entropically driven nature of hydrophobic interactions in RPLC correlates with the calculated number of water molecules released upon adsorption. These findings suggest that a hydrophobic effect similar to that of proteins during folding occurs on the surface of silica-based packing materials when the peptide is adsorbed onto the chromatographic resin surface.

As current theories only focus on hydrophobicity as the driving principle of adsorption in RPLC columns, it is of great interest to accurately predict and determine the hydrophobic contribution to the RPLC separation process. Some authors have suggested that the methylene selectivity (α_{CH_2} value) best describes column hydrophobicity[70]. This argument is mainly based on the idea that a methylene group is the most hydrophobic group and can cover all surface areas without being influenced by steric factors. Thus, even if there were any contributions, they would be minimal. Another advantage of this parameter is that it can be a proxy for surface carbon content.

3.3 Column Selection Systems

Empirical chromatographic factors or the application of analytical techniques described in Section 3.2 do not reflect the complex, diverse mutual chemical interactions between analytes, solvent molecules, and the mobile and stationary phases. These chemical or surface descriptors do not achieve the reliability and robustness required to become rational, objective, and reproducible column characterisation and classification procedures.

To reliably determine the properties of a column that directly contribute to its quality and separation power, it becomes crucial to design *in-situ* experiments using different test compounds, eluents, experimental conditions, and calculation procedures to determine proper independent system constants with physical relevance. As a solution for this problem, column selection systems (CSSs) have been proposed as tools for classifying stationary phases according to their ability to separate selected analytes, also referred to as probes, under controlled conditions.

The first CSS was developed by Walters in 1987[71] using 12 different C₁₈ columns. He proposed classifying columns according to hydrophobicity, silanol activity, and column efficiency. Hydrophobicity is evaluated by the selectivity factor from an injection of anthracene and benzene ($\alpha_{A/B}$), whereas silanol activity is estimated using N,N-diethyl-m-toluamide (DETA, previously called DEET; $k_{DETA/A}$) and the retention factor of nitrobenzene (k_{NB}). After two years of Walters' contribution, Engelhardt and Jungheim proposed a similar but more straightforward CSS[72]. They characterised 43 C₈ and C₁₈ columns by means of their hydrophobicity measured by the relative retention of ethylbenzene and toluene using gradients of methanol with 10-50% H₂O. In both scales, molecular probes were selected for the easy detection of aromatic rings at 254 nm

and a noticeable increase of hydrophobicity with extended chains or the addition of rings to the shared minimal structure.

In 1989, Kimata et al.[70] developed the Tanaka method for the characterisation of RPLC columns. The method consists of a set of four chromatographic runs that vary the contents of the mobile phase and, thus, the retention behaviour with the resin. The different solutions can consist of H₂O mixed with methanol or phosphate buffer at pH 2.7 or 7.6. The power and popularity of this method rely on its capability to render qualitative information on the column about its hydrophobicity selectivity, steric selectivity, hydrogen bonding capacity (considered as well as silanol activity), and ion exchange capability at different pH levels. Parameters of the CSS are based on the retention and capacity factors of injections of the following molecules: uracil, pentylbenzene, butylbenzene, triphenylene, o-terphenyl, caffeine, phenol and benzylamine. It is important to mention that hydrophobicity under this scale is represented by the selectivity of a methylene group described by parameter α_{CH_2} . As previously mentioned in the Section 3.2.4, this chemical group can be considered the most hydrophobic compound because it cannot form hydrogen bonds, and its surface free energy is relatively low. This change in the approach to estimating hydrophobicity makes the Tanaka method more significant than its counterparts while maintaining the probes' detectability.

The parameters determined in the Tanaka method are:

- **k_{PB} – Hydrophobicity:** Represents the hydrophobic retention capacity of the column measured by the injection of n-pentylbenzene. Dependent on carbon load and surface area.
- **α_{CH_2} – Hydrophobic selectivity:** Represents the capacity of the column to selectively resolve a methylene group through the injection of n-pentylbenzene and n-butylbenzene.
- **$\alpha_{T/O}$ – Steric selectivity:** Accounts for the capacity of a stationary phase to discriminate compounds of identical elemental composition but different three-dimensional structures. It highly depends on the length of the bound chain, bonding density, and the type of bonding. Injection of T: triphenylene and O: o-terphenyl.
- **$\alpha_{C/P}$ – Hydrogen bonding capacity (or silanol capacity):** Reflects the H-bonding capacity of the column proportional to the number of available silanol groups present on the surface of the resin. Injection of C: caffeine and P: phenol.
- **$\alpha_{B/P, pH\ 7.6}$ – Ion exchange capacity at pH 7.6:** accounts for the total number of free silanol groups on the column. At this pH, most silanols possess a negative charge, while benzylamine is

protonated with a positive charge, thus forcing electrostatic interactions. Injection of benzylamine and phenol.

- **$\alpha_{B/P, pH 2.6}$ – Ion exchange capacity at pH 2.6:** Accounts for the total number of acidic silanols on the column. Since there is no protonation, electrostatic interactions should not be expected (theoretically). Injection of benzylamine and phenol.

Since its introduction, the Tanaka method has slowly become a standard characterisation protocol for commercial RPLC columns because it is easy and relatively fast to perform. Cruz et al.[73] introduced radar plots to visualise the parameters obtained in the Tanaka method into a seven-dimensional space that facilitates the comparison between columns. However, the applications of the method continued without any standardisation or statistical reliability analysis of the technique. It remained like that until 2005 when Petersson and Euerby[74] assessed the robustness of the protocol using design of experiments (DoE) coupled with principal component analysis (PCA). Their findings suggest that the Tanaka method is susceptible to small changes in the percentage of organic modifier; therefore, precision in preparing the mobile phases is required to obtain reliable, reproducible, and meaningful results. Assuming all of these precautions are taken into consideration, then the Tanaka method/procedure should be sufficient to create a standard protocol that all users and manufacturers could use to rank and rate commercial columns. ACDLabs have developed a database with Tanaka parameters for hundreds of RPLC columns as a first attempt to achieve this standardisation in the market[75]. An example of results from the Tanaka method can be observed in Figure 3.7.

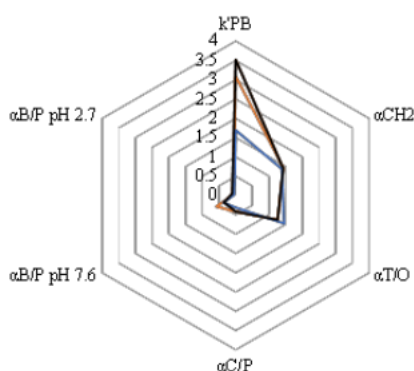


Figure 3.7 | Graphical results of the Tanaka method. Values corresponding to an RPLC column. Values from the ACDLabs database (black), the tested column (orange) and a column of the same type (blue).

Despite the effectiveness of the Tanaka method for empirical RPLC column performance comparison, there is a poor understanding of the fundamental retention mechanisms inside the column, which drive peptide partitioning and retention. A deeper fundamental understanding of peptide retention mechanisms is essential to fully understanding the performance of thousands of available column resins and to optimising RPLC separation methods.

3.3.1 Quantitative Structure-Retention Relationships in Column Selection Systems

The field of chromatographic retention prediction has introduced the Quantitative Structure-Retention Relationships (QSRR) approach. These relationships are models that correlate the retention behaviour of some probes and some empirical or fundamental descriptors for these molecules, using thermodynamic parameters determined from the adsorption interactions inside the columns.

The first QSRR models followed a linear regression-based analysis utilising linear free energy relationships (LFER). These equations assume a linear relationship exists between activation and reaction energies while entropies remain constant within a homologous series[76]. Regarding chromatography retention, Abraham and his colleagues established an equation that considers solubility-related solute parameters (SP) corresponding to simple fundamental solution properties[77, 78]. In the same work, they determined that SP could be the specific retention volume (V_G^0), or the partition coefficient of an analyte, transforming their QSRR model into the linear solvation energy relationships (LSER) model following the equation:

$$\log k = c + rR_2 + s\pi_2^H + a \sum \alpha_2^H + b \sum \beta_2^H + vV_x \quad (3.3)$$

where the parameters are fundamental physicochemical properties of the solutes. R_2 is a modified polarisability parameter that characterises the ability of a solute to interact with π - or n -electron pairs, π_2^H is an estimate of the dipolarity, α_2^H and β_2^H represent the hydrogen-bond acidity and basicity, respectively, and V_x is the McGowans characteristic molecular volume.

Another prominent QSRR model used for solute retention time modelling is the Linear Solvent Strength Model proposed by Baczek and Kaliszan[79]. The equation includes the total dipole moment (μ), electron excess charge of the most negatively charged atom (δ_{min}) and the water-accessible surface area (A_{WAS}). The model has the form:

$$\log k_w = k'_1 + k'_2\mu + k'_3\delta_{min} + k'_4A_{WAS} \quad (3.4)$$

where $k'_1 - k'_4$ are regression coefficients. In 2007, this QSRR equation was implemented to classify RPLC stationary phases through PCA[80]. The results were positive, and they stated that combining QSRR methods with statistical analyses could create a clear difference between columns. However, there are still questions about the relevance of these models and their lack of fundamental physicochemical insightfulness. It can be argued that QSRR models and, especially, the parameters in those equations that have been proposed are subjective and unreliable because not enough information is provided in terms of how relevant they are to the dynamic retention mechanisms occurring in the system they describe.

3.4 Using peptides as column descriptors

Since the molecular descriptors estimated from small molecules have been mistakenly believed to be sufficiently accurate or scientifically relevant, there has been very limited research on using other molecules as adsorbates, such as real-world biomolecules. As seen in Chapter 2, biomolecules such as peptides have entirely different characteristics from small molecules. Their size and complex heterogeneous structures create different complex physicochemical interactions with the chromatographic system that cannot be simply represented by column descriptors obtained with small molecules. Therefore, more time should be invested in obtaining reliable column descriptors based on injecting peptide probes as adsorption probes.

Plass, Valko, and Abraham[81] were the first to exploit the use of peptides to characterise RPLC packing materials under fast gradient elution conditions using derivative sequences from the tripeptide Z-AX_{aa}V-OMe. The Z (benzyloxycarbonyl) and OMe (methyl ester) were used to protect the N-terminal and C-terminal sites, respectively. Although they could not estimate any

column characteristics, their findings suggested that the use of gradient flow and the solvation equation developed by Abraham was robust enough to distinguish diastereomeric peptides with respect to their different hydrogen bond basicities.

In 2010, Mant, Cepeniene and Hodges[82] compared the effectiveness amongst silica-based RP-HPLC stationary phases, i.e. polar embedding and polar end-capping C₁₈, n-alkyl and aromatic phenyl hexyl, through the utilisation of the peptide standard sequence Gly-Gly-Leu-Gly-Gly-Ala-Ala-Leu-Gly-X-Leu-Lys-Lys-amide. X was substituted with Gly, Ala, Val, Ile in order of increasing hydrophobicity, Phe, Tyr as aromatic sidechains, and Tyr due to its polar character. Results showed that different columns and mobile phase conditions complemented each other for the resolution of specific peptide pairs, and more than one column and solvent type is required for maximum flexibility in both analytical and preparative applications. They also explored the same composition and minimal sequence variation (SCMSV), which questions if, in the absence of any variation in sidechain nearest-neighbour or peptide confirmation effects, a reversed-phase packing can separate peptide pairs when the only difference between them is the location of single amino acid, as well as how does the nature of the substituted amino acid affect the degree of resolution. Standard techniques to protect the N-terminal of a peptide are the addition of either N^α-acetylated or a free α -amino group[83].

3.4.1 Peptides in CSS

The work of Field et al. is probably the only complete adaptation of a CSS using peptides[84-87]. They designed arguably the first column characterisation system implementing peptides as probes. An array of 13 diverse stationary phases was assessed to provide a representative range in the selectivity and peak capacity of commonly used commercial pre-packed columns. The variations in characteristics included the base silica, the ligand density, the effect of alkyl lengths, the use of polar embedded group (PEG) and phenyl phases. Additionally, a column with a weak cationic exchange phase was introduced to the protocol to supply information on the hydrogen binding interactions between a protonated moiety in the column surface and the polar and charged peptide species. Multiple probes were derived from Bovine GLP-2 as the base sequence since it contains two well-defined hydrophilic and hydrophobic regions. In addition, the lack of cysteine reduces the probability of sulphide bridge formation shuffling.

Additionally, the distribution of residues is optimal for analysing the effect of common peptide degradation mechanisms. Trends within the data analysed through Principal Component Analysis deciphered groupings of electrostatic interactions amongst phases where positive charge delta values, a normalised measure of retention time between two peaks, were represented by columns with a negative character. In contrast, a negative delta value correlates with a positive character. Additionally, an intermediate region of low electrostatic interactions was found for the stationary phases with neutral character.

Furthermore, the method supported the hypothesis that trifluoroacetic acid (TFA) masks some peptide-column interactions, resulting in similar results within columns despite their polar character.

Chapter 4

Fundamentals of Liquid Chromatography

This Chapter is an in-depth exploration of the theory and principles underlying liquid chromatography (LC) systems. It was written as a reference material for beginners and advanced chromatographers. It begins by detailing the anatomy of an LC system and discusses essential physicochemical operational considerations within LC systems, including dwell volume, extra-column volume, and hold-up volume, which are vital for understanding and optimising chromatographic separations. Additionally, this Chapter covers elution profiles, differentiating between isocratic and gradient elution methods, and their impact on the separation of analytes. It finishes by discussing key retention parameters and examines different modes of liquid chromatography, emphasising the fundamental analyte/solute retention mechanisms and interfacial interactions critical for RPLC whilst covering other LC modes as well. Advanced topics such as chromatographically determined adsorption isotherms and the LC methodology to measure them are also discussed.

4.1 LC system anatomy

A liquid chromatography (LC) system is composed of different instrument modules (Figure 4.1) that work synchronously to deliver the mobile phase containing the analyte/solute through the chromatographic column and record the separation output while controlling parameters such as mobile phase composition, column temperature and mobile phase flow rate using appropriate chromatographic detectors and associated chromatographic modules.

The different LC system elements are:

1. **Mobile phase lines**—The liquid solvent is the mobile phase that carries the analyte/solute through the column. The mobile phase can be used as a single solvent or as a mixture and is chosen based on the properties of the sample and the stationary phase.
2. **Degasser**—This device removes dissolved gases, such as oxygen and nitrogen, from the mobile phase solvents before they enter the pump and column.
3. **Solvent delivery pump**—The pump delivers the mobile phase (solvent) through the system at a controlled and usually fixed flow rate. Depending on the complexity of the mixture, solvents are mixed through a binary or quaternary pump.
4. **Autosampler**—This module has two critical parts: the sample vial rack and the liquid sample injector module. The former is where the sample vials are placed, and the latter introduces the liquid sample into the mobile phase flow path.
5. **Column oven**—This regulates the temperature of the chromatographic column during analysis, helping to maintain consistent separation conditions.
6. **Column**—This item is the heart of chromatography and contains a packed solid stationary phase. The analytes/solutes in the mobile phase interact with the solid stationary phase differently, typically due to differing retention behaviour, causing them to move at different rates through the column and thus separate.
7. **Detector**—This measures the concentration of the separated components in the mobile phases they elute from the column. Common/standard detectors include UV-Vis, fluorescence, and mass spectrometers.
8. **Fraction collector**—This module collects the different mobile phase fractions of a sample as they elute from the chromatography column. Each fraction corresponds to a specific time or volume and contains separated compounds of interest.
9. **Communication bus and PC**—This system controls the LC modules as a whole unit and collects the detector signals to produce a chromatogram.

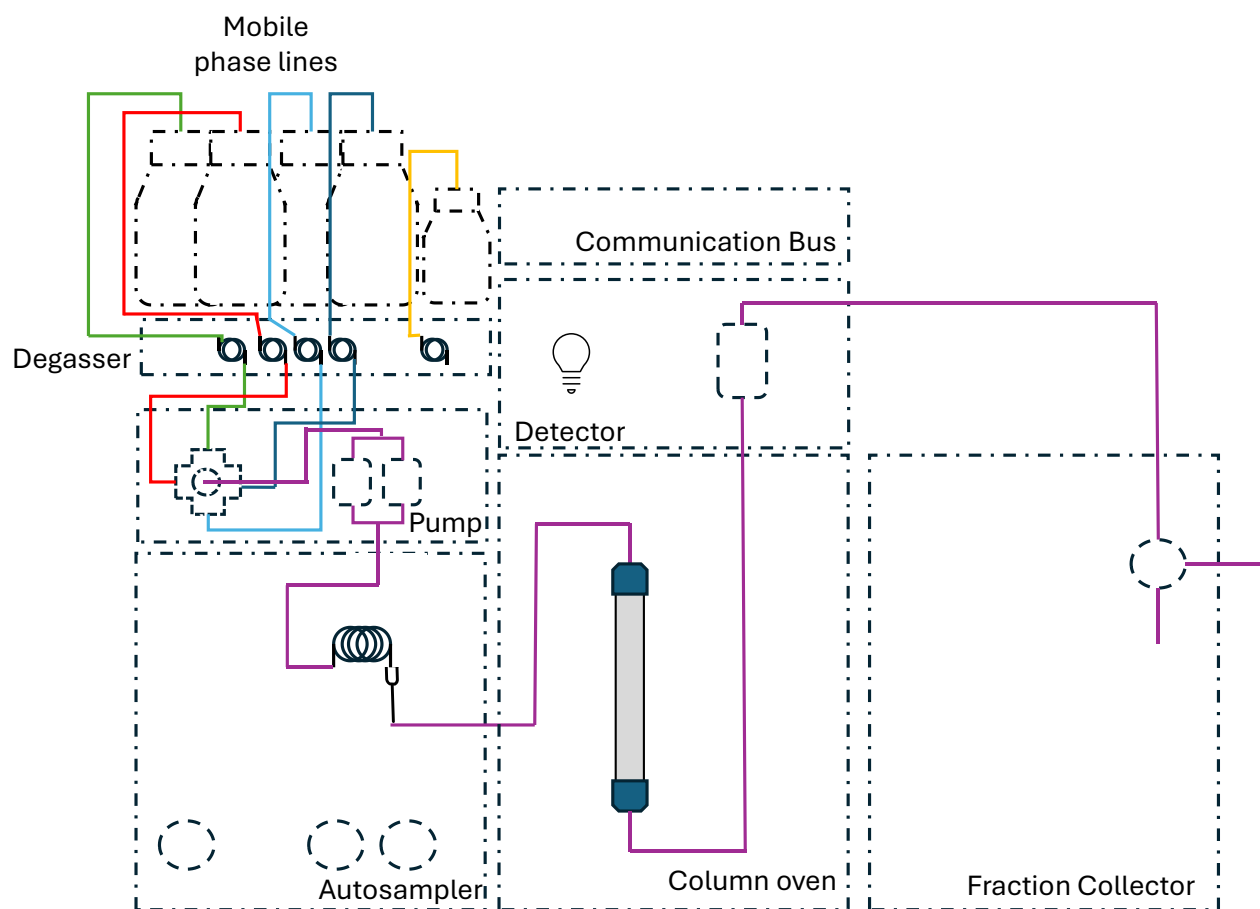


Figure 4.1 | General schematic diagram of a modular LC system.

4.1.1 Volumes in the LC system

As seen from Figure 4.1, all LC systems are composed of a network of tubing and chambers filled with the mobile phase. Due to the different physical scales of chromatography, from nanolitres to hundreds of litres in terms of mobile phase consumption, all pieces of hardware are of various lengths and internal volumes. This variability can hinder the transferability of chromatographic methods within systems during scale-up or scale-down processes due to the over-broadening of the mixed analytes/solutes chromatograms before the column or the separated compounds after the column.

Understanding the different volumes within an LC system helps to standardise a chromatographic method independent of the instrumentation or scale in which it has been developed. However, this task is not straightforward, as each LC instrument manufacturer has different

dimension standards. Volumetric information is usually unavailable, especially for tubing, as these are installed on the spot, depending on space and availability. However, static and dynamic experimental techniques can be used to characterise all of the volumetric characteristics of any LC system.

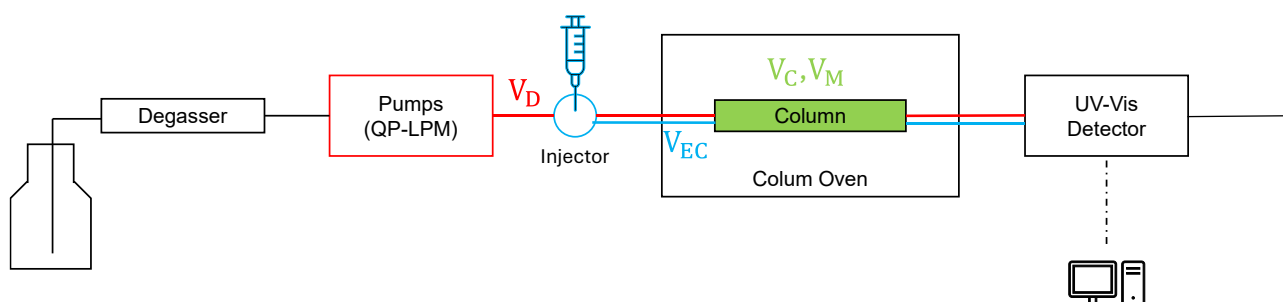


Figure 4.2 | Graphical summary of volumes in an LC system.

Dwell volume (V_D)

When multiple solvents are mixed at the pump, a delay can be observed from the point the mixing was programmed until the selected composition of solvent is recorded in the detector, exposing the sample to a short isocratic (constant organic concentration; see Section 4.2.1) front in the initial mobile phase composition[88, 89]. This delay in volume is called the dwell volume (V_D), whereas the equivalent time is referred to as dwell time (t_D). Some authors also use the term gradient delay volume or GDV, but the majority prefer t_D and V_D .

The value of V_D represents the volume from the mixing point of solvents to the detector (Figure 4.2). Since this delay is a function of the specifics of the solvent mixing system, it is a unique characteristic property of each LC system.

There are two types of mixing systems available for which different considerations are needed to define the dwell volume: (i) high-pressure mixing systems, where the dwell volume comprises the mixing chamber, the connecting tubing and the autosampler loop; and (ii) low-pressure mixing systems, where the volume of a pump and additional tubing are added to the components because the solvents are mixed upstream the pump.

Fortunately, despite the disparity of the dwell volume between instruments, it can be measured easily through a dynamic technique. Commonly, the mobile phase which flows through the

system contains a UV-absorbing solute component at a concentration, which results in an approximate detection signal of 1AU in a UV-transparent solvent such as methanol, water, and acetonitrile. The technique consists of mixing two mobile phase lines, where, for example, line A is pure solvent and line B has the UV-adsorbing component. Although V_D can be estimated by instantaneously changing lines A to B as the difference between the detector response time and the programmed time according to the definition; a gradient method is preferred as it accounts for true mixing.

Solvent lines A and B are mixed gradually over a 0-100%B ramp/gradient for a significant experimental time (ca. 20 min). The operational column is replaced by a zero-dead volume union or a capillary tubing of known volume. The flow rate should be sufficiently high to generate enough back pressure to confirm the complete operation of the valves. Then, the value of t_d is obtained graphically from the generated curve by drawing a tangent to the linear part of the ramp/gradient curve and extrapolating the baseline to find the intersection. The time between the start of the gradient and the intersection is the t_D value, which is multiplied by the flow rate to obtain V_D (method depicted in Figure 4.3).

Alternatively, the following equation can also be used with the same experimental results:

$$t_D = t_{50} - \frac{1}{2}t_G \quad (4.5)$$

where t_{50} represents the time at the midpoint of the gradient curve, i.e., when the baseline has increased by 50% of the maximum height and t_G is the ramp/gradient time.

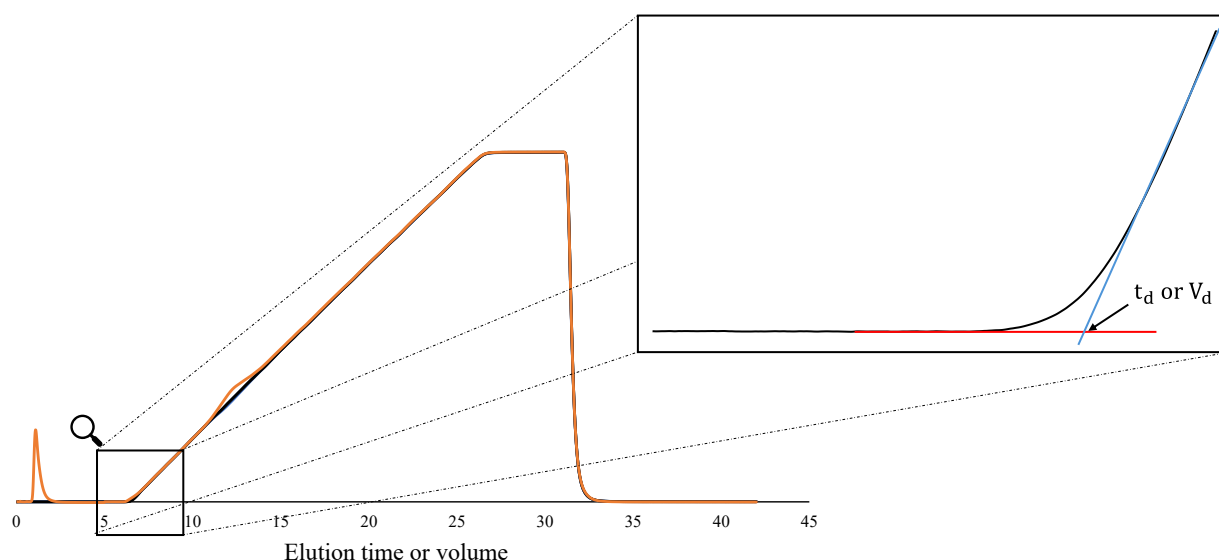


Figure 4.3 | Graphical calculation of dwell volume.

Extra-column volume (V_{EC})

The extra column volume refers to all systematic volumes that an analyte is exposed from the injector to the detector, excluding the separation column (Figure 4.2). This is particularly important for molecules with high dispersion coefficient (D_m) in the tubing, which can affect the separation. This phenomenon is called extra-column dispersion (ECD). Dispersion is quantified as the second moment or variance ($\sigma_{\text{observed}}^2$) of the Gaussian peak observed at the detector, which is an accumulative value for the flow path, as contained within equation 4.6. The ECD is equivalent to $\sigma_{\text{observed}}^2$ but removes the contribution of the column (σ_{column}^2).

$$\begin{aligned} \sigma_{\text{observed}}^2 = & \sigma_{\text{injector}}^2 + \sigma_{\text{tubing,pre-column}}^2 + \sigma_{\text{column}}^2 \\ & + \sigma_{\text{tubing,post-column}}^2 + \sigma_{\text{detector}}^2 \end{aligned} \quad (4.6)$$

The value of V_{EC} is estimated based on a similar dynamic approach as V_D , where the column is removed or replaced by a tubing of known length. However, in this approach, the composition of the mobile phase remains constant throughout the runs. Varying quantities of a UV-absorbing component are injected into the system at different flow rates until a peak is observed. The

difference between the apex of the peak and the programmed injection time at each set of experiments are values of t_D which is then multiplied by the different flow rates to obtain V_D .

In general, V_{EC} should be minimised to reduce the effect of the ECD. Since not much can be done at the injector as the injection loop is generally fixed, the impact of ECD can be reduced by reducing the dimensions of the tubing (equation 4.7)[90] or the detector cell volume (equation 4.8)[91]:

$$\sigma_{\text{tubing}}^2 = \frac{\pi^2 L_{\text{tubing}}^2 d_{\text{tubing}}^4}{48 + 384\pi L_{\text{tubing}} \frac{D_m}{F}} \quad (4.7)$$

$$\sigma_{\text{detector}}^2 = \frac{V_{\text{cell}}^2}{\theta_{\text{detector}}} \quad (4.8)$$

In cases where fraction collection is required, the contribution of the tubing after the detector should be considered. This extra tubing will contribute to the ECD, but most chromatography users generally ignore this contribution as it can become confusing. Since tubing after detectors tends to be wider and longer than the rest of the tubing in the system, the general recommendation is to reduce the length of the tubing as well. For the purpose of this Thesis, V_{EC} values are restricted by the definition, and the volume of the tubing post-detectors is not considered.

Hold-up volume (V_M)

The total column volume (V_C) is composed of several components that can be grouped into accessible and inaccessible volumes. Specifically:

1. Accessible (included) volumes:
 - Interparticle (Interstitial or External Pore) Volume (V_{inter} or V_i): This refers to the space between particles in the column packing material.
 - Intraparticle (Internal or Open Pore) Volume (V_{intra} or V_p): This volume exists within the individual particles of the stationary phase.

2. Inaccessible (excluded) volumes:

- Stationary-Phase (Support Solid) Volume (V_{SP}): This represents the solid material that makes up the stationary phase.
- Pore-Close Volume (V_{CP}): The volume of pores is too small for molecules to enter.
- Volume of the Bonded Ligand (V_{ligand}): The space occupied by the ligand bonded to the stationary phase.

Understanding these volumes allows the chromatographer to characterise a column regardless of the LC system, making method transferability across systems or dimensions easier. However, estimating all of these separate values is challenging for most LC modes/types and irrelevant for many applications.

Instead, chromatographers estimate the overall mobile phase accessible volume, which is referred to as the mobile phase hold-up volume (V_M). The theoretical definition of the V_M volume is defined by equation 4.9, where the accessible volumes are measured as a cumulative variable. Experimentally, it can be defined as the time needed for the mobile phase to carry an ideally unretained molecule to migrate from the inlet to the outlet of the column, referred to as hold-up time or dead time (t_M). Multiplying t_M by the flow rate F , yields V_M :

$$V_M = V_i + V_p = V_c - V_{SP} - V_{CP} - V_{ligand} \quad (4.9)$$

It is important to note that the term void volume (V_0) has been used interchangeably with V_M , however, this term is considered an obsolete term as the IUPAC defines it as “the volume increase to the total available volume to the mobile phase flow by all tubing and connections” [92, 93]. In this case, extra-column volume (V_{EC}) is preferred to avoid confusion.

The V_M value is very important as it is used to correct experimental retention data, isolate the separation phenomena to the stationary phase, and minimise inaccessible volumes inside the chromatographic column. In ideality, corrected retention data remains constant within systems regardless of the scale, assuming the media in the chromatography column is the same. However, due to the natural porosity of most RPLC packing materials, the solvation of the bonded phase causes deviations in V_M measurements.

As chromatography solid media interact with the mobile phase, solvation occurs, and solvated ligand chains and solvent molecules can create a 'repulsive' layer at the entrance of the pores, obstructing access to the analytes/solutes and thus modifying the hold-up volume. This phenomenon indicates that different solutes in the mobile phase could experience different hold-up volumes depending on the physicochemical properties of each molecule. Therefore, the hold-up volume has been divided into kinetic and thermodynamic components. The former is confined to the volume of the moving mobile phase and does not contribute to retention. In other words, it is the volume required to transport a solute through the column without interacting with the stationary phase (ideal non-retained volume). In contrast, the latter term accounts for the volume of the mobile phase required to reach the equilibrium between the mobile and stationary phases. It considers the interactions between the solute and the stationary phase, reflecting the time it takes for the solute to equilibrate between the two phases. In summary, the kinetic hold-up volume is more about the physical transport of the solute, and the thermodynamic hold-up volume involves the equilibrium and interactions between the solute and the stationary phase. With these definitions, the thermodynamic hold-up volume is considered dependent on the size of the analyte; therefore, it varies among molecules, and, contrarily, the kinetic hold-up volume is independent of the solute since there is no static adsorption portion[94, 95].

Due to the high complexity surrounding the determination of the hold-up volume in RPLC, quantification techniques are divided into static methods that operate with an immobile liquid phase and dynamic methods whose measurements are performed while the mobile phase percolates through the column at a constant velocity[96].

The most popular static measurement is pycnometry-based. It is also called the weight difference method, which consists of weighing the LC column when it is sequentially filled with solvents of different densities. Then, the pycnometric void volume can be calculated by the weight difference using the following equation[97]:

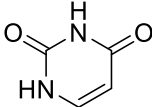
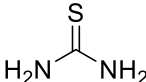
$$V_0^{\text{pyc}} = \frac{m_x - m_y}{\rho_x - \rho_y} \quad (4.10)$$

where m refers to the mass of the column filled with either solvent x or y , and ρ is the density of the individual solvents. Since there are no dynamic interactions during this method, results are considered to be the maximum thermodynamic hold-up volume. Commonly used solvents used for these purposes include methanol, carbon tetrachloride, acetonitrile, and chloroform[98]; usually, these are chosen depending on the difference between the density of the solvents. A larger solvent density difference improves the measurement accuracy of the void volume because the method becomes more sensitive to small changes in the mass-to-density ratio. The maximum density difference is achieved by weighing the column filled with one solvent, followed by a drying stage with helium[99] or nitrogen gas[100]. Although this method provides the most accurate result, a significant problem limiting its popularity is that the gas drying stage could potentially damage the packing material, making the measurement flawed.

Dynamic methods to approximate hold-up volumes are more common in routine chromatographic operations than pycnometry. With their adaptability and versatility, these dynamic methods have distinct advantages over static techniques. The presence of the mobile phase includes the possible solvation effects, thereby eliminating the risk of overestimating the true practical hold-up volume. The simplicity of the methods allows for fast analysis with a non-destructive impact on the packing material, reassuring their applicability in various chromatographic operations.

The most common strategy requires injecting an unretained neutral marker molecule into the mobile phase and subsequently recording the neutral marker molecule as the first negative (or positive peak). However, there are uncertainties around how reliable is the assumption about the unretained natural marker molecule being completely unretained. Strictly speaking, no true unretained solutes are possible as any molecule different to the mobile phase will exhibit different chemical interactions with the stationary phase relative to the solvent phase present. An ideal unretained molecule must be small enough to access all accessible volumes whilst also being perfectly hydrophilic and hydrophobic balanced to avoid solvation with the stationary phase. Some molecules (Table 4.1) have been proposed for these determinations, but acetone and uracil are most commonly used for most chromatographic modes. However, the elution volume of these small molecules has been shown to depend on the mobile phase's composition and temperature. Retention volumes tend to increase with decreasing mobile phase strength and temperatures[101, 102].

Table 4.1 | Proposed unretained molecules for liquid chromatography. Small molecules reported in the literature for V_M measurements.

Name and structure	Advantages	Disadvantages
 Acetone	Good UV response due to $\pi \rightarrow \pi^*$ transition at the double bond ($\lambda_{\max,1} = 195 \text{ nm}$) and $n \rightarrow \pi$ transition at the non-bonding electrons ($\lambda_{\max,2} = 274 \text{ nm}$).	60-70% error compared to pycnometric measurements[102].
 Uracil	Good UV response due to double bonds ($\lambda_{\max,1} = 202 \text{ nm}$) and ring ($\lambda_{\max,2} = 258 \text{ nm}$).	20-30% error compared to pycnometric measurements[102].
 Thiourea	Good UV absorption at $\lambda_{\max} = 236 \text{ nm}$.	Tailing of peaks in the middle range of organic solvent composition (60-95%)[103].
 Deuterium oxide (D_2O)	Similar or lower retention than uracil[104].	Detection of this molecule is only achieved through a refractive index detector[104].
Inorganic and organic salt solutions Particularly $NaNO_3$ and KBr	- Simple and easy to detect via UV. - Size range allows multiple interpretations.	Inactive silanol groups could exclude charged species from the pores of the resin (Donnan exclusion)[98, 105].

Two more complex approaches can produce more accurate V_M values; measuring the mobile-phase excess isotherms and using a homologous chemical series of analytes. The former technique will be fully described in the isotherm Section 4.6.1 of this Chapter. The latter method involves plotting the logarithm of the retention time and homologous number, i.e., carbon number and fitting a linear relationship[106]. The intersection with the y-axis, where the number of carbons equals zero, is then argued to represent the column's hold-up volume.

This second procedure is more commonly used in gas chromatography, providing accurate results using linear hydrocarbons between C6 and C12 in length. Therefore, many attempts to

incorporate the homologous series method into RPLC have been investigated. The main challenge of this method is the appropriate selection of homologous molecule family that would provide enough information to calculate the value of V_M accurately. A minimum of four chemical homologues is necessary to extrapolate the trendline to void volume correctly[107]. The homologues must be chosen according to practical characteristics like solubility, detectability, and availability. Deviations in the linearity of the correlations reflect multiple chemical interactions between the analytes and the stationary phase other than hydrophobic interactions between the analytes and the stationary phase.

The inaccuracy of the experimental estimations of the hold-up volume causes most of the RPLC errors related to volume and retention time adjustments. Regardless of the method chosen to determine the void volume, it is relevant to mention that they all have different assumptions depending on the chromatographic mode. In the case of RPLC, it is assumed that the primary partitioning mechanism of the solutes governs the separation mechanisms. However, other secondary retention effects, such as hydrogen bonding between the solutes and the unreacted surface silanol present in the stationary phase, solute adsorption onto the surface of either the silica or the bonded phase, and size exclusion mechanisms depending on the pore size, can affect the hold-up volume estimates.

4.2 Mobile Phase Elution Profiles

Mobile phase elution profiles refer to the changes in mobile phase composition with respect to time elapsed or volume eluted. These changes are generally achieved by mixing two or more solvents at the quaternary pump and are, therefore, represented as a volume fraction or percentage relative to line A, the primary mobile phase solvent.

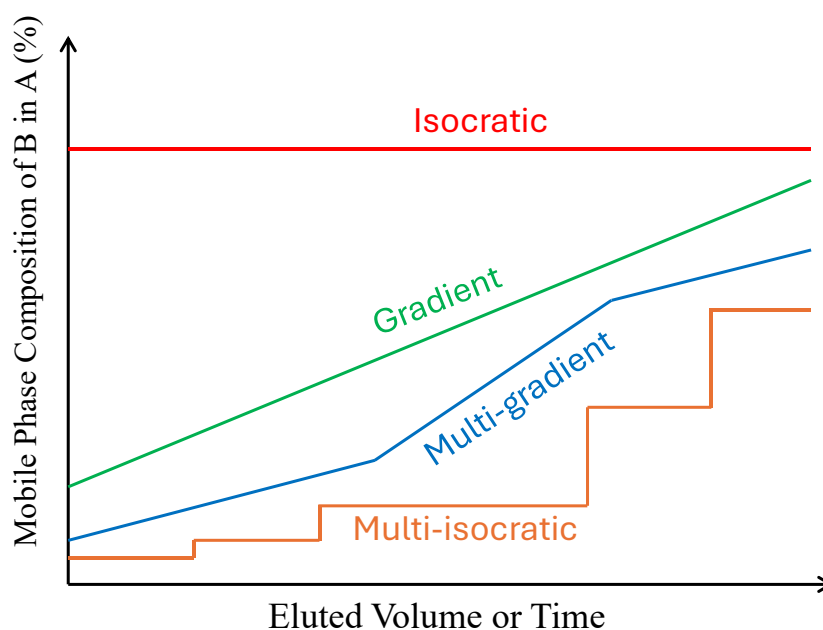


Figure 4.4 | Mobile phase elution profiles used in chromatography.

4.2.1 Isocratic

In isocratic elution, the composition of the mobile phase remains constant throughout the separation process. This type of elution is simple to set up and reproduce, making it easier to control and replicate, leading to consistent results.

In RPLC, this elution profile is often used for separating analytes with similar polarities or when the sample components are well separated under a single mobile phase condition, making it suitable for routine analyses and quality control. However, it may not be ideal for complex mixtures with components with a wide range of polarities, as some components may elute very slowly or not at all, resulting in long run times and poor resolution. Indeed, elution times can increase

exponentially for significantly different analytes, making the sole use of isocratic profiles impractical for many separation processes.

4.2.2 Gradient Elution Profiles

In contrast to isocratic elution, the composition of the mobile phase changes continuously or in steps during gradient elution. Gradient elution can help elute compounds that would otherwise be difficult to separate, improve peak resolution, and shorten run times by gradually increasing or decreasing the strength of the mobile phase, typically by altering the proportion of organic solvent. This highly versatile approach can be fine-tuned to optimise the separation of specific analytes and is the dominant approach used for research and production scale RPLC. However, it requires a quaternary pump and careful method development to ensure reproducibility and accuracy. A comparison between the two types of elution profiles can be seen in Table 4.2.

Table 4.2 | Comparative of gradient and isocratic elution methods. Adapted from Schellinger and Carr[108].

Metric	Gradient elution	Isocratic elution
Retention factor (k) range	$k \gg 15$	$1 < k < 15$
Peak capacity	High (10 or more solutes)	Low (less than 10 solutes)
Method transfer	Fair	Excellent
Quantitation (baseline disturbance)	Background adjustment is necessary for higher gradient slopes.	Baseline disturbance is rare.
Re-equilibration time	Slow	Instant
Method development	Poor	Good
Instrumentation simplicity	Requires binary or quaternary pump.	Standard LC system

4.2.3 Complex

Multi-isocratic Elution Profiles

Multi-isocratic elution combines the principles of isocratic and gradient elution by using a series of different isocratic steps, where the mobile phase composition remains constant within each step but changes between steps. This approach offers better separation for complex mixtures than a single isocratic run while maintaining some of the simplicity and reproducibility of traditional isocratic elution. During multi-isocratic elution, the mobile phase composition is kept constant for a defined period or volume, and then it is changed to a new, constant value for the next step. This method allows for improved separation of mixtures with components with a range of polarities, as each step can be optimised for specific groups of analytes. While it is not as straightforward as single-isocratic elution, multi-isocratic elution is generally simpler to implement and reproduce than gradient elution, allowing for individual optimisation of each step and simplifying method development. This technique is particularly useful in quality control and routine analyses where reproducibility and simplicity are crucial, offering a balanced approach between ease of use and enhanced resolution for complex mixtures.

Multi-gradient

Multi-gradient elution employs several distinct gradient segments, each tailored to enhance the separation of specific groups of analytes. This allows for greater flexibility and precision in manipulating the mobile phase strength to resolve compound groups with varying polarities and affinities for the stationary phase.

By strategically designing these gradient segments, chromatographers can improve peak resolution, reduce analysis time, and enhance overall separation efficiency. Multi-gradient elution is particularly useful in large-scale chromatography.

4.3 Retention parameters

4.3.1 Retention factor k

The retention factor (k), also known as the retention factor or the partition ratio, represents the equilibrium between the number of molecules of a solute in the stationary phase (N_{SP}) and the number of molecules in the mobile phase (N_{MP}). The number of molecules can also be expressed in terms of solute concentration in the volume of each phase ($N_{SP} = c_{SP}V_{SP}$ and $N_{MP} = c_{MP}V_{MP}$). This transformation allows the chromatographic system to be described based on the analyte distribution coefficient ($K = \frac{c_{SP}}{c_{MP}}$) and the phase ratio of the column ($\Phi = \frac{V_{SP}}{V_{MP}}$). All these definitions are contained within equation 4.11. Note that the volume of the mobile phase (V_{MP}) can also be written as the hold-up volume (V_M):

$$k = \frac{m_{SP}}{m_{MP}} = \frac{c_{SP}V_{SP}}{c_{MP}V_M} = K\Phi \quad (4.11)$$

Since the chromatographic column is a closed system, it is difficult to measure k values based on the previous definitions. The main difficulty resides in measuring the concentration at, or the volume of, the stationary phase as this information is not available and, in some chromatographic modes, it is not well defined. However, it is possible to find k values based on the total time (t_r)/volume (V_r) a sample component travels through the packed chromatographic column in relation to the hold-up volume (equation 4.12). Note from Figure 4.5 that t_r and V_r are measured at the first moment (mean; μ) of the Gaussian curve, i.e., the apex of the peak:

$$k = \frac{V_r - V_M}{V_M} = \frac{t_r - t_M}{t_M} \quad (4.12)$$

A higher k value indicates that a more significant proportion of the solute molecules interact with or are retained by the stationary phase than those in the mobile phase. Conversely, a lower k value indicates that more molecules are in the mobile phase and fewer are retained in the stationary phase. In other words, the capacity factor reflects the strength of the interaction between

the analyte component and the stationary phase and, to some respect, isolated from the effects of the mobile phase.

It is important to mention that the capacity factor k is correctly defined during isocratic elution based on the equilibrium definition. Other definitions for gradient elution will be described in Chapter 10. However, most experimentalists report k -values for simplicity regardless of the elution mode.

4.3.2 Selectivity factor

The separation factor, commonly called the selectivity factor, describes the relative separation between two solutes. It measures how well the chromatographic system can distinguish between two different compounds. The selectivity factor is defined as the ratio of the retention factors (k) of two solutes, typically denoted as solute X and solute Y:

$$\alpha = \frac{k_Y}{k_X} \text{ when } k_Y > k_X \quad (4.13)$$

When $\alpha > 1$, the chromatographic system can separate the two solutes. The higher the value, the better the separation. When $\alpha = 1$, the two solutes have the same retention factor; therefore, they cannot be separated by the chromatographic system. Negative values are not theoretically possible as α is defined such that the more strongly retained solute is in the numerator.

The calculation of α is practical as it provides a quick evaluation of the separation efficiency and peak overlapping. However, as it is only calculated from the retention factor and does not consider peak sharpness, it does not provide the complete picture of the separation process. Therefore, the information provided by this value is complementary to the following two additional retention parameters.

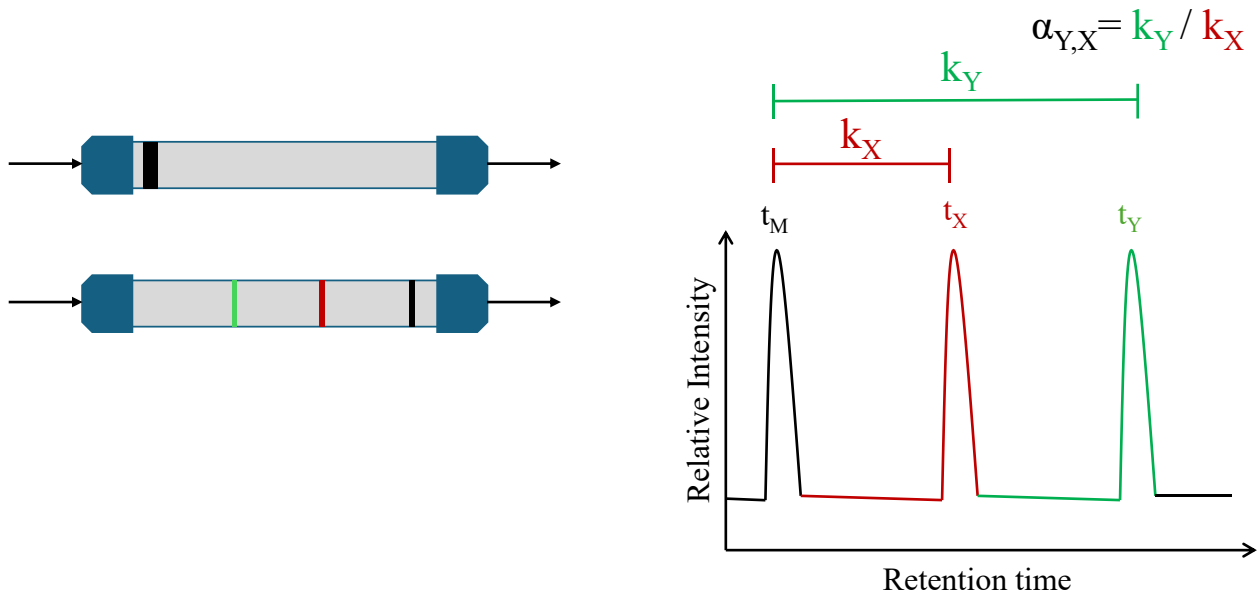


Figure 4.5 | Graphical summary of chromatographic retention parameters.

4.3.3 Number of theoretical plates

The number of theoretical plates (N) describes the efficiency of a separation column. It measures how well a column can separate different compounds and indicates its performance. The term "theoretical plate" is derived from distillation theory, where it represents a hypothetical stage in which the solute reaches equilibrium between the mobile and the stationary phases. As observed from equation 4.14, the calculation of N assumes a peak in the shape of a well-defined Gaussian curve where the ratio between the first (mean; μ , t_r or V_r) and the second (standard deviation; σ) moments of the curve are the primary descriptors.

$$N = \left(\frac{t_r}{\sigma^t} \right)^2 = \left(\frac{V_r}{\sigma^V} \right)^2 \quad (4.14)$$

The standard deviation can be expressed as the width of the peak based on different fractions of the peak height (H) and as a ratio between the peak height and area (H/A):

$$N = 16 \left(\frac{t_r \text{ or } V_r}{W_{0.1H}} \right)^2 = 5.54 \left(\frac{t_r \text{ or } V_r}{W_{0.5H}} \right)^2 = 2\pi \left(\frac{t_r \text{ or } V_r \times H}{A} \right)^2 \quad (4.15)$$

A common practice is to express column efficiency as height equivalent to a theoretical plate (HETP), which converts the efficiency measured by N into a physical dimension dependent on the length of the column (L) (equation 4.16). This provides a standardised measure that makes it easier to compare the efficiency of columns with different lengths:

$$\text{HETP} = \frac{L}{N} \quad (4.16)$$

Since N and HETP are inversely proportional, they have very similar interpretations. Smaller HETP (high N) values indicate higher column efficiency, as each theoretical plate is shorter, meaning the column can produce sharper peaks in a shorter distance. Larger HETP (low N) values indicate lower column efficiency (broader peaks), suggesting that the column requires a greater length to achieve the same level of separation. These parameters are affected by column length, particle size, mobile phase flow rate, column packing quality and temperature.

4.3.4 Resolution

Resolution (R_S) is very similar to α . It quantifies the degree to which two peaks are distinct and separate from each other. High resolution indicates that the peaks are well separated, while low resolution suggests that the peaks may overlap:

$$R_S = \frac{2(t_{r,2} - t_{r,1})}{W_1 + W_2} = \frac{1.18(t_{r,2} - t_{r,1})}{W_{0.5H,1} + W_{0.5H,2}} \text{ when } t_{r,2} > t_{r,1} \quad (4.17)$$

When R_S is equal to unity, peaks are close and may overlap slightly, but individual compounds can often be distinguished. Values higher than 1.5 are typically considered baseline resolution, indicating that the two peaks are well-separated and can be quantified accurately. Any peaks with R_S values lower than 1 have poor separation. Such peaks overlap significantly, making it difficult to distinguish between the compounds.

Resolution is particularly useful because it can also be expressed in terms of all previously described retention parameters (Equation 4.18). This allows chromatographers to optimise separations based on the R_S value. For example., to improve resolution, N can be increased by using a longer column, smaller particle size, or optimising the mobile phase flow rate. The selectivity can be changed by adjusting the composition of the mobile phase, temperature or stationary phase., whereas adjusting the mobile phase strength or column characteristics can optimise k :

$$R_S = \frac{\sqrt{N}}{4} \left(\frac{\alpha - 1}{\alpha} \right) \left(\frac{k_2}{1 + k_2} \right) \text{ when } k_2 > k_1 \quad (4.18)$$

4.3.5 Symmetry factor

The symmetry factor (S), or tailing factor (T), describes the shape of a chromatographic peak. It provides information about the symmetry of the peak, indicating whether the peak is symmetrical or exhibits tailing or fronting. In equation 4.19, $W_{0.05H}$ is the width of the peak at 5% of the peak height and f is the distance from the peak apex to the leading edge (left side) of the peak at 5% of the peak height:

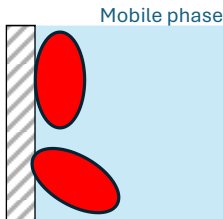
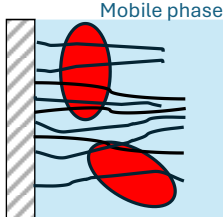
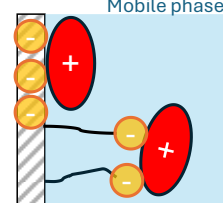
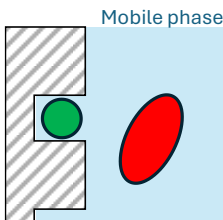
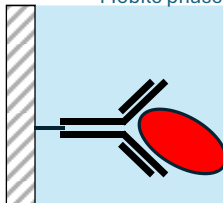
$$S = \frac{W_{0.05H}}{2f} \quad (4.19)$$

S values equal to 1 indicate a perfectly symmetrical peak. Any other values indicate tailing ($S > 1$), where the peak has a longer tail on the right side, or fronting ($S < 1$), where the peak has a longer front on the left side. This parameter is affected by the column packing quality, type of interaction with the stationary phase, mobile phase composition, overloading and detector response.

4.4 Modes/Types of Liquid Chromatography

Liquid chromatography modes are classified depending on the retention mechanisms that govern the separation process. These mechanisms are (1) adsorption, (2) partition, (3) ion exchange, (4) size exclusion and (5) affinity[109].

Table 4.3 | Summary of types of liquid chromatography.

	Diagram	Retention	Elution
Adsorption		Competition between the analyte and the mobile phase to bind into the surface of the support.	The elutropic strength overcomes the binding strength of the analyte to the support.
Partition		Competition between the energies of separation of the analyte between the liquid mobile phase and the support coated/bounded with a ligand (stationary phase).	The solubility of the analyte in the mobile phase is thermodynamically preferred over the binding to the ligand.
Ion exchange		Interaction/Exchange of ions in the analyte with ionic groups of opposite charge fixed to the support.	A counter-ion is added to the mobile phase, which displaces and releases the analyte.
Size exclusion		The analyte travels through a network of pores of different sizes in the support. This creates a sieving effect, separating analytes by size, specifically hydrodynamic volume.	Constant system. The mobile phase carries the analyte constantly. Large analytes elute faster as they get excluded from some of the pores.
Affinity		Selective and reversible chemical reaction between the analyte and an immobilised binding agent (affinity ligand) bound to the support.	The analyte is released by disrupting the affinity between the affinity ligand and the analyte. This is done with the aid of an elution buffer at a certain pH or enriched with a bio-specific competing agent or a chaotropic agent.

4.5 Adsorption and Partition Processes on RPLC

RPLC is particularly interesting due to the separation process conditions. As introduced in the previous section, the most studied interaction in RPLC is the partitioning effect of the analytes/solutes between the stationary and surrounding mobile phases. However, both adsorption and partitioning take place during the RPLC separation[110]. The initial interaction of the solute molecules with the hydrophobic ligands immobilised onto the surface of the stationary phase involves adsorption on the surface of the stationary phase, specifically at the upper level of the bonded hydrophobic ligands. Once adsorbed, the solute molecules partition between the hydrophobic surface ligand layer and the mobile phase[111, 112]. This continuous exchange determines the overall retention and separation of the solutes.

4.5.1 Isotherms

An adsorption isotherm is a graphical representation that describes the equilibrium relationship between the concentration of a solute/adsorbate in the stationary phase/adsorbent and its concentration in the mobile phase at a constant temperature. This distribution affects how the solute migrates through the chromatographic system.

From a surface chemistry perspective, isotherms aid the study of adsorption mechanisms, maximum adsorption capacities and properties of the adsorbents. Isotherms are crucial for optimising chromatographic separations as they help predict how solutes behave under different conditions and can guide the selection of appropriate stationary and mobile phases and other operational parameters. They are the key determinants of equilibrium adsorption behaviour and thermodynamics.

Types of adsorption isotherms

In RPLC, the shapes of adsorption isotherms can be divided into 4 groups:

- **Convex (favourable):** These exhibit a convex Langmuirian shape, indicating that a large amount of adsorption on the adsorbent (stationary phase) is reached at low adsorbate concentrations in the mobile phase.
- **Concave (unfavourable):** These have a concave anti-Langmuirian shape, requiring high concentrations of adsorbate in the mobile phase to achieve significant adsorption on the stationary phase.
- **S-shape:** These are a combination of the other two types. As the adsorbate concentration in the mobile phase increases, adsorption on the stationary phase becomes more accessible.
- **Linear:** This describes a direct, proportional relationship between the concentration of a solute in the stationary and mobile phases at equilibrium. This type of isotherm could be an isolated group, as most isotherms start as linear adsorption if the concentration at the mobile phase is kept low.

The shape of the isotherm influences the chromatographic peak shape[113] (see Figure 4.6). A linear isotherm generally leads to Gaussian-shaped peaks. This behaviour is because the solute distribution between the mobile and stationary phases remains constant as the solute concentration changes. A convex isotherm often results in tailing peaks. At higher concentrations, the solute molecules have fewer available adsorption sites, causing them to desorb more slowly, which leads to tailing. Concave isotherms tend to produce fronting peaks. As the solute concentration increases, the solute molecules adsorb more readily, causing the front of the peak to move faster than the tail. S-shaped isotherms reflect complex adsorption behaviour, resulting in non-standard peak shapes, combining elements of both fronting and tailing. Dealing with this type of isotherm is challenging but, fortunately, is rare in RPLC.

For all RPLC scales, the ideal adsorption isotherm is typically linear due to its simplicity, predictability, and ease of use. Linear isotherms lead to symmetrical peaks, consistent retention times, and straightforward method development, making them highly suitable for precise analytical work and scalable preparative processes. However, when they occur, understanding and managing non-linear isotherms is essential for optimising chromatographic separations across various applications, especially process-scale chromatography. The next Chapter discusses non-linear chromatography and its applications on a manufacturing scale.

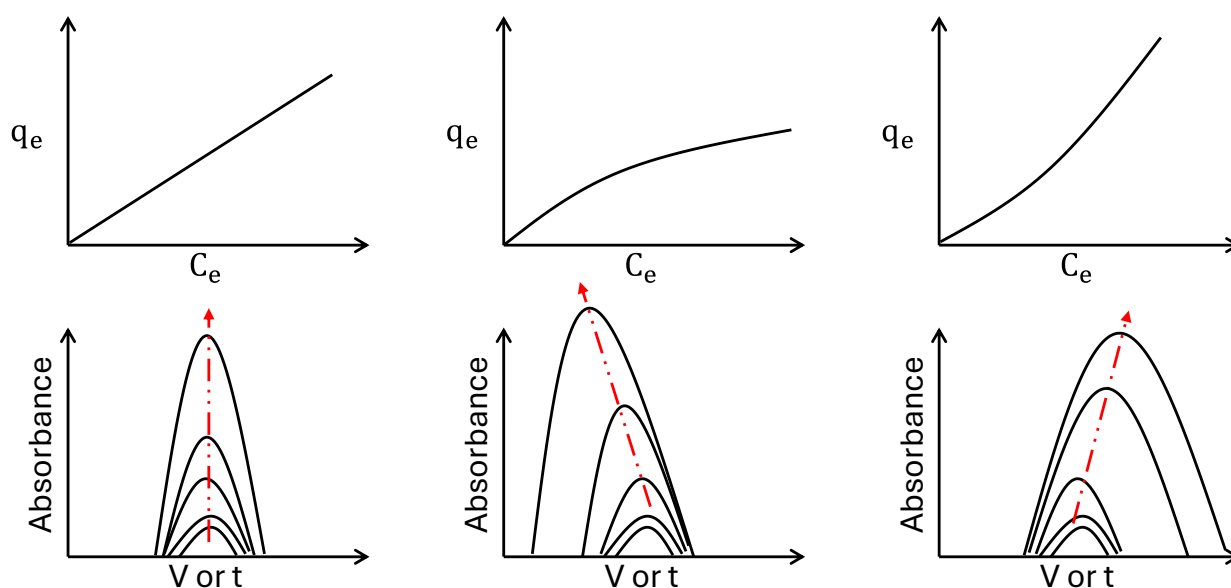


Figure 4.6 | Effect of the adsorption isotherm shape on chromatographic peak shape.

The IUPAC has a more detailed general classification of adsorption isotherms divided into 6 types[45]. Although this classification system is designed for application to gas-solid adsorption systems, it can also be applied to adsorption in liquid-solid systems to understand the mechanisms of the adsorption process and its relationships to the isotherm shape. These 6 adsorption isotherm types are depicted in Figure 4.7.

- **Type I:** Monolayer adsorption on microporous materials. Adsorption on a surface with a finite number of identical sites leads to monolayer coverage.
- **Type II:** Multilayer adsorption on nonporous or microporous materials.
- **Type III:** Weak adsorbate-surface interaction without saturation. The adsorbate-adsorbate interactions are stronger than adsorbate-surface interactions.
- **Type IV:** Multilayer adsorption with capillary condensation in mesoporous materials. Initial adsorption follows a Type II pattern, leading to capillary condensation. This capillary condensation forms a hysteresis loop.
- **Type V:** Weak adsorbate-surface interaction with capillary condensation in mesoporous materials. Weak initial adsorption (Type III) followed by enhanced adsorption at higher pressures. Exhibits hysteresis due to capillary condensation.
- **Type VI:** Stepwise multilayer adsorption on uniform nonporous surfaces. Represents adsorption on a uniform nonporous surface, where distinct adsorption steps correspond to the

formation of successive layers. Each step reflects the completion of a monolayer before the next begins.

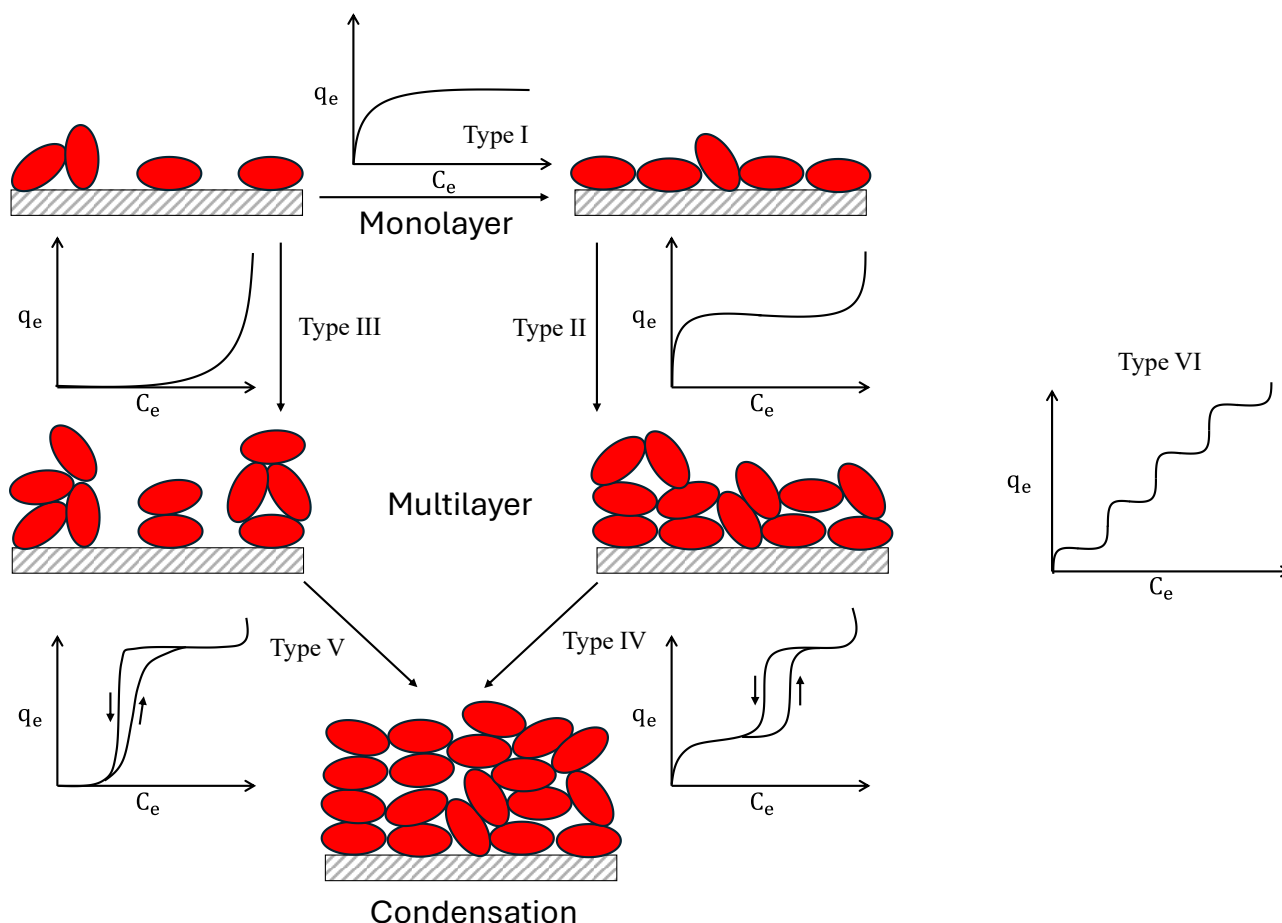


Figure 4.7 | IUPAC classification of gas phase adsorption isotherms. See text for description.

After experimentally measuring isotherms, they are commonly modelled through mathematical isotherm equations. The equations range from fully empirical equations to quite sophisticated mathematical models with a strong fundamental basis. Modelling is particularly useful because it allows extrapolation and interpolation of the experimental data and a deeper understanding of the isotherm by extracting key adsorption descriptors, such as maximum adsorption capacity, used to compare the performance of different adsorbents.

Many authors have used modelling to obtain mechanistic information about the adsorption behaviour and the isotherm shape. However, unlike the detailed comparable studies of gas-solid adsorption processes, the physical meaning of most isotherm models has not been thoroughly studied for liquid-solid adsorption processes.

Table 5.3 summarises the most commonly cited models in the literature. The Langmuir isotherm is the most frequently used model to fit liquid-solid chromatographic separations, including pharmaceuticals. It is a theoretical model of a monolayer chemical adsorption that describes macroscopic homogenous adsorption systems. Empirical models are fit-to-purpose equations lacking physical meaning. The Brunauer–Emmett–Teller (BET) model is a theoretical multi-layer homogenous adsorption model, where the adsorption energy in the first adsorption layer differs from the energy associated with the subsequent multilayers.

Wang and Guo[114] reported an intensive analysis of 13 models. They concluded that all studied empirical models (linear, Freundlich and Sips models) are derived from the Langmuir theory. Also, they noticed that the frequency of fitted models in the literature is Langmuir > Freundlich > Temkin, Sips, linear, Redlich-Peterson > Others.

Table 4.4 | Summary of adsorption isotherm models. In all models q_e and C_e are the equilibrium concentration of the adsorbate in the stationary phase (mg_a/g_{SP}) and the mobile phase (mg_a/L_{MP}), respectively.

Model	Equation	Constants
Linear (Henry's Law)	$q_e = KC_e$	K ($L_{MP} \cdot g_{SP}^{-1}$) – partition coefficient
Freundlich	$q_e = K_F C_e^{1/n}$	K_F ($L_{MP}^{1/n} \cdot mg_a^{1-1/n} \cdot g_{SP}^{-1}$) – partition coefficient n – constant
Redlich-Peterson	$q_e = \frac{K_{RP} C_e}{1 + a_{RP} C_e^g}$	K_{RP} ($L_{MP} \cdot g_{SP}^{-1}$) – partition coefficient a_{RP} ($L_{MP}^g \cdot mg_a^{-g}$) – constant g – exponent
Langmuir	$q_e = \frac{q_m K_L C_e}{1 + K_L C_e}$	K_L ($L_{MP} \cdot mg_a^{-1}$) – ratio of the adsorption and desorption rates q_m ($mg_a \cdot g_{SP}^{-1}$) – maximum adsorption capacity
Brunauer–Emmett–Teller (BET)	$q_e = \frac{q_{mBET} K_{BET1} C_e [1 - (n_{BET} + 1)(K_{BET2} C_e)^{n_{BET}} + n_{BET} (K_{BET2} C_e)^{n_{BET}+1}]}{(1 - K_{BET2} C_e) \left[1 + \left(\frac{K_{BET1}}{K_{BET2}} - 1 \right) K_{BET2} C_e - \left(\frac{K_{BET1}}{K_{BET2}} \right) (K_{BET2} C_e)^{n_{BET}+1} \right]}$	K_{BET1} and K_{BET2} ($L_{MP} \cdot mg_a^{-1}$) – adsorption equilibrium parameters in the first and upper layers q_{mBET} ($mg_a \cdot g_{SP}^{-1}$) – maximum adsorption capacity n_{BET} – constant
	$q_e = \frac{q_{mBET} C_{BET} C_e}{(C_s - C_e) \left[1 + (C_{BET} - 1) \frac{C_e}{C_s} \right]}$	C_{BET} ($mg_a \cdot g_{SP}^{-1}$) – maximum adsorption capacity q_{mBET} ($mg_a \cdot g_{SP}^{-1}$) – maximum adsorption capacity C_s ($mg_a \cdot L_{MP}^{-1}$) – solubility of the absorbate

4.5.2 Inverse chromatographic methods to measure isotherms

Static adsorption isotherm measurements

Closed-vessel experimental adsorption systems with direct static measurement of adsorbate concentrations is the most common and practical technique to measure adsorption isotherms for any liquid-solid system[115-118]. This involves preparing a series of adsorbent and adsorbate solutions, mixing them to reach equilibrium, and then analysing the residual adsorbate concentration to determine the adsorption capacity. This process starts by accurately weighing a known mass of adsorbent ($m_{\text{adsorbent}}$) and adding it to beakers containing different initial concentrations of adsorbate (C_0) solutions of known volume (V). After allowing the mixtures to equilibrate, the adsorbent is separated by filtration or centrifugation, and the residual adsorbate concentration C_e is measured using techniques like spectrophotometry or chromatography. The equilibrium adsorption capacity q_e is calculated using equation 4.20 and plotted against the equilibrium concentration to create the isotherm (q_e).

$$q_e = \frac{(C_0 - C_e)V}{m_{\text{adsorbent}}} \quad (4.20)$$

While the beaker method for measuring adsorption isotherms is a straightforward and convenient laboratory technique, it is not typically the most practical or efficient method for studying adsorption on solid-liquid chromatographic systems. Chromatography involves dynamic processes wherein solutes continuously interact with the stationary and mobile phases as they move through the column, which differs from the static equilibrium conditions studied in a beaker. RPLC resins are even more difficult to measure in stationary/static conditions as they do not disperse well in solution. The lack of added pressure, like the one produced in the LC system, causes phase hydrophobic collapse or dewetting in the pores[119].

Inverse liquid chromatography

Unlike traditional chromatography, which aims to describe the properties of analytes or to separate complex analyte mixtures by injecting the sample into multiple commercial columns, inverse liquid chromatography (ILC) consists of injecting one or multiple known molecules, called probes, into commercial chromatographic columns to elucidate their physicochemical adsorption properties or characteristics.

These experiments are performed in an LC system, and the amount of analyte or adsorbate adsorbed by the column's adsorbent is indirectly measured based on the detector's signal. Since LC detectors measure concentration changes in the eluent post-column and not in the stationary phase, mobile phase conditions must be selected to allow complete analyte/solute desorption. In other words, the stationary-mobile phase system should always return to initial conditions after the perturbation with the adsorbate. This is equilibrium in inverse chromatography.

The following inverse chromatographic techniques can be used to study LC adsorption isotherms and solute behaviour.

Frontal analysis (FA)

Frontal analysis consists of continuously introducing the solute, with a fixed and known concentration in the mobile phase, into a chromatographic column until the outlet concentration equals the inlet concentration. This creates a breakthrough curve that can be analysed to extract adsorption data[120]. This method is generally performed by premixing the mobile phase to a concentration of organic solvent that desorbs/elutes the analyte completely. This premixed mobile phase is split into two LC lines, for example, A and B, as seen in Figure 4.8. To reduce experimental time, the adsorbate is dissolved in line A at a known concentration typically close to its maximum solubility in the mobile phase. The valves are changed dynamically between A and B throughout the experiment to test adsorption and desorption. Mixing A and B during the adsorption step can also test other concentrations.

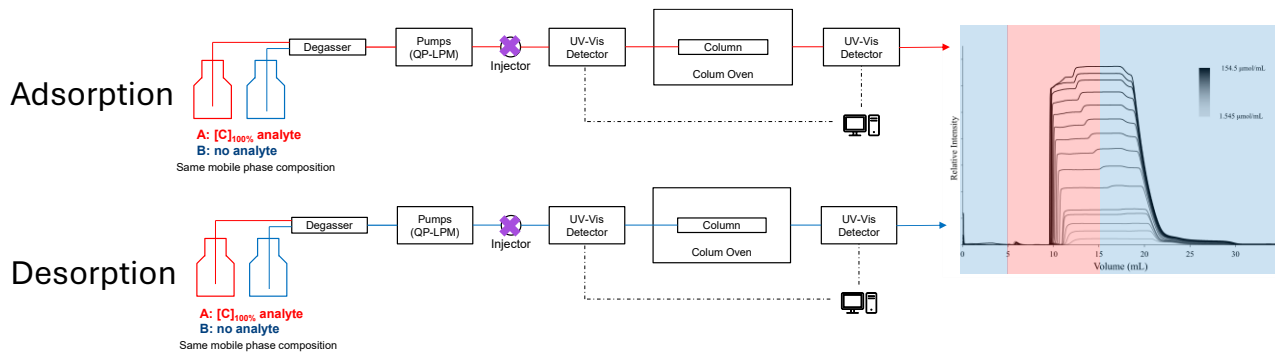


Figure 4.8 | Experimental set-up for frontal analysis.

When the valve changes to line A, the adsorbate is gradually taken up by the adsorbent until saturation at the set mobile phase condition is reached. At this first point of equilibrium, the detector reaches a plateau corresponding to the signal of the concentration of adsorbate in line A ($V_{e,1}$). When the valve returns to line B, the adsorbate is released as it desorbs from the adsorbate and dissolves in the mobile phase, directly reflected at the detector. The second equilibrium is reached ($V_{e,1}$), once the signal returns to baseline. For both the adsorption and desorption processes, the amount adsorbed or desorbed by the adsorbate is measured by multiplying the net volume (V_n) of the mobile phase needed to reach its respective equilibrium by the concentration of analyte ($C_{\text{adsorbate}}$) during the step (equation 5.21).

$$q_e = \frac{n_{\text{ads}}}{m_{\text{adsorbent}}} = \frac{C_{\text{adsorbate}} V_n}{m_{\text{adsorbent}}} = \frac{\text{Area}}{m_{\text{adsorbent}}} \quad (4.21)$$

In the case of desorption, the area to be calculated is simply the area under the curve after the valve returns to line B, which is adjusted by the hold-up volume. During adsorption, the area under the curve is the content of adsorbate in the mobile phase; therefore, the amount adsorbed on the stationary phase is calculated as the area over the curve, with the upper limit at the plateau height once $V_{e,1}$ is reached. Therefore, equations 5.22 and 5.23 can be used for adsorption and desorption, respectively. The chromatographic areas can be observed in Figure 4.9.

$$q_e^{\text{ads}} = \frac{\text{Area}}{m_{\text{adsorbent}}} = \frac{C_e(V_{e,1} - V_M) - \int_{V_M}^{V_{e,1}} c \, dV}{m_{\text{adsorbent}}} \quad (4.22)$$

$$= \frac{F \left[C_e(t_{e,1} - t_M) - \int_{t_M}^{t_{e,1}} c \, dt \right]}{m_{\text{adsorbent}}}$$

$$q_e^{\text{des}} = \frac{\int_{V_M}^{V_{e,2}} c \, dV}{m_{\text{adsorbent}}} = \frac{F \int_{t_M}^{t_{e,2}} c \, dt}{m_{\text{adsorbent}}} \quad (4.23)$$

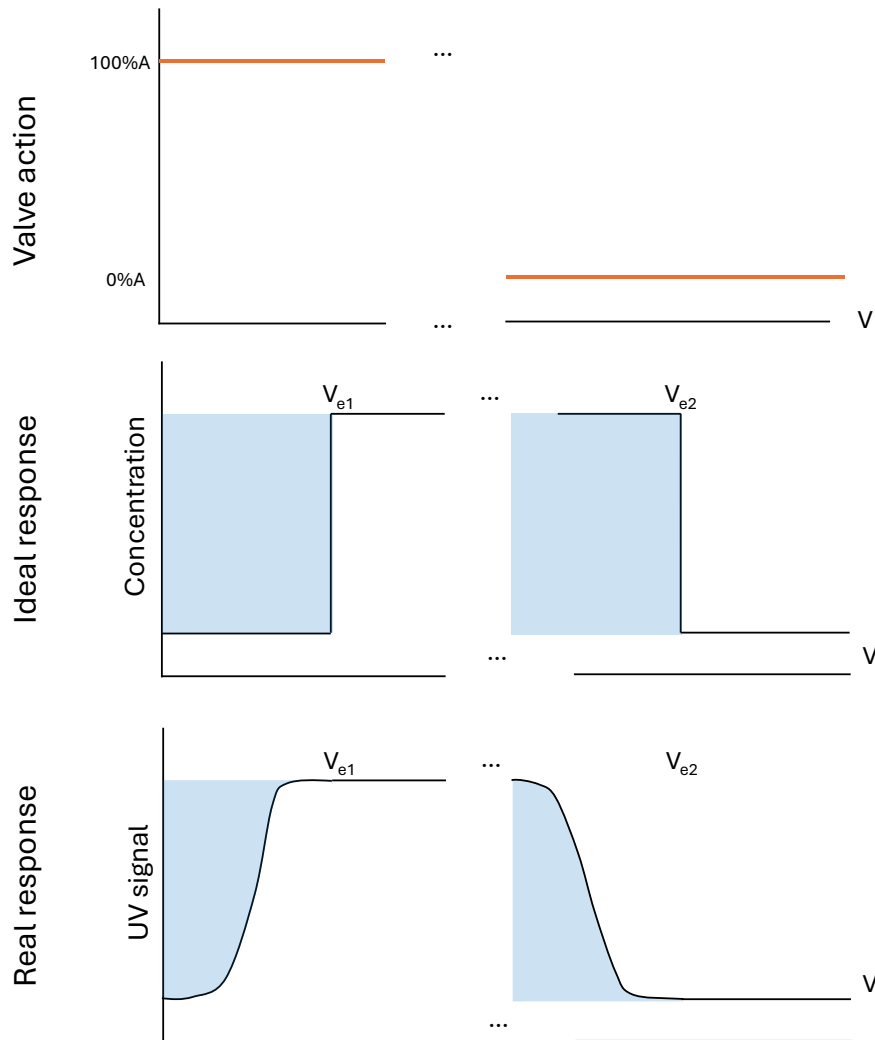


Figure 4.9 | Detector response to frontal analysis experiments.

Frontal analysis is a practical technique as it is easy to set up and interpret the data while maintaining precision and high accuracy. Another advantage is that the continuous flow approach mimics continuous injection, which is similar to real-world manufacturing-scale chromatography, thus leading to better insights into the behaviour of the analytes. This technique can also be used to analyse competitive adsorption by dissolving different analytes in other lines and creating different mixtures through the experiment. In terms of error, this technique is superior to other techniques, as each area calculated for adsorption and desorption is an individual point in the isotherm[121]. However, like any method, it has also disadvantages.

Running FA methods is time-consuming; they require extended runs to ensure equilibration at all stages. Similarly, this makes FA resource-intensive, where high amounts of solvents and pure adsorbate increase experimental costs. In fact, this method might not be preferred for high-value molecules like biomolecules as the amounts required of pure material to create an entire adsorption isotherm are excessive.

Despite being time-consuming and resource-intensive, frontal analysis provides straightforward and often more reliable data for determining adsorption capacity and isotherm shape. It is less sensitive to minor experimental variations and better suited for studying complex adsorption phenomena. On the other hand, the complexity of the elution method in data interpretation and sensitivity to conditions can be significant drawbacks, particularly for researchers new to these techniques or when working with highly variable systems.

Peak maxima (PM)

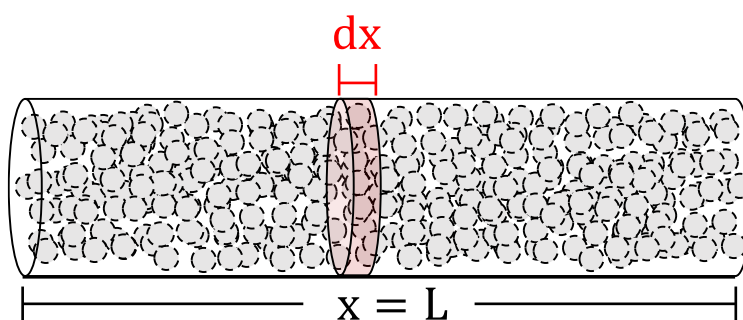


Figure 4.10 | Representation of an infinitesimal chromatographic band migrating through a column.

The peak maxima (PM), pulse, or elution method involves injecting a small, concentrated solute pulse into the chromatographic system and analysing the resulting eluted peak profile. The LC system setup is depicted in Figure 4.11. The analysis of this method is less straightforward than FA. This method is based on the adsorption analysis of an infinitesimal column section representing the injection zone moving through the column, as shown in Figure 4.10 above. The movement of the solute along the column is expressed by equation 4.24, which is a dynamic version of equation 4.12. In this equation, only the peak maxima, thus the name of the technique, is considered, i.e., the first moment of the Gaussian curve or the mean of the distribution. The complete derivation of equation 4.24 can be found in Appendix II.

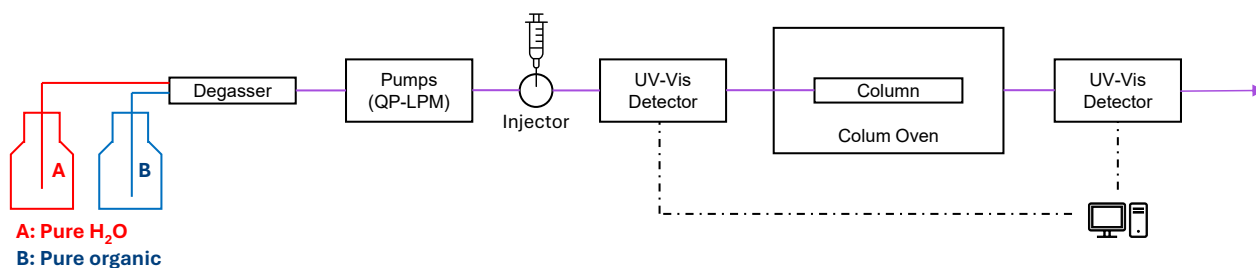


Figure 4.11 | Experimental set-up for peak maxima.

$$V_r = V_M + V_s \frac{dq_e}{dC_e} \quad (4.24)$$

When rearranging the equation, the net retention volume at the apex of the peak (V_n) is equivalent to the rate of adsorption (dq_e/dC_e) (equation 4.25). It is important to notice that adsorption can be measured in terms of any representative units. Since measuring the volume of the stationary phase (V_{SP}) is challenging, calculations can be done based on the mass, volume, or surface area of the stationary phase (A_s):

$$V_n = V_r - V_M = V_{SP} \frac{dq_e}{dC_e} = m_{\text{adsorbent}} \frac{dq_e}{dC_e} = A_s \frac{dq_e}{dC_e} \quad (4.25)$$

The concentration of the analyte in the stationary phase is calculated through numerical integration. This approach is particularly useful as calculations are relatively simple and quick to compute; however, due to the nature of the integration, the error is cumulative as each point in the isotherm is highly dependent on the previous point (Figure 4.12):

$$q_e = \int_0^C \frac{V_n}{m_{\text{adsorbent}}} dC_e = \int_0^C \frac{V_n}{V_{SP}} dC_e \quad (4.26)$$

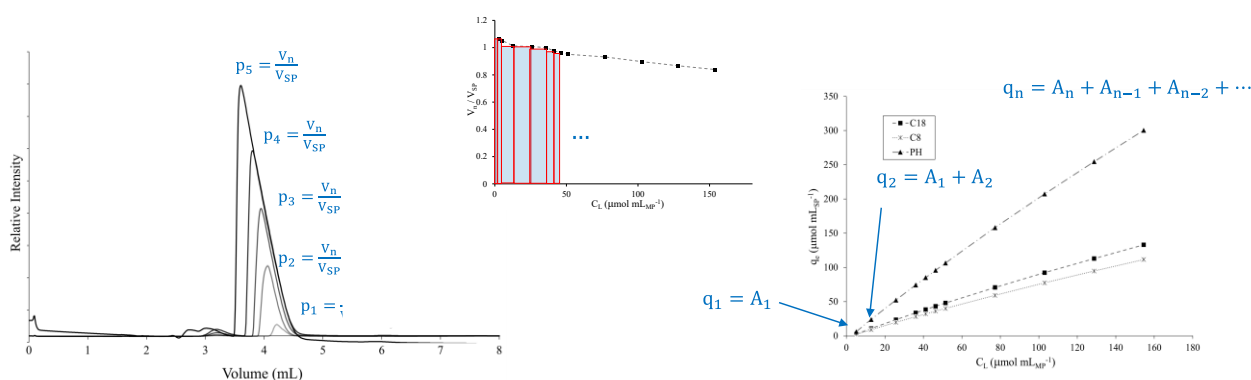


Figure 4.12 | Graphical summary of isotherm determination through peak maxima experiments.

One of the primary advantages of the elution method is its ability to reflect the dynamic conditions encountered in actual chromatographic processes. Unlike frontal analysis, which involves constant concentration profiles at equilibrium, the elution method captures the transient interactions between the solute and stationary phase as the solute peak moves through the column.

This method is generally faster and more efficient than frontal analysis. Since it involves the injection of a minor, discrete pulse of solute, the total time required for an experiment is significantly reduced. This rapid analysis is particularly beneficial in high-throughput environments where time or the amount of sample used are critical factors.

Elution of a characteristic point methods

Inverse elution methods by a characteristic point assume that the adsorption isotherm has no inflexions, and the adsorption behaviour is maintained throughout the entire adsorption isotherm, be it type I or III for example. These methods require a single experimental point at the

maximum adsorption capacity; the rest of the isotherm points are determined by considering the ratios as those of similar triangles. The equivalent for FA and PM are frontal analysis by characteristic point (FACP) and elution by characteristic point (ECP), respectively:

$$\frac{C_{\max}}{h_{\max}} = \frac{C_{e,2}}{h_{e,2}} \quad (4.27)$$

In equation 4.27, $C_{\max}$ and $h_{\max}$ have different meanings. For PM, they represent the sample concentration and the peak height of the experimental peak, respectively. For FA, they represent the analyte concentration in the mobile phase and the plateau height. Variables $C_{e,2}$ and $h_{e,2}$ are the values of concentration and height of a projected smaller peak or plateau overlapping with the boundary of the experimental peak or breakthrough curve.

The main advantages of these techniques are the reduction of time, mobile phase, and analyte requirements. However, they require some base knowledge of adsorption isotherms and their effect on peak shape. Additionally, using a single reference point can cause a loss of sensitivity to changes in the adsorption isotherms due to any inflexion points present in the isotherm. Therefore, it is recommended that additional confirmatory points for the isotherm be injected at a lower concentration of analyte.

Inverse method (IM)

The inverse method is an advanced technique used in chromatography to determine adsorption isotherms. Unlike traditional methods, which directly measure solute concentration changes during adsorption, the inverse method infers adsorption isotherms by analysing changes in chromatographic peak profiles. This approach is beneficial for systems where direct adsorption measurement is challenging or impractical.

The inverse method procedure begins with the same setup as PM, where one or multiple small, concentrated pulses of the solute are injected into the chromatographic system. The resulting elution profiles are recorded at different solute concentrations. The resultant peak shapes, retention times, and width changes are analysed.

The next step involves deriving the isotherm from the injection elution profiles by iteratively solving the liquid chromatography mass balance equation. The most commonly used model for IM is the equilibrium-dispersive model with the Martin-Synge algorithm[122, 123]. The solution of all mass balance models requires the presumption of the type of the isotherm, making IM an intrinsically biased method. If the wrong isotherm is preselected, the model will converge into a solution regardless of whether it is representative of the adsorption behaviour between the adsorbent and the adsorbate, or not.

Several key differences emerge when comparing the inverse method with traditional methods like FA and PM. Traditional methods directly measure changes in solute concentration in the stationary or mobile phase, providing straightforward data collection and analysis. In contrast, the inverse method infers adsorption isotherms indirectly from changes in peak profiles, offering a non-invasive alternative with detailed peak analysis. Traditional methods often require multiple injections and extensive experimentation to gather comprehensive data. In contrast, the inverse method can generate comprehensive isotherms from a single set of experiments, reducing the experimental workload.

Despite its advantages, the inverse method also has significant disadvantages. It requires sophisticated mathematical modelling and statistical techniques to accurately infer adsorption isotherms from peak profiles, making data analysis complex and computationally demanding. Additionally, precise control of experimental conditions, such as flow rate, temperature, and injection volume, is necessary to ensure accurate data. The method relies on highly sensitive LC detectors to capture subtle changes in elution profiles, and the initial development and validation can be time-consuming and complex. There have been attempts to debias the IM; however, more research needs to be done[123].

4.6 Other interactions in RPLC

Since RPLC has dedicated most of its history to applications involving the separation of small particles, where very punctual and well-studied interactions take place between the packing material and the analytes, it is acceptable to use only on simple chromatographic descriptors such as those discussed in the previous paragraph. Nonetheless, the mass-transfer effects and thermodynamic interactions involved during the separation of peptide-based molecules through LC

are complex by nature. All types of interactions and properties can be divided into three groups depending on the elements of the system involved (Table 4.5, Figure 4.13).

It is unclear whether the partitioning effect between the analyte/solute in the mobile and stationary phases or the adsorption of the analyte/solute into the stationary phase drives the separation of peptide-based systems in RPLC. Both phenomena may have a synergic effect, but little decisive research has been reported on this critical research question.

However, this issue has been addressed in HIC because the dynamics between the hydrophobic stationary phase and the surface hydrophobicity of the biomolecule are more evident. Here, adsorption of peptide-based systems is driven by the increase in entropy that accompanies the reduction in the total hydrophobic surface area[124]The hydrophobic patches on the surface of the biomolecule contact the hydrophobic ligands on the beads as they flow through the column, releasing water and, thus, decreasing the total wetted area. This liberation of water molecules from hydrophobic regions of the column increases the system's entropy, increasing the enthalpy of interaction between the solute and solid phase.

Table 4.5 | Parameters involved in peptide separation through RPLC.

Properties of RPLC Components			Interaction between RPLC components		
<i>Peptides</i>	<i>Solvent</i>	<i>Media/Resin</i>	<i>Peptide-Solvent</i>	<i>Resin-Solvent</i>	<i>Peptide-Resin</i>
Molecular mass	Chemical composition	Bed porosity	Molecular volume	Ligand charge	Binding equilibrium
Residue distribution	pH	Particle size distribution	Molecular diffusivity	Ligand hydrophobicity	Maximum capacity
Hydrophobicity	Buffer capacity	Particle porosity	Charge		Pore accessibility
	Ionic strength	Pore size distribution	Hydrophobicity		Pore diffusivity
	Temperature	Ligand density	Solubility		Surface diffusivity
	Conductivity	Swelling	2 nd Osmotic viral coefficient		
	Interstitial velocity				
	Gradient steepness				

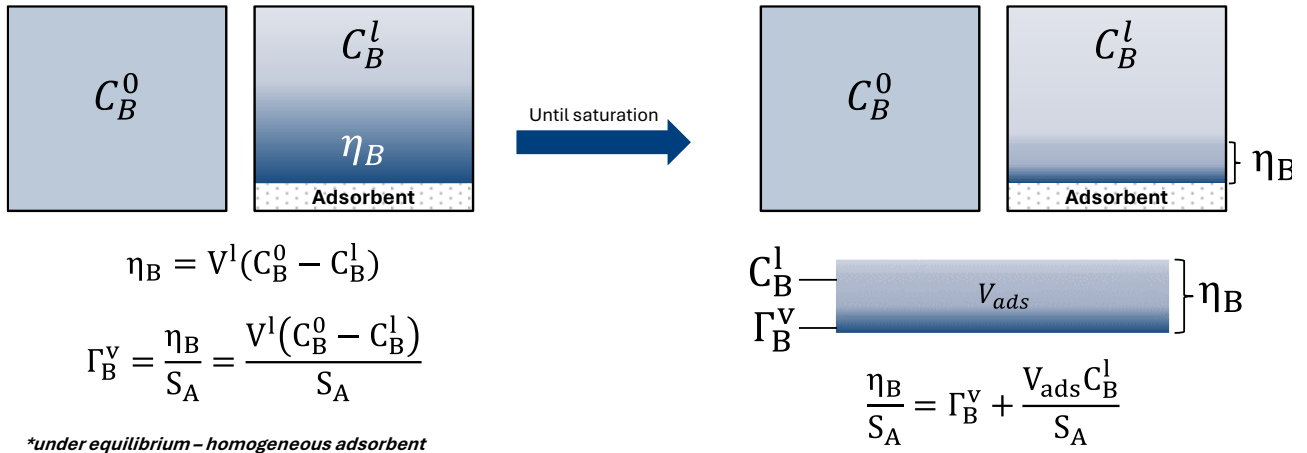


Figure 4.14 | Gibbs excess adsorption theory.

As static adsorption is assumed for these models, the above equations cannot be directly applied to chromatography. The most common approach to calculate excess isotherms in a dynamic chromatographic system is through the minor disturbance or perturbation method, which involves generating a concentration disturbance and measuring the retention of mobile phases with increasing compositions[126, 127]. An injection of pure solvent, either water or organic, under infinite dilution (minimal injection volume) would produce a peak. Equation 4.24 can describe the retention volume of that disruption peak in terms of the excess (equation 4.29). Two peaks are generally formed when water is injected into the equilibrated system with a fully formed adsorption liquid layer at the V_M value of the column (Figure 4.15). The first one is due to the elution of pure water. The second one, the water vacancy peak, is produced because water molecules disrupt the equilibria between the mobile and stationary phases. This disruption produces an instant state where the excess of organic solvent elutes immediately after the water[105].

$$V_R = V_M + S_A \frac{d\Gamma_B^v}{dC_B^l} \quad (4.29)$$

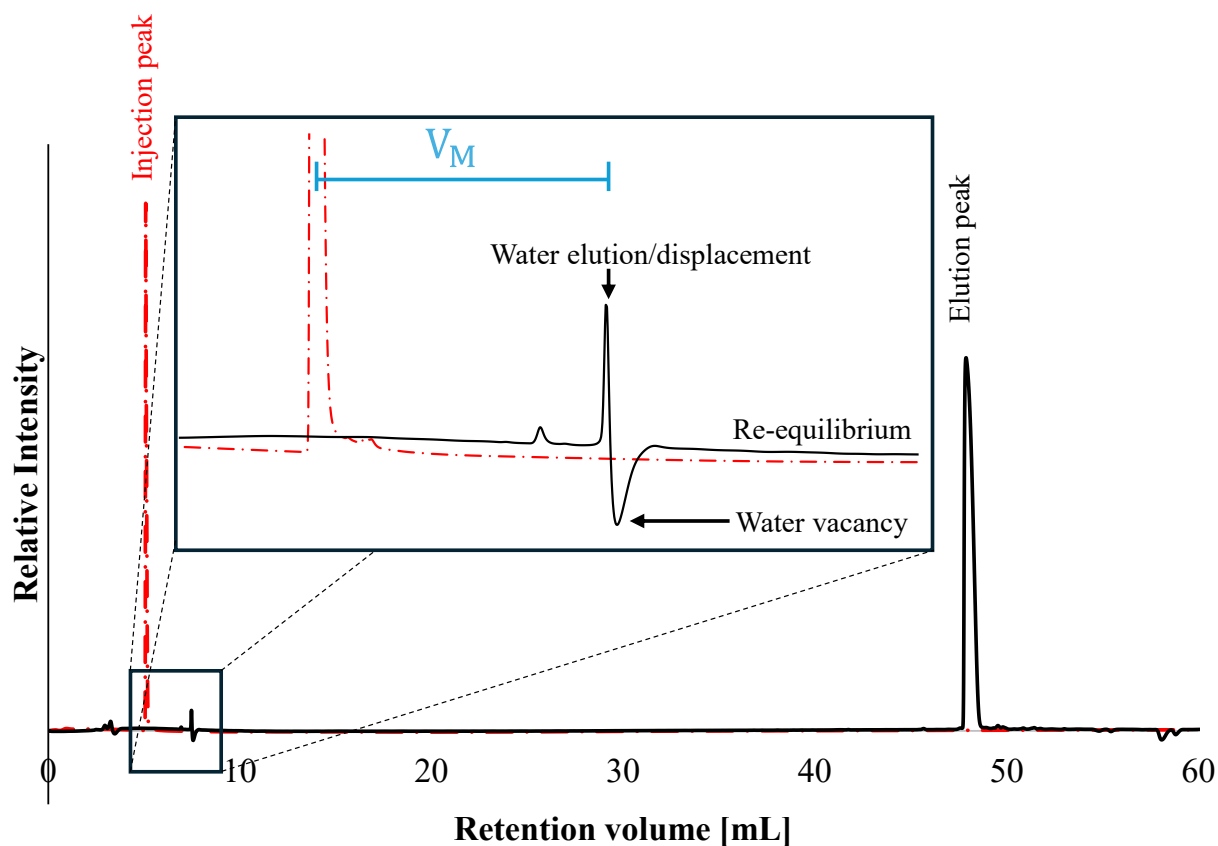


Figure 4.15 | Effect of small solvent injection into a fully equilibrated chromatographic column.

Since the pure solvent is injected, it would be expected that the peaks generated through the disruption method should have a retention volume similar to the hold-up volume; however, that is never the case. The retention volume of the disruption peak depends on the composition of the mobile phase (Figure 4.16). This occurs because the hydrophobic bonded groups on the surface of the column media can adsorb the organic component in the binary mobile phase, creating a molecular organic interface/interphase layer that becomes part of the stationary phase[128]. The integral average of the excess isotherm points obtained from the minor disturbance peaks gives the value of the hold-up volume (equation 4.30). This interphase and its composition play a critical role in the RPLC separation process. Excess adsorption in pure solvent is zero ($\Gamma_B^v(0\%) = \Gamma_B^v(100\%) = 0$):

$$V_m = \frac{\int_{C_B^I(0\%)}^{C_B^I(100\%)} V_r dC_B^I}{C_B^I(100\%)} \quad (4.30)$$

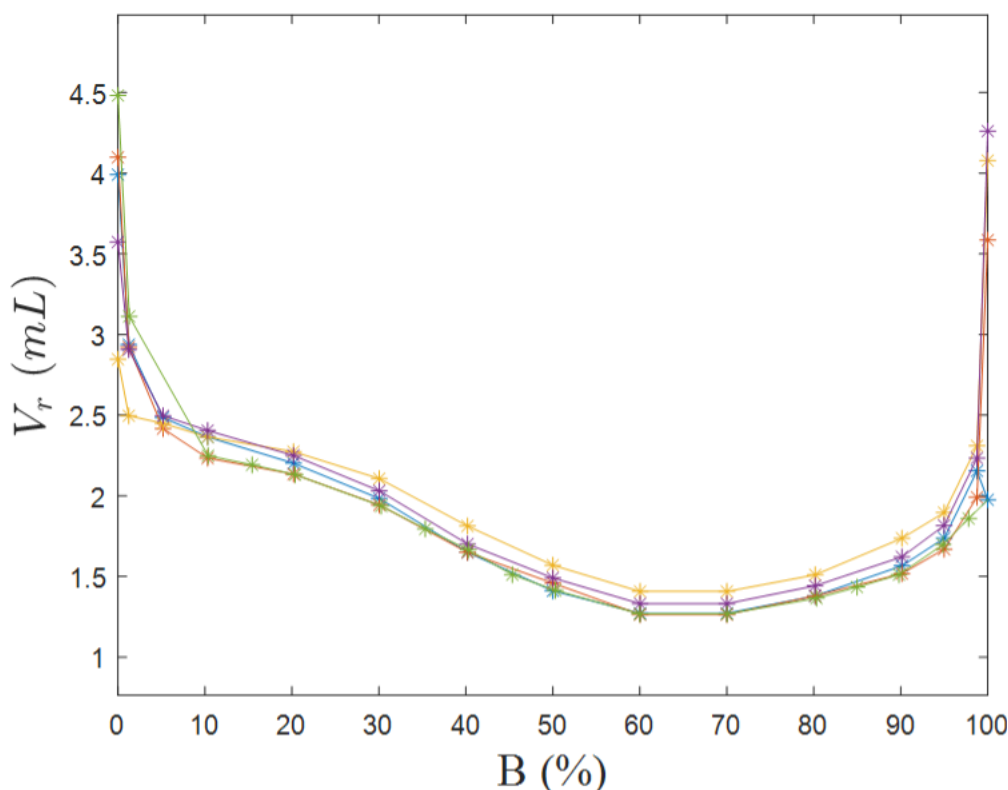


Figure 4.16 | Retention volume of perturbation peaks for different volume percentages of organic solvent in different RPLC columns.

Given that the boundary between these phases is not clearly distinguished, it becomes challenging to determine which portion of the organic phase should be considered part of the packing material with the interphase layer, and thus no longer part of the mobile phase. Different adsorption regions will be observed in heterogeneous surfaces like silica-based resins due to the presence of hydrophobic alkyl chains and hydrophilic residual silanols. In a mathematical representation, it can be inferred that organic solvent B molecules mainly accumulate on the hydrophobic surface domains of the resin/media (Equation 4.31). In contrast, the polar solvent (water) is attracted to the hydrophilic parts of the surface (Equation 4.32). Values V^{hpho} and V^{hphi} are the liquid volume of solvent B and water adsorbed into the hydrophobic and hydrophilic sites, respectively.

$$\frac{\eta_B}{S_A} = \Gamma_B^{\text{hpho}} + V^{\text{hpho}} C_B^l \quad (4.31)$$

$$\frac{\eta_w}{S_A} = \Gamma_w^{\text{hphi}} + V^{\text{hphi}} C_w^l \quad (4.32)$$

The total excess of solvent B along the isotherm can be described by equation 4.33. Values D_B and D_w are the molar density of solvent B and water, respectively. The full derivation of all equations in this Section can be found in Appendix III.

$$\Gamma_B^T = V^{\text{hpho}} (C_B^{\text{hpho}} - C_B^l) - V^{\text{hphi}} \left(C_w^{\text{hphi}} \frac{D_B}{D_w} - D_B + C_B^l \right) \quad (4.33)$$

As seen in Figure 4.17, the excess isotherm can be divided into three regions:

- I. The concentration of B is low. Any increase of B in the mobile phase will rapidly increase the excess adsorption of B until Γ_B^V is reached.
- II. The adsorbent surface is completely filled. Further increase of B in the mobile phase only raises its concentration in the bulk liquid phase. Thus leading to a **linear decrease** of excess adsorption on the surface until concentration B in the bulk liquid phase is equal to the concentration on the adsorbent surface ($\Gamma_B^V=0$).
- III. Excess isotherm continues to decrease below zero until Γ_B^V is reached. The concentration in the bulk liquid becomes very high, which starts replacing some water molecules from the surface.

During linear decrease (Region II), equation 4.33 can be simplified to equation 4.34, given that $V^{\text{hpho}} = V_B$, $C_B^{\text{hpho}} = D_B$, $V^{\text{hphi}} = V_w$, and $C_w^{\text{hphi}} = D_w$ are true in that region. With this final equation $V_B + V_w$ approximates V_{ads} . From here, interphase layer thickness (τ) can be computed through the equation 5.35[111], where v_m is the molar volume of the solvent:

$$\Gamma_B^T = V_B D_B - (V_B + V_W) C_B^l \quad (4.34)$$

$$V_{\text{ads}} = S_A C_{\text{ads}} v_m = S_A \tau \quad (4.35)$$

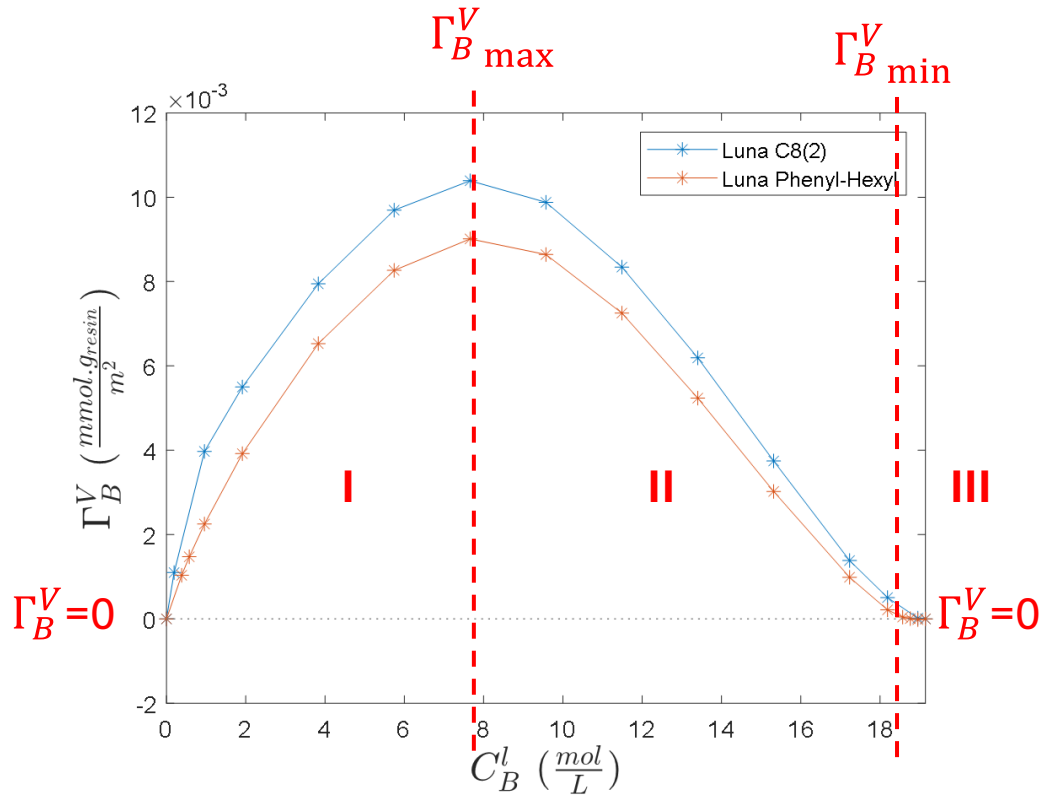


Figure 4.17 | Excess adsorption isotherms of an organic modifier onto RPLC columns. See text for description.

Chapter 5

Chromatography at the manufacturing scale

5.1 Definitions

5.1.1 Purity, yield and throughput

Purity or quality refers to the degree to which a separated compound is free from contaminants or impurities. It is often expressed as a percentage and indicates the proportion of the desired compound relative to the total amount of material collected in the fraction. High purity means that the target compound is present with minimal impurities.

Yield refers to the amount of the target compound recovered from the chromatography process relative to the initial amount of that compound in the starting mixture. It is also often expressed as a percentage and represents the process's efficiency in recovering the desired compound.

Throughput refers to the quantity of material the chromatography system can process in a given period. It is usually measured as the amount of feed or sample that can be handled per unit of time, such as grams per hour or litres per hour. Throughput reflects the capacity and speed of the chromatography process.

Each of these parameters is critical in evaluating the performance and suitability of a chromatography method for a given application. Balancing these factors is often necessary to optimise the separation process for the desired outcome (Figure 5.1).

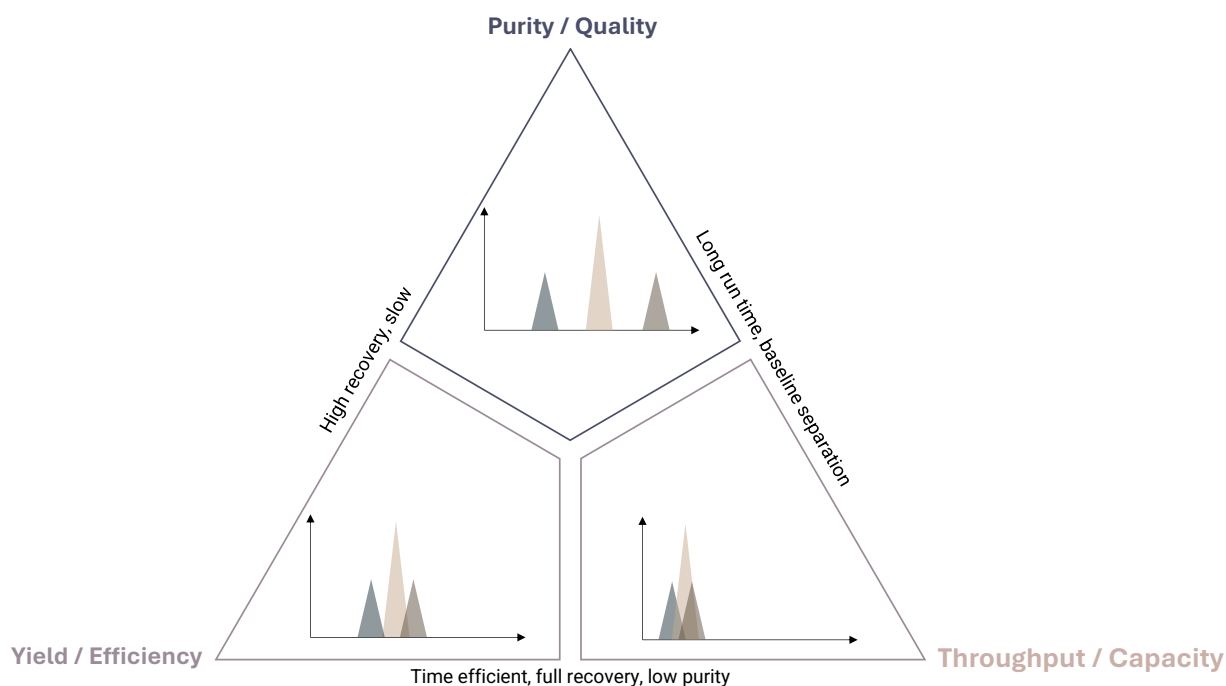


Figure 5.1 | Ternary diagram of chromatographic separation performance.

5.1.2 Analytical and preparative scale

The type of chromatography employed significantly affects experimental or production outcomes, with the three primary scales of chromatography—analytical, preparative, and semi-preparative—each serving distinct purposes based on scale, objectives, and methodologies.

Analytical chromatography is used to analyse mixture composition, focusing on identifying, quantifying, and characterising the compounds present. It operates on a small scale with low sample volumes and flow rates, providing high resolution and precise detection. This method is widely used in quality control, bioanalysis, and research, where achieving high purity is essential for accurate analysis. Quantities of material analysed will be in the μg -mg range.

Preparative chromatography, in contrast, is designed to isolate and purify large quantities of specific compounds. This large-scale process uses larger columns and higher flow rates to enhance throughput and yield. While detector sensitivity is lower than in analytical methods, the focus is on maximising the recovery of purified compounds for pharmaceutical development, biochemical production, or further synthetic reactions. The quantities of material analysed can range from many grams to many kg.

Semi-preparative chromatography bridges the gap between analytical and preparative methods, balancing the need for both analysis and preparation. It operates on a medium scale, purifying moderate quantities with intermediate column sizes and optimised flow rates for both speed and resolution. This method is particularly useful for purifying compounds for research, method development, and small-scale synthesis. Quantities of material analysed can range from mg to many grams.

In both analytical and preparative chromatography, the goals differ. Analytical chromatography prioritises high purity for detailed analysis, while preparative chromatography focuses on high yield and purity to collect usable amounts of the target compound. Semi-preparative chromatography is a versatile approach for applications requiring analysis and production.

5.1.3 Linear and nonlinear chromatography

The distinction between linear and nonlinear chromatographic behaviour is essential for understanding various chromatographic separation processes. In linear chromatography, equilibrium isotherms are straight lines originating from the isotherm origin, meaning that band shapes and retention times are unaffected by the sample concentration or composition. The peak height in the chromatogram is directly proportional to the amount of each component injected, making this approach ideal for analytical applications, particularly when sample concentrations are low.

Nonlinear chromatography, on the other hand, is characterised by equilibrium isotherms that deviate from linearity. In this case, the concentration of a component in the stationary phase is not directly proportional to its concentration in the mobile phase, and this relationship depends on the presence of other components in the mixture. Key chromatographic parameters such as chromatographic band shape, peak maxima, and retention time are affected by both the sample composition, the injected amount and the solution solvent (mobile phase) composition. Nonlinear chromatography is more relevant to preparative chromatography, where larger sample loads are used to isolate significant quantities of target compounds[128, 129].

In RPLC, most substances follow nonlinear adsorption behaviour, complicating the prediction of retention and band shapes due to this non-linear adsorption. To tackle these complexities, nonlinear chromatography can be modelled through ideal or nonideal approaches, helping

researchers to isolate factors contributing to deviations in linearity and better understand the underlying thermodynamics and kinetics.

Understanding the difference between linear and nonlinear chromatography is particularly important when transitioning from analytical-scale separations, typically operating under linear conditions, to preparative-scale processes, where nonlinear behaviour prevails due to higher concentrations and adsorbate-adsorbent interactions.

5.1.4 Ideal and nonideal chromatography

In ideal chromatography, several simplifying assumptions are made to facilitate the analysis of chromatographic behaviour. The primary assumption is that column efficiency is infinite, implying that axial dispersion is negligible and that mass transfer kinetics for adsorption/desorption occur instantaneously. Under these conditions, the internal surface inside the porous chromatographic particles remains in equilibrium with the surrounding solution. As a result, elution band profiles are governed entirely by the thermodynamics of phase equilibria, where the band shape reflects the injection profile. Ideal linear chromatography results in identical elution profiles, with variations dependent only on the retention factor or the slope of the linear isotherm. However, this scenario is impractical and rarely applicable in real-world chromatography.

Nonlinear ideal chromatography, while also theoretical, is a valuable model as it focuses on the influence of thermodynamics on the band profile, excluding the impact of mass transfer kinetics and axial dispersion. This makes it helpful in understanding the thermodynamic properties of separation processes.

In contrast, nonideal chromatography assumes finite column efficiency, where different factors contribute to band broadening. Several models exist to describe this behaviour. One model might consider finite axial dispersion, while another may focus on the kinetics of specific processes, such as adsorption-desorption, while ignoring dispersion. The general rate model of chromatography is the most comprehensive, accounting for all sources of nonideal behaviour, including molecular diffusion, eddy diffusion, and the kinetics of mass transfer steps. These nonideal models are essential for understanding real-world chromatography, where numerous factors cause deviations from ideal behaviour.

Table 5.1 | Summary of differences between chromatographic scales.

Parameter	Analytical	Semi-Preparative	Preparative
Purpose	Analyse and quantify components in a mixture	Balance between analysis and purification	Isolate and purify large quantities of target compounds
Scale	Small-scale, low-sample volumes	Medium-scale, moderate sample volumes	Large-scale, high sample volumes
Column Size	Small columns	Intermediate column sizes	Large columns
Flow Rate	Low flow rates for precision	Intermediate flow rates to balance speed and resolution	High flow rates to maximise throughput and yield
Focus	High-resolution and precise detection	Moderate resolution with a focus on moderate purity and yield	Maximising recovery of the target compound (yield)
Typical Applications	Quality control, bioanalysis, research	Research, method development, small-scale synthesis	Pharmaceutical development and production
Purity Requirement	High purity is required for accurate analysis	Moderate purity	High purity, but yield is often prioritised
Detector Sensitivity	High sensitivity, detailed analysis of sample components	Moderate sensitivity	Lower sensitivity compared to analytical chromatography
Sample Handling Capacity	Handles very small sample amounts (milligrams)	Handles moderate sample amounts (grams)	Handles large sample amounts (grams to kilograms)
Throughput	Low	Moderate	High
Use of Nonlinear Chromatography	Rarely needed, it operates in linear conditions	Can experience nonlinear behaviour at times	Often operates under nonlinear conditions due to large sample loads

5.2 Scaling Up RPLC for Peptide Purification

Scaling up RPLC for peptide purification from the laboratory to the manufacturing scale presents several technical challenges that require careful consideration to maintain both peptide purity and overall process efficiency. This Section outlines the principles and practical considerations necessary for successful scale-up, ensuring that key chromatographic properties are preserved while optimising production throughput.

5.2.1 Differences Between Lab-Scale and Industrial-Scale RPLC

The fundamental separation principles of RPLC remain consistent between laboratory and manufacturing scales, but the transition to a larger scale requires adjustments in several key factors. These include:

- **Column Size:** In lab-scale RPLC, smaller columns (e.g., 4.6 mm i.d.) are typically used, while in industrial-scale processes, column diameters can reach 30 cm or more. This change in column dimensions significantly affects pressure requirements, flow rates, and overall system dynamics.
- **Flow rate and volume:** Lab-scale separations often operate at low flow rates (e.g., 0.5–1 mL/min), while large-scale processes can operate at flow rates of several litres per minute. The increased volume introduces challenges in maintaining consistent gradient profiles across the column length and width.
- **System Dynamics:** Scale-up often leads to changes in system behaviour, such as decreased backpressure per column volume, longer diffusion paths, and greater solvent consumption. Addressing these dynamics is critical to avoid losses in separation efficiency.

5.2.2 Theoretical Framework for Scaling

The scale-up of RPLC is governed by several scaling laws and principles aimed at maintaining chromatographic performance while accommodating higher throughput. These principles include:

- **Linear Velocity and Flow Rate Adjustments:** One of the most important factors in scale-up is ensuring that the linear velocity of the mobile phase across the stationary phase remains consistent between processing scales. The flow rate (F) can be scaled up using the following relationship:

$$F_{\text{large}} = F_{\text{small}} \times \left(\frac{d_{\text{large}}}{d_{\text{small}}} \right)^2 \quad (5.36)$$

where d_{large} and d_{small} are the diameters of the large- and small-scale columns, respectively.

This ensures that the retention times and resolution observed at the analytical scale are

preserved in the scaled-up process. The ratio of the squared diameters is the scale-up factor where the length of the columns remains constant, which is the most commonly used approach.

- **Maintaining the Gradient Profile:** In RPLC, peptides are typically separated using a gradient of increasing organic solvent concentration. Maintaining the same gradient slope based on column volume and profile becomes critical to ensure consistent elution patterns as the system is scaled up. This may involve adjusting solvent delivery systems to ensure the gradient remains uniform throughout the column.
- **Column Length Considerations:** While column diameter typically increases during scale-up, column length may also need to be adjusted to preserve separation efficiency. Taller columns often have longer bed heights to maintain resolution for complex peptide mixtures, which increases the risk of higher backpressures. In the case that the column length must change for the scale-up process, equation 5.36 changes to:

$$F_{\text{large}} = F_{\text{small}} \times \frac{V_c^{\text{large}}}{V_c^{\text{small}}} \quad (5.37)$$

where the ratio of the column volume of both columns is the universal scale-up factor.

5.2.3 Impact on Gradient Optimization for Large-Scale Systems

In laboratory-scale RPLC, mobile phase gradient optimisation is often straightforward with small solvent volumes and short column lengths. However, during scale-up, several adjustments must be made. In large-scale systems, managing solvent consumption is crucial due to the significant volumes required to maintain gradient flow. This high solvent usage not only increases costs but also raises environmental concerns. Therefore, optimising solvent usage by fine-tuning gradient slopes is essential to minimise waste.

Additionally, careful management of elution times is necessary to prevent excessive dilution of peptide fractions. Changes in flow rate and column dimensions can alter peptide retention times, necessitating the re-optimization of gradient parameters to avoid issues such as co-elution or peak broadening.

Furthermore, solvent mixing efficiency can vary in large systems, impacting gradient reproducibility. This challenge can be addressed by employing advanced mixing systems or implementing real-time gradient monitoring to ensure consistent solvent mixing before it enters the column.

5.2.4 Pressure and Temperature Considerations in Scale-Up

Scaling up RPLC often introduces significant challenges related to pressure and temperature control.

Firstly, pressure control is crucial. Larger column dimensions and higher flow rates typically result in elevated system backpressures. Pressure control strategies must be implemented to avoid exceeding the pressure limits of the chromatography system or damaging columns. These can include adjusting flow rates or using high-pressure-resistant columns.

Secondly, temperature management is essential. Temperature is a critical parameter in peptide separation, influencing both resolution and retention times. As column sizes increase, the risk of uneven temperature distribution across the column also rises. Using temperature-controlled or jacketed columns can help maintain consistent temperature profiles during large-scale runs.

5.2.5 Importance of particle size considerations

One of the most critical factors in improving peak resolution during scale-up is the particle size of the stationary phase in the column. The particle size directly affects the efficiency of separation, peak sharpness, and the overall pressure requirements of the system.

- **Smaller Particle Size for Better Resolution:** Columns packed with smaller particles (e.g., sub-2 micron or 3-micron particles) typically provide better resolution because they offer a higher number of theoretical plates, meaning more separation "zones" along the column length. This leads to sharper peaks and improved separation of closely eluting peptides, which is essential for effective peak shaving. However, smaller particle sizes also result in higher column backpressures, which can be a challenge on a large scale.

- **Trade-Off Between Resolution and Pressure:** The benefit of improved resolution from smaller particles must be weighed against the increased system backpressure. In large-scale RPLC, the use of very small particles may necessitate specialised high-pressure pumps or system components to handle the increased pressure, potentially adding to the complexity and cost of the process[130]. Therefore, a balance must be struck between achieving optimal resolution and maintaining manageable system pressures.
- **Larger Particle Size for Lower Pressure:** While larger particles (e.g., 5 microns or higher) result in lower backpressure, they also lead to broader peaks and reduced resolution. In some cases, this may be acceptable if the separation requirements are less stringent, but for closely related peptides, using larger particles can compromise the effectiveness of peak shaving[130]. When using larger particles, optimisation of other factors, such as flow rate and gradient steepness, becomes even more critical to maintain sufficient separation.
- **Monolithic Columns and Core-Shell Particles:** In some cases, advanced stationary phases like monolithic columns or core-shell particle technologies can be used to balance resolution and pressure. Core-shell particles with a solid core and a porous outer layer offer high resolution similar to small particles but at lower backpressure[131, 132]. This can be particularly advantageous in manufacturing-scale applications where system pressure is a limiting factor.

By carefully selecting the appropriate particle size for the stationary phase, peak resolution can be optimised while backpressure challenges are minimised, making peak shaving more precise and effective during large-scale peptide purification.

5.3 Peak Shaving

Peak shaving refers to the selective cutting or removal of chromatographic peaks, parts thereof, to capture the most concentrated or purest portion of a target compound while discarding or recycling the less pure fractions. In the context of peptide purification at the manufacturing scale, peak shaving serves several purposes:

- **Enhancing Purity:** By capturing only the central, most resolved portion of a peak, impurities that elute at the tail or front of the peak are excluded, thereby improving the purity of the final product.

- **Maximizing Yield:** In some cases, the shoulder regions of a peak may contain significant amounts of the desired peptide. Peak shaving enables the collection of these regions while minimising contamination from impurities.
- **Improving Efficiency:** In large-scale production, peak shaving allows for more precise control over product recovery, reducing the need for extensive reprocessing or repurification, which saves time and resources.

In Figure 5.2, a depiction of analytical and preparative approaches to separating a theoretical complex peptide sample is presented. Both scales present similar multi-gradient elution profiles as a scale-up suggestion of each other. Although the resulting peaks are not resolved to standard analytical scale objectives (purity over yield), the peaks are very well defined (Gaussian-shaped) as the analytes follow chromatographic linear adsorption behaviour due to the small injection per column volume.

Due to the larger injection at the preparative scale, the chromatographic peaks usually follow non-ideal, non-linear adsorption, causing the peak to exhibit a Langmuir-like shape isotherm. To produce better separation, general practice is to use a narrower first gradient for elution and increase the flow rate. Since the primary objective is yield and high throughput, peak shaving is used to the extent of the main chromatographic peak cluster, where shorter fractions are collected, and the main peak is expected to be more concentrated (shown in green). All of those fractions are tested for purity through an optimised analytical method, and those above the purity threshold are pooled together to obtain a final purity (shown in purple). This approach improves recovery despite the poor chromatographic method.

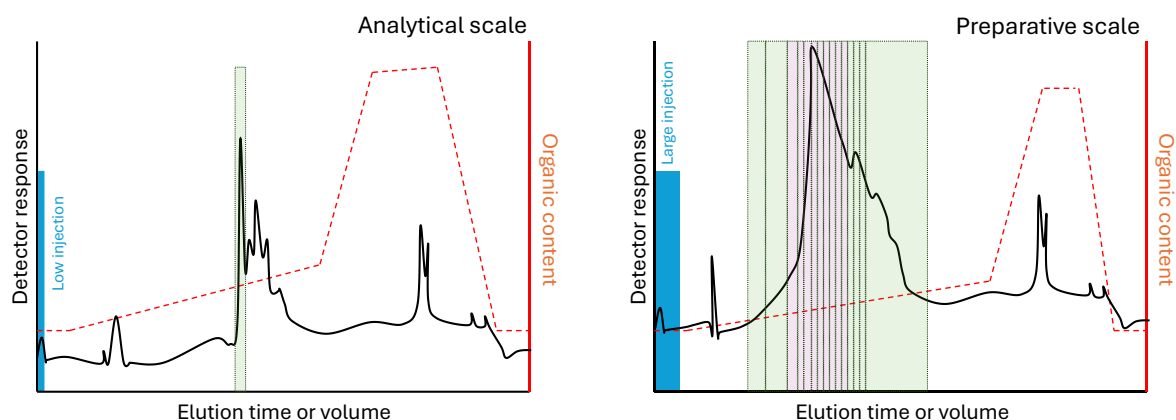


Figure 5.2 | Analytical and preparative scale method comparison. See text for description.

Chapter 6

Materials and Methods

6.1 Materials

6.1.1 General

HPLC-grade acetonitrile (ACN, $\geq 99.8\%$ purity), methanol (MeOH, $\geq 99.8\%$), ethanol (EtOH, $\geq 99.8\%$), isopropanol ($\geq 99.5\%$), methylene chloride ($\geq 99.5\%$), hexane ($\geq 95\%$), were purchased from Fisher Scientific (USA). Type I (deionised, DI) water with a resistivity of $18\text{ M}\Omega\cdot\text{cm}$ was obtained from a Purelab Chorus 1 system (Elga LabWater, USA). Trifluoroacetic Acid (TFA, $\geq 99.0\%$) was purchased from Sigma-Aldrich (USA).

The following reagents were purchased from Merck (Germany): Uracil (99%), Ethylbenzene (anhydrous 99.8%), Propylbenzene (98%), Butylbenzene (99%), Pentylbenzene (99%), Hexylbenzene (97%), o-Terphenyl (99%), Triphenylene (98%), Caffeine (ReagentPlus, 99%), Benzylamine, Phenol ($\geq 99\%$), Phosphoric acid (ACS reagent, $\geq 85\text{ wt. \%}$), Potassium phosphate monobasic (ACS reagent, $\geq 99.0\%$), Potassium phosphate dibasic (ACS reagent, $\geq 99.0\%$), Potassium hydroxide (ACS reagent, $\geq 85\%$, pellets).

6.1.2 Peptides

Peptides P1 to P6, as proposed by Krokhin and Spicer[133], were purchased from GenScript (USA) with a minimum purity of 96% and no termini modification (-NH₂ n-terminal and -OH c-terminal). Table 6.1 contains information about these standard peptides, and the complete structures can be found in Figure 6.1. These peptide standards were selected due to their increasing hydrophobicity and lack of secondary structures. They are suitable as retention time markers for peptides and tryptic protein fragments, as they display uniform retention distribution under isocratic and gradient conditions in RPLC settings. Under gradient elution, they cover a narrow ACN% concentration range (ca. 25%) for nano-, micro- and normal-flow gradient systems.

Additional features, such as a molecular weight of around 900 Da and C-terminal adjacent Asp, make these peptides easily detectable by most MS/MS search engines for protein identification using ESI and MALDI MS techniques.

Table 6.1 | Characteristics of peptide standards. Sequences were obtained from Krokhin and Spicer[133]. Hydrophobicity scales were calculated manually according to the scales from the reference cited.

ID*	Peptide sequence	No. residues	m/z (+1)*	Hydrophobicity			
				cLogP	Kyte-Doolittle[134]	Tanford[135]	Ponnuswamy[136]
P1	LGGGGGGDGSR	11	888.413	-5.72	-0.709	0.190	0.052
P2	LGGGGGGDFR	10	892.428	-5.01	-0.380	0.298	0.278
P3	LLGGGGDFR	9	891.469	-4.18	0.089	0.394	0.530
P4	LLLGGDFR	8	890.510	-2.67	0.675	0.515	0.845
P5	LLLLDFR	7	889.551	-1.14	1.429	0.670	1.250
P6	LLLLLDFR	8	1002.635	-0.39	1.725	0.778	1.3425

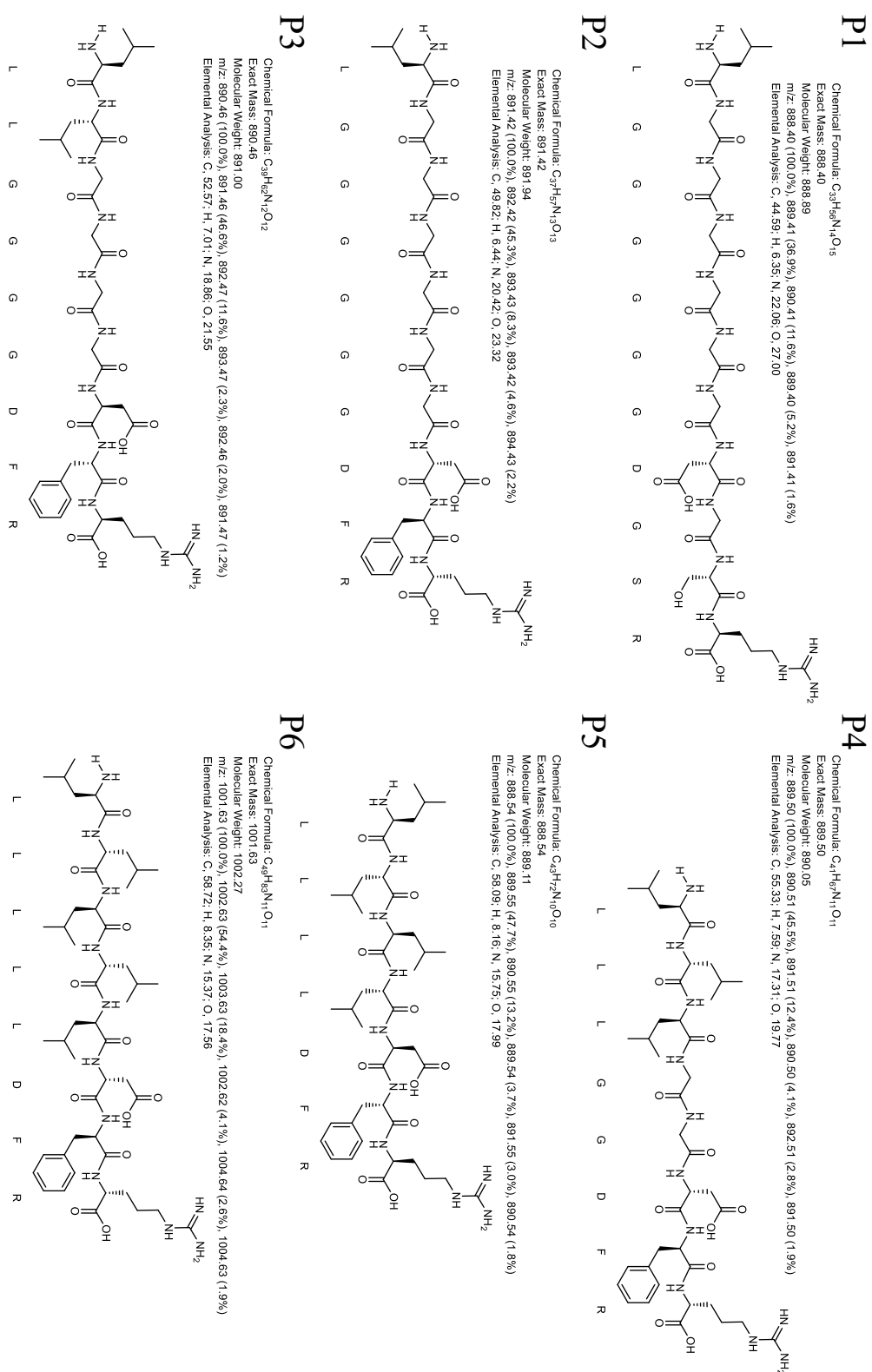


Figure 6.1 | Full structure of Krokhin peptide standards. Chemical structures of the compounds were generated using ChemDraw software (version 23.0).

Non-purified human glucagon(1-29) (Sequence: HSQGTFTSDYSKYLDSRRAQDFVQWLMNT, m/z^{+1} : 3481.827) was purchased from GenScript (USA) with no termini modification with a 27.0 ± 0.73 crude purity based on the HPLC area. The three-dimensional structure of glucagon can be found in Figure 6.2. Due to the high level of impurities, glucagon samples were filtered and sonicated to reduce the amount of dissolved gas before injection into the LC systems.

Using this impure glucagon as a test peptide system in the development of LC methods offers several key advantages. As a relatively small peptide with 29 amino acids, its size allows for faster method optimisation and analysis than for larger proteins or peptides. Its well-characterized structure and physicochemical properties provide a solid foundation for method adjustment, making predicting and controlling its chromatographic behaviour during separations easier.

Glucagon is also commercially available in a stable and highly purified form, ensuring consistent performance in LC method development. Its hydrophobic nature makes it ideal for RPLC. The aromatic residues in its sequence enable easy UV absorbance detection, simplifying the monitoring process during chromatographic runs.

Furthermore, glucagon contains ionisable amino acid residues, offering flexibility in method development using mobile phases with added ion-pairing agents. Once an LC method is optimised for glucagon, it can serve as a model for other peptides of similar size and composition, streamlining method transfer to different applications. The stability, simplicity, and well-known characteristics of glucagon make it a highly effective peptide probe for developing robust and scalable LC methods.

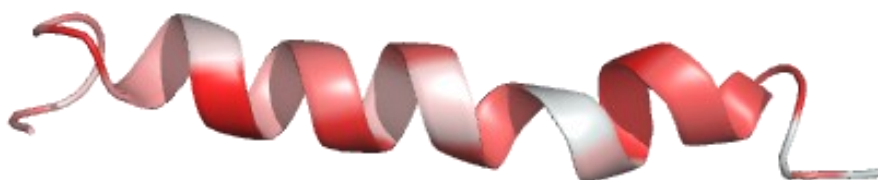


Figure 6.2 | Secondary structure of glucagon(1-29). The red colour depicts regions of higher hydrophobicity according to the Eisenberg scale[30].

6.2 LC instruments

6.2.1 Analytical scale

All analytical experiments were conducted using a modular Prominence LC-40 (Shimadzu, Japan), which will be referred to throughout this Thesis as the 'analytical LC'. Figure 6.3 shows a generalised schematic of the analytical instrumentation setup. Mobile phase solutions were continuously filtered with a 10 μm in-line filter and degassed using a DGU-405 degassing unit and an LC-40D solvent delivery pump. The solvent delivery pump was equipped with a low-pressure mixing system composed of a quaternary proportioning valve that controls the mixing of solvents and a parallel dual-head high-pressure pump to maintain steady mobile phase flow rates. The column temperature was controlled using a CTO-40C column oven suited with a 6-port column selection valve. The extra column and dwell volumes were estimated for each valve position as independent systems (Table 6.2). The chromatographic data was recorded during injection and elution by positioning a dual-channel SPD-40 UV/Vis detector and an SPD-M40 photodiode array (PDA) detector before and after the column, respectively. All samples were injected using a SIL-40C autosampler. Sample vials were kept at 15 $^{\circ}\text{C}$ using the temperature-control function of the autosampler. The LC was controlled using a communication bus module CBM-40lite and a 64-bit desktop computer running on Windows 10 Pro version 21H1 and Shimadzu LabSolutions Version 5.110.

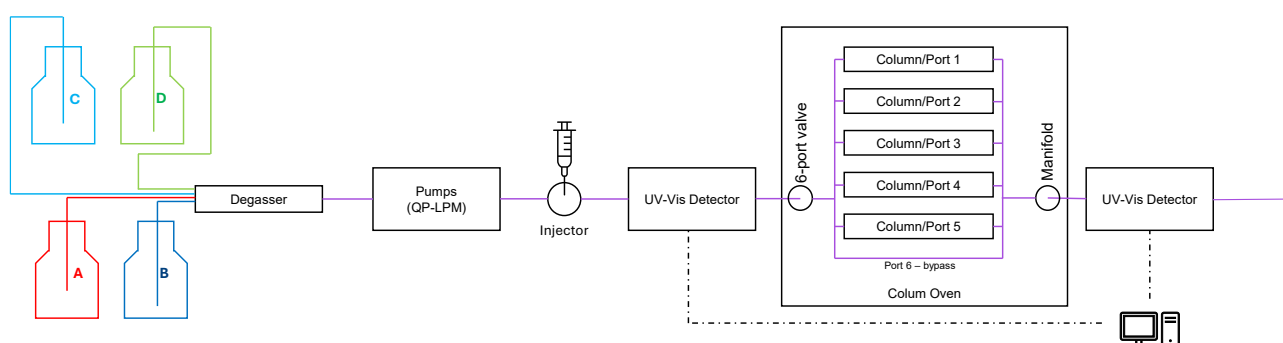


Figure 6.3 | Schematic of analytical HPLC used for experimentation.

Table 6.2 | Extra-column and dwell volumes of analytical LC. Extra-column volumes were calculated by injecting 0.1 µg of uracil and caffeine in water with 0, 25, 50, 75 and 100% (v/v) of methanol. Dwell volumes were estimated using gradient elution from 0-100% (v/v) water to 0.1% acetone in water in 20 min with no column attached and a zero-dead volume connector. The numbers shown are the average and standard deviations of all runs.

Port valve	Pre-column detector		Post-column detector	
	Extra-column volume	Dwell volume	Extra-column volume	Dwell volume
1	86.70 ± 2.42 µL	1.37 ± 0.09 mL	172.3 ± 2.55 µL	1.27 ± 0.12 mL
2	87.89 ± 2.59 µL	1.59 ± 0.06 mL	173.2 ± 2.61 µL	1.50 ± 0.12 mL
3	87.65 ± 2.22 µL	1.45 ± 0.02 mL	172.3 ± 2.42 µL	1.38 ± 0.05 mL
4	88.85 ± 2.38 µL	1.48 ± 0.01 mL	173.2 ± 1.90 µL	1.40 ± 0.06 mL
5	88.35 ± 2.63 µL	1.48 ± 0.01 mL	173.1 ± 2.29 µL	1.40 ± 0.06 mL
6	87.85 ± 3.47 µL	1.49 ± 0.01 mL	167.3 ± 2.11 µL	1.48 ± 0.15 mL

This HPLC instrument configuration with pre- and post-column detectors offers several advantages over a standard single-detector HPLC system, particularly for monitoring system performance and method development. One key benefit is assessing column efficiency and monitoring the baseline. The chromatographer can identify signal changes by comparing signals before and after the column, which helps detect issues such as column degradation or fouling. This setup also enables more effective troubleshooting by identifying contaminants, leaks, or mobile phase changes early in the process.

Additionally, having a pre-column detector helps distinguish sample matrix components that might interfere with the analysis without being retained by the column. This attribute is especially useful for identifying solvent front or buffer changes during gradient elution, ensuring that the system is functioning as expected. The pre-column detector can also detect injection issues, ensuring the sample enters the HPLC system correctly.

Quantitatively, comparing the signals from both detectors can provide valuable insights into sample recovery, the presence of irreversible adsorption events, and analyte stability. By monitoring the analyte before and after the column, users can determine if any sample degradation occurs under the chromatographic conditions. Overall, this dual-detector setup enhances method reliability, allowing for more accurate peak identification and system diagnostics.

For all chromatographic experiments, the mobile phase comprised DI water (solvent line A) and ACN (solvent line B) for peptide standards. For glucagon, 0.1% TFA (v/v) was added to both A

and B. These mobile phases were delivered in separate lines in the LC systems and mixed with the quaternary pump.

6.2.2 Preparative scale

An AKTA Pure 25 L (Cytiva, formerly GE Healthcare Life Sciences, USA) was employed for all preparative-scale chromatographic separations. This instrument will be referred to as the 'preparative LC'. The system was fitted with two binary system pumps with a flow rate range of up to 25 mL/min. A system pressure monitor is connected to the pumps. The pumps were used to mix two lines (A and B) that mix at an attached 5-mL mixer mounted with an in-line filter to ensure homogenous mixing and filtering to remove any debris in the mobile phase. Sample injection was performed manually through a fitted injection valve with a 200 μ L loop. Signal was detected using an LED UV monitor with a fixed wavelength of 280 nm. Data was collected using the UNICORN software version 7.3.0.473, which was installed in a Windows 10 Pro version 22H2. The dwell and extra-column volumes of this instrument were not calculated as the software makes these adjustments directly.

6.2.3 Columns

All RPLC columns used for this Thesis were acquired directly from the manufacturer. Some columns in Table 6.3 were used for other experiments before this Thesis. Therefore, a regeneration procedure was performed before running any new experiments. The column was connected in the reverse direction and flushed with a series of organic solvents at a flow rate from 0.2 to 0.5 mL/min adjusted to keep a pre-column pressure lower than 50 bar for 10CV. The solvents and order were DI water, isopropanol, methylene chloride, hexane, methylene chloride, isopropanol and water.

New columns (Table 6.4) were tested for packing integrity by recording the back pressure of running a mobile phase of pure DI water at different flow rates to confirm linearity.

Table 6.3 | Properties of columns used in Chapters 7-10. Information was obtained directly from the manufacturer.

Manufacturer	Column	I.D. x L (mm x mm)	Particle size (μm)	Pore size (Å)	Surface area (m ² /g)	Carbon load (%)	End-capping
Phenomenex	Luna C8 (2)	4.6 x 150	5	100	400	13.5	TMS
Phenomenex	Luna C18 (2)	4.6 x 150	5	100	400	17.5	TMS
Phenomenex	Luna Phenyl-Hexyl	4.6 x 150	5	100	400	17.5	TMS
Phenomenex	Luna Prep C18 (2)	4.6 x 250	10	100	No information	No information	No information
Waters	Spherisorb ODS1	4.6 x 250	5	80	220	No information	No
Waters	Symmetry C18	4.6 x 75	3.5	100	335	19	Proprietary

Table 6.4 | Properties of columns used in Chapter II. Information was obtained directly from the manufacturer.

Manufacturer	Column	I.D. x L (mm x mm)	Particle size (μm)	Pore size (Å)	Surface area (m ² /g)	Carbon load (%)	End-capping
Phenomenex	Luna C5	4.6 x 250	5	100	440	12.5	TMS
Phenomenex	Luna C8(2)	4.6 x 250	5	100	400	13.5	TMS
Phenomenex	Luna C8(2)	4.6 x 250	10	100	400	13.5	TMS
Phenomenex	Luna C8(2) PREP	4.6 x 250	10	100	No information	No information	No information
Phenomenex	Luna C18(2)	4.6 x 250	5	100	400	17.5	TMS
Phenomenex	Luna C18(2)	4.6 x 250	10	100	400	17.5	TMS
Phenomenex	Luna C18(2) PREP	4.6 x 250	10	100	No information	No information	No information
Phenomenex	Luna Phenyl-Hexyl	4.6 x 250	5	100	400	17.5	TMS
Phenomenex	Luna Phenyl-Hexyl	4.6 x 250	10	100	400	17.5	TMS
Phenomenex	Luna Phenyl-Hexyl PREP	4.6 x 250	10	100	No information	No information	No information

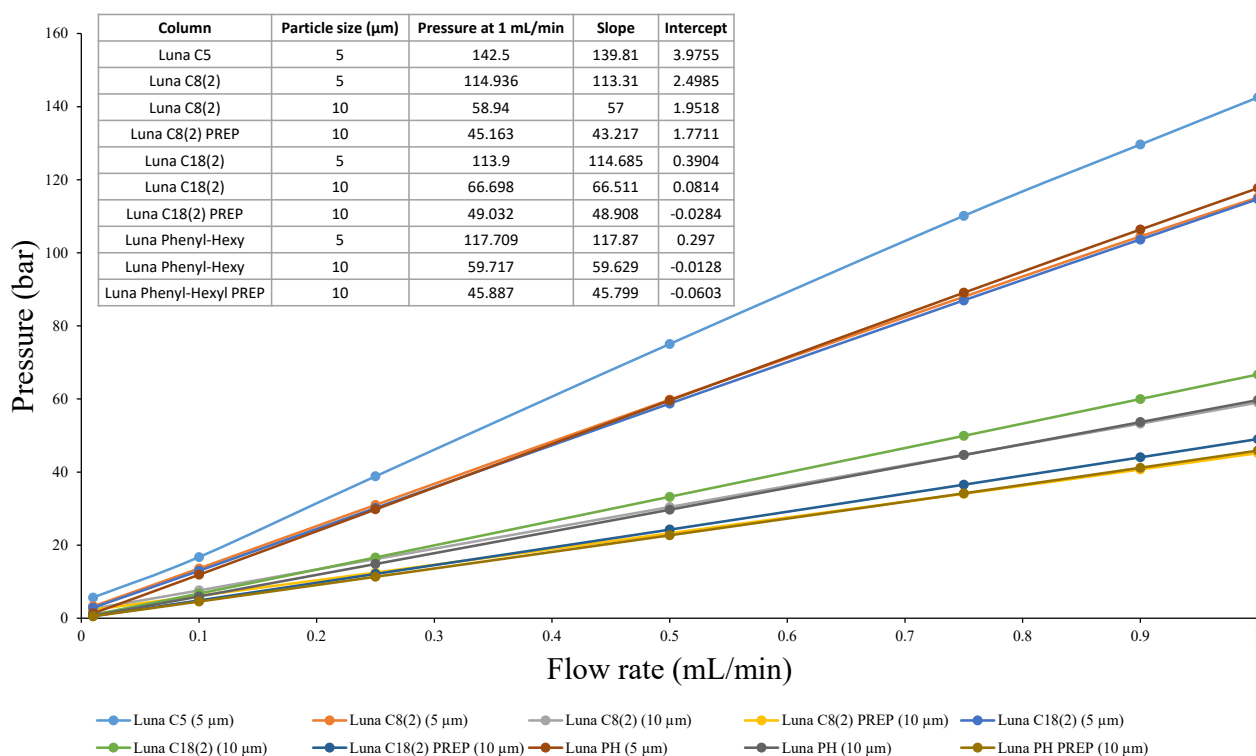


Figure 6.4 | Back pressure test of pure DI water into columns in Table 6.4.

6.3 Methods

6.3.1 Tanaka method

As discussed in Section 3.3, the Tanaka method for column selectivity evaluation is a widely used approach in RPLC that evaluates and compares the performance of different columns. This method provides a systematic way to characterise columns based on their retention behaviour, efficiency, and interactions with solutes. It is advantageous during method development and column selection, offering insights into how well a column can separate specific types of compounds.

At the core of the Tanaka method is a series of probe compounds that highlight different adsorption interaction mechanisms during chromatography. These solutes are chosen to assess various column properties such as hydrophobicity, steric selectivity, hydrogen bonding, and ion-exchange interactions. The method evaluates column selectivity based on five key parameters: hydrophobicity, steric selectivity, hydrogen bonding capacity, ion-exchange interactions, and the

presence of metal impurities, which can influence retention behaviour, particularly for compounds capable of forming metal chelates (see Section 3.3).

The Tanaka method followed in this Thesis is described in Table 6.5. The column temperature was controlled by preheating for 30 minutes in the oven before the first injection. Samples were dissolved in pure solvents to avoid precipitation inside the column due to high concentrations. The method was expanded by injecting increasing the number of arene species used to determine parameter α_{CH_2} which describes the column hydrophobicity selectivity based on a single methylene (-CH₂-) group of difference between molecular probes. The structures of all probes can be found in Figure 6.5.

Table 6.5 | Tanaka test conditions. Runs 1-4 were run as described in the literature[137]. The flow rate was programmed at 0.8 mL/min for all runs to ensure a maximum column pressure of 70 bar. Temperature was kept constant at 40 °C. UV detection was performed at $\lambda = 254$ nm. The sample injection volume was 10 μ L. All samples were dissolved in pure methanol for Run 1 and Expansion and pure water for all the remaining Runs.

	Run 1	Expansion	Run 2	Run 3	Run 4
Parameter	k'_{PB} , $\alpha_{T/O}$ and α_{CH_2}	α_{CH_2}	$\alpha_{C/P}$	$\alpha_{B/P}$, pH = 7.6	$\alpha_{B/P}$, pH = 2.7
Property tested	Hydrophobicity, Steric selectivity	Hydrophobicity	Hydrogen bonding capacity	Ion-exchange capacity at pH7.6	Ion-exchange capacity at pH2.7
Mobile phase composition	Methanol/H ₂ O 80:20	Methanol/H ₂ O 80:20	Methanol/H ₂ O 30:70	Methanol/Phosphate buffer 20 mM 30:70 pH = 7.6	Methanol/Phosphate buffer 20 mM 30:70 pH = 2.7
Sample	0.1 mg/mL Uracil 10 μ L/mL Butylbenzene 1 mg/mL <i>o</i> -Terphenyl 10 μ L/mL Pentylbenzene 0.1 mg/mL Triphenylene	Uracil (0.1 mg/mL) All 10 μ L/mL: Ethylbenzene Propylbenzene Butylbenzene Pentylbenzene Hexylbenzene	Uracil (0.1 mg/mL) Caffeine (0.5 mg/mL) Phenol (0.5 mg/mL)	Uracil (0.1 mg/mL) Benzylamine (0.5 mg/mL) Phenol (0.5 mg/mL)	Uracil (0.1 mg/mL) Benzylamine (0.5 mg/mL) Phenol (0.5 mg/mL)

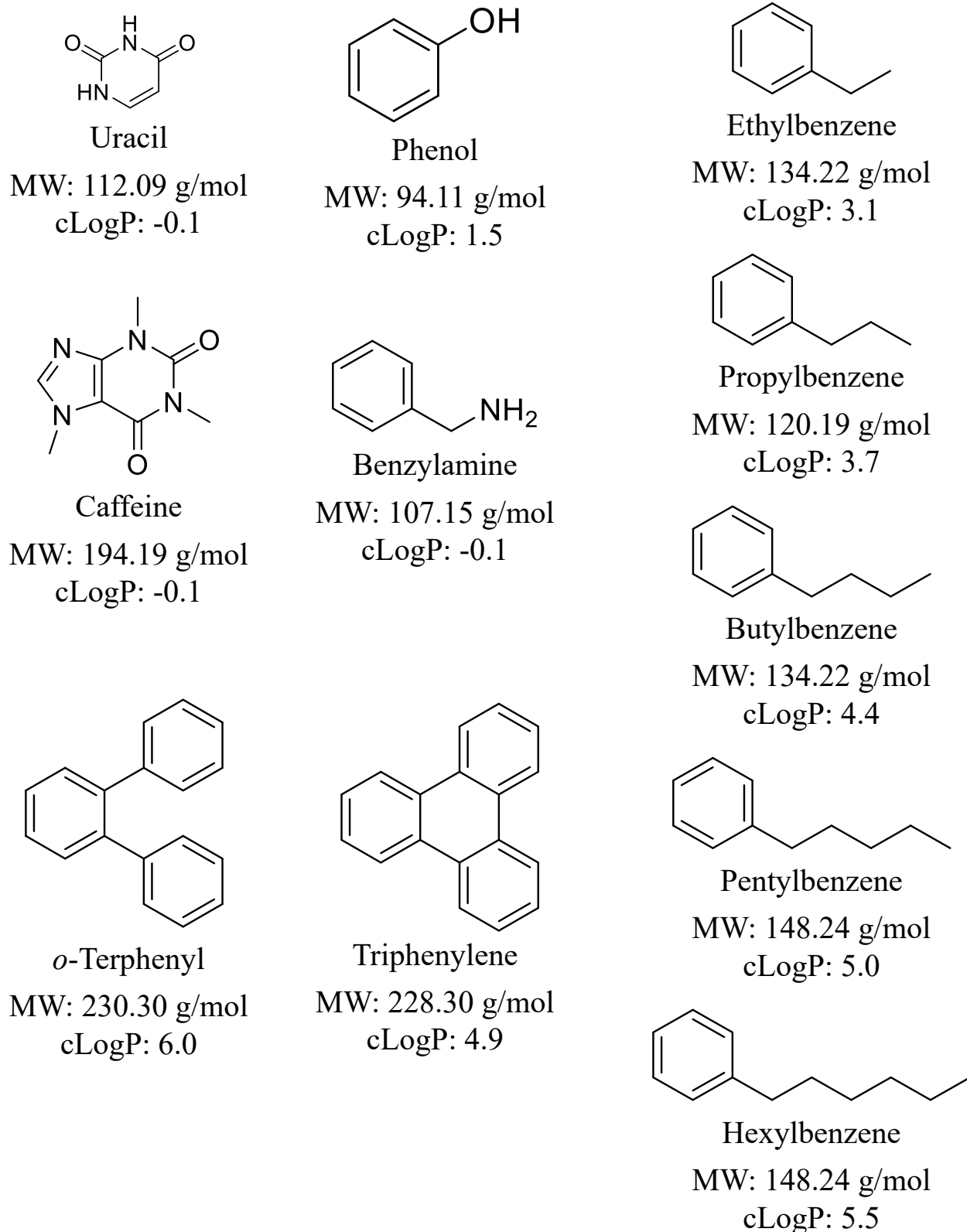


Figure 6.5 | Chemical structure of the probes used in the Tanaka method. Chemical structures of the compounds were generated using ChemDraw software (version 23.0). Molecular weight (MW) and computational LogP (cLogP) were computed using XLogP3 3.0 (PubChem release 2021.10.14).

The Tanaka method was used to characterise column properties comprehensively. It allows for direct comparison between different RPLC columns, making it easier to select the most suitable column for specific molecular separations.

6.3.2 Design of experiments (DoE)

The effects of key chromatographic parameters were assessed using a comprehensive experimental programme supported by a design of experiment (DoE) methodology. The DoE was designed in Minitab version 20.3.0 on a PC running Windows 11 Home version 24H2. The key chromatographic factors chosen are listed in Table 6.6. A full factorial design was generated for all factors except temperature, which was introduced through an L18 design to produce the statistical significance of the effect while reducing the number of experiments. The final number of runs needed was 48 per peptide, giving a total of 432 injections and counting triplicates.

Table 6.6 | Factors used for the design of experiments in Chapter 7.

Factor	Levels			%ACN		
				Lower	Central	Upper
Peptide	P1	P2	P3	-	-	-
%ACN	Lower	Central	Upper	P1	10.0	14.0
Flow rate	1 mL/min	0.5 mL/min	-	P2	18.0	23.0
Ion-pairing agent	0.1% trifluoroacetic acid (pH=1.97)	0.1% formic acid (pH = 2.68)	20 mM Potassium phosphate buffer (pH = 7.00)	P3	28.0	33.0
Temperature	25 °C	40 °C	-	-	-	-

Peptides were injected into column #2 of Table 6.3 in the analytical LC. The key output variables responses considered for this analysis were retention volume (V_r), the area under the curve, peak height, and tailing factor calculated at different peak heights. Using V_r values over t_r is preferred because the latter is a function of the flow rate; thus, it skews the effect analysis.

The effect screening was performed using the standard least squares procedure to assess the fitting of a factor-crossed model through the analysis of variance (ANOVA). Additionally, a principal component analysis (PCA) was also performed on the results to reduce the dimensionality of the experiments.

DoE is a statistical approach that systematically varies experimental parameters to evaluate their effects on the separation outcome. It provides a structured and efficient way to optimise complex chromatographic processes, offering benefits in method development, efficiency, and understanding of the separation process[138].

Using DoE in RPLC for peptide analysis offers several advantages, particularly when developing robust, optimised methods. One of the primary advantages of DoE in RPLC peptide chromatography is its ability to optimise multiple variables simultaneously. In traditional approaches, one variable is changed at a time (OVAT), which can be inefficient and may overlook important interactions between variables. DoE, on the other hand, allows simultaneous variation of factors such as mobile phase composition, gradient slope, flow rate, column temperature, and pH. This approach strategically reduces the experimental plan (total number of runs) while maximising the information gained compared to OVAT designs. By analysing how these factors influence critical outcomes like resolution, retention time, and peak shape, DoE enables a comprehensive system understanding from a molecular interactions perspective. This statistical approach also accelerates method development, leading to better-performing methods less prone to failure.

6.3.3 Frontal analysis for small molecules

Caffeine was used as a reference adsorbate to compare the agreement between FA and PM methods. Both methods were carried out under 40 °C and with a mobile phase solution of methanol/water (30/70, v/v) in accordance with the Tanaka method[70] and to ensure the desorption of the analyte.

FA experiments were performed with a saturated batch with 35 g/L caffeine in MeOH/H₂O (30/70, v/v) and connected to line A of the analytical LC. A second solution of MeOH/H₂O (30/70, v/v) without caffeine was connected to line B. Mixing of these two solutions was performed using the delivery pump unit of the LC to adjust the concentration of caffeine in the mobile phase entering into the column. This allows the production of fronts of different heights that translate into more adsorption data points without changing the LC setup or manually preparing multiple caffeine solutions at various concentrations.

The breakthrough curves of the analyte before and after the column were measured by the pre- and post-column detectors. Signal intensities were transformed to the concentration of analytes, and the mobile volume was subtracted from the post-column detector signal. The area between two curves on the front part was calculated and represented as the moles of analytes adsorbed (n_{ads}). The equilibrium concentration at the front part of the breakthrough curve was calculated using equation 4.22. Similarly, equilibrium concentration was calculated from the rear part of the curve through equation 4.23. Since the manufacturer could not disclose the total mass of the adsorbent ($m_{\text{adsorbent}}$) due to proprietary reasons and the column packing had to be preserved, the $m_{\text{adsorbent}}$ was replaced for the volume of the stationary phase, calculated as $V_{\text{SP}} = V_{\text{C}} - V_{\text{M}}$.

6.3.4 Peak maxima for peptides and small molecules

For PM experiments, caffeine in methanol/water (30/70, v/v) solution of the following concentrations were prepared in 2 ml vials respectively and placed into the autosampler in HPLC: 0.5, 1, 2.5, 5, 7, 8, 9, 10, 15, 20, 25, 30 mg/mL. The following injection volumes were taken for each vial: 1 μL , 10 μL , 25 μL , 50 μL , 100 μL . Each volume of injection was repeated in triplicates for result accuracy. The same method was carried out for all three target columns and with an additional run without any column for comparison. The system was initially equilibrated using the mobile phase solution between testing for each column.

The retention time was obtained from the main peak in the chromatograph. Equilibrium concentration was calculated by modifying equation 4.26 into:

$$q_e = \int \frac{V_r - V_m}{V_s} dC_L \quad (6.38)$$

where V_r is the retention volume, F_m is the flow rate of the mobile phase, t_r is the retention time, V_s is the volume of the stationary phase, V_m is the void volume of the column.

Adsorption data were fitted via multiple-objective optimisation using the following models:

Freundlich model:

$$q_e = K_F C_e^{1/n} \quad (6.39)$$

Redlich-Peterson model:

$$q_e = \frac{K_{RP} C_e}{1 + \alpha_{RP} C_e^g} \quad (6.40)$$

Langmuir model:

$$q_e = \frac{q_m K_L C_e}{1 + K_L C_e} \quad (6.41)$$

BET model:

$$q_e = \frac{q_m K_{BET1} C_e}{(1 - K_{BET2} C_e)(1 - K_{BET2} C_e + K_{BET1} C_e)} \quad (6.42)$$

Anti-Langmuir model:

$$q_e = \frac{\hat{q}_m \hat{K}_L C_e}{1 - \hat{K}_L C_e} \quad (6.43)$$

The best-fitting model was selected by minimising the mean squared error.

6.3.5 Excess isotherm and V_M measurement using MDM

The volume of the mobile phase, also known as the thermodynamic void volume (V_M), as well as the intermediate layer thickness (τ) were obtained from the excess isotherms measured by the minor disturbance method as described by Kazakevich and McNair[126]. The total amount of solvent B per unit of surface area (η_B/S_A) can be derived from mass transfer analysis resulting from equation 4.28. When thermodynamic equilibrium is achieved during an isocratic run, equation 4.29 can be expressed in terms of retention volume (V_R) and V_M .

The HPLC pump was used to mix water (line A) and acetonitrile (line B) in a step series from 0 to 100%B. Each step was run for 20 mL, allowing an equilibration volume of 10 mL between steps. After equilibration, 1 μ L of pure ACN was injected into the column to cause a disruption whilst maintaining infinite dilution of the organic modifier in the mobile phase. The retention volume of the first resulting disturbance peak, either positive or negative, detected by the post-column detector was adjusted by subtracting the retention volume of the injection peak detected by the pre-column detector.

These retention volumes over the concentration range were computed to determine V_M , as explained by the solution to equation 4.30. Excess isotherms were plotted by solving equation 4.28 for excess adsorption using the correspondent calculated V_M to give the following equation:

$$\Gamma_B^v = \int_0^{C_B^1} \frac{V_R - V_M}{S_A} dC_B^1 \quad (6.44)$$

The value of τ was estimated by fitting a negative linear regression to the points of the isotherm where the adsorbent surface is filled (steady reduction of excess adsorption); thus, the equation is true and the intersection with the y-axis represents the maximum concentration of pure solvent B (C_{ads}) at the intermediate layer according to the layer model[139]. From here, the thickness was computed through the following equation 4.35.

6.3.6 SMART method for peptide separation

The present study proposes a new SMART (Systematic, Multi-parametric, Adaptive, Responsive, and Tailored) algorithm for RPLC method development and optimisation. A flowchart of the algorithm can be found in Figure 6.6.

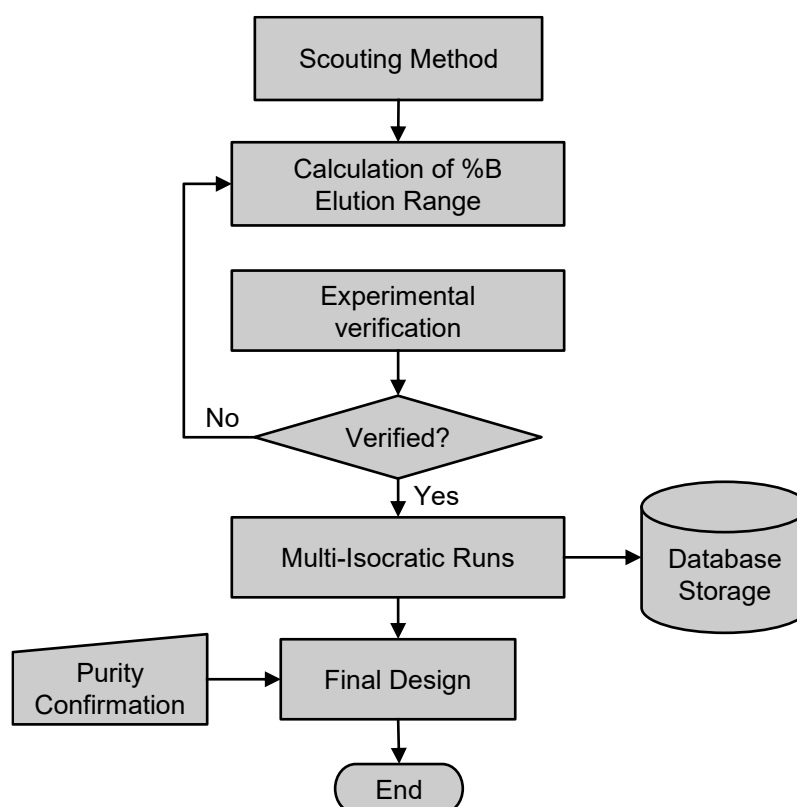


Figure 6.6 | Diagram of the SMART algorithm. See text for description.

The algorithm initiates with a scouting method encompassing 2-3 linear gradient elution methods experiments. These linear gradients are used to compute the average values of k^* and ϕ_B^* for each peak of interest. The steepness of the linear gradients is significantly different, which provides a more realistic understanding of the solvent pump's mobile phase mixing capacity.

The mean value of ϕ_B^* serves as an initial estimate for the elution range during isocratic runs. The isocratic ranging test is chosen in accordance with the predicted behaviour of the analyte. From these runs, a plot of $\log k$ vs ϕ_B is generated. If the S-value is significantly low, the range is extended. These experiments are iteratively performed until sufficient information is gathered for all species (peaks) of interest.

Upon collecting this data, a final experimental design is formulated based on the separation objectives. Typically, final designs comprise three steps: (1) pre-elution or flushing of early species that exhibit less affinity to the resin, (2) focalised elution region, which can be a combination of multi-isocratic runs ensuring optimised separation between the primary species of interest, and (3) flushing of the column, which involves an organic-rich mobile phase that elutes all species exhibiting higher adsorption to the resin. This comprehensive approach ensures a robust and efficient separation process.

Analytical RPLC columns Luna C18(2), Luna C8 and Luna phenyl-hexyl (PH) with dimensions of 150 x 4.6 mm (L x ID) were purchased from Phenomenex (USA). All columns had an average particle size (d_p) of 5 μm , average pore size (n_p) of 100 Å, specific surface area (S_A) of 400 m^2/g , and TMS end-capping according to the manufacturing information. Columns were connected to valve positions 1, 2, and 3 of the Shimadzu system, respectively. Extra pre-column and post-column volumes were estimated to be the same for all ports, around 87 and 172 μL , respectively. Dwell volumes varied between ports, resulting in 1.27 (port 1), 1.49 (port 2), and 1.38 mL (port 3).

Scale-up experiments were conducted in the system described above and in an AKTA Pure 25 system (Cytiva, USA). Solvents were filtered and sonicated to reduce the amount of dissolved gas. An RPLC column Luna C18(2) with dimensions of 150 x 10 mm (L x ID). This column will be referred to as the 'semi-prep column'.

To compare the performance of the purification process, crude and fraction-collected samples were run through the C18 analytical column. The mobile phases were the same as described in the instrumentation. Gradient steepness was 2.5 %B per geometrical column volume.

Chapter 7

Initial experiments

I extend my appreciation to Chingis Idrissov for his support in the execution of the Tanaka experiments for this Chapter during his placement in our lab through the Undergraduate Research Opportunities Programme in July-August 2021.

This Chapter aims to propose an initial set of experiments to determine the descriptive fundamental parameters that describe the peptide separation through RPLC. This constitutes (1) column selection and (2) the selection of representative chromatographic parameters.

7.1 Column suitability and selection

Although the main objectives of the project involve not being dependent on small molecules to mimic peptide separation studies, it is sensible to use a well-established technique during the early stage of the experiments to confirm the suitability of the chromatographic columns. Since the Tanaka method is the most complete CSS developed so far, it has been selected for this purpose. Also, the existence of a database with information in several columns, specifically the ACD labs database[75] provides an ideal initial reference for understanding the chromatographic performance of commercial columns.

Despite the robustness of the Tanaka method, it is indeed a time-consuming methodology. Therefore, a pre-screening method is proposed. This initial method involves a small injection of a mix test sample (2 μ L) composed of uracil (0.10 mg/mL), toluene (10 μ L/mL), dimethyl phthalate (0.2 μ L/mL) and naphthalene (1 mg/mL), allows for a quick yet thorough, initial assessment of the performance of the column. These specific compounds are standard RPLC

probes commonly used by manufacturers to certify column performance after production, and therefore, most columns would come with reference retention values for these molecules.

Six columns (Table 6.5) were meticulously selected from a pool of 25 columns based on the pre-screening mix test sample tested for evaluation using the Tanaka method according to the proposed method in Section 6.3.1. The thoroughness of the pre-selection process increases the likelihood of positive Tanaka results and reduces the number of experiments needed. Section 3.3 describes the parameters resulting from this method.

The results show consistency between the data found in the ACD labs database and many of the columns tested in this work, as seen in the visualisation of results presented in Figure 7.1. It can be asserted that the method is quick and robust, with slight variations to data reported in the literature for most of the columns tested. The most significant difference can be observed with ODS1, where lower $\alpha_{B/P}$ values at pH 7.6 and 2.7 were observed.

The small number of disparities observed can be attributed to different factors. On the one hand, the data in the ACD database does not include enough column information such as batch number or date of manufacture, length or internal diameter of the column. These parameters can affect the retention of the components used in the Tanaka method. In the case of ODS1, the column had been used previously for other applications, which could drastically affect the comparison with newly packed columns. Another valid source of disparity is the packing efficiency of the column. Although it is quite consistent in commercially available columns, batch-to-batch differences between them commonly arise from variations in column packing efficiency. Additionally, Petersson and Eurby[74] reported that a slight change in methanol content, temperature, and the pH of the mobile phase greatly affects observed results.

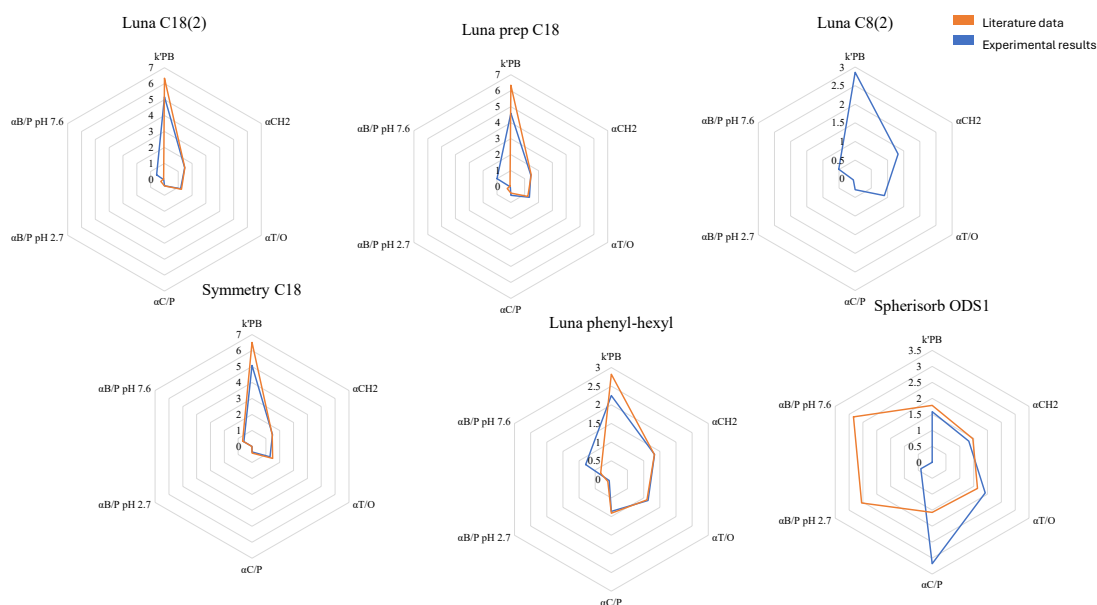


Figure 7.1 | Visualisation of the Tanaka method results for RPLC columns. Radar charts are used for the simple representation of results comparing data obtained in this project (blue) and from the ACD labs database[75] (orange).

As seen in the chromatograms (Figure 7.2 - Figure 7.5), the following elution orders were consistent within columns:

- Run 1: uracil → butylbenzene → *o*-terphenyle → pentylbenzene → triphenylene.
- Run 2: uracil → caffeine → phenol.
- Run 3 and 4: uracil → benzylamine → phenol.

Only Luna Phenyl-Hexyl and Spherisorb ODS1 showed deviations from that behaviour in runs 1 and 2, respectively. Peak swapping in the phenyl-hexyl column was expected since more selectivity is obtained through π -stacking between the benzene rings. Since the structure of *o*-Terphenyle contains two extra benzene rings relative to pentylbenzene, higher retention is expected from the latter in a phenyl-hexyl column.

Deviations in the Spherisorb ODS1 column are also expected to occur due to the lack of end-capping species. However, an unusually high population of silanols is suspected because caffeine (Run 2) and the charged species benzylamine (Run 3) experience high affinity to the resin, probably caused by hydrogen bonding. Once again, it is sensible to speculate that ageing, previous misuse or wrong storage conditions permanently damaged the column. These results prove that

the Tanaka method can be used as a column verification step at different steps in a chromatographic workflow and as a qualitative evaluation of column ageing.

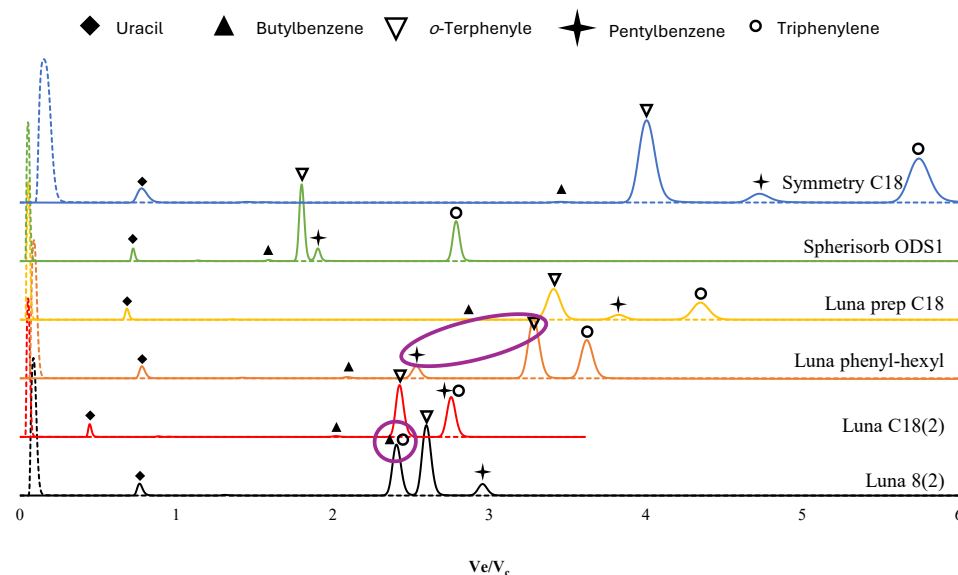


Figure 7.2 | Chromatograms from Run 1 of the Tanaka method. UV responses from before and after the column are represented by the dotted and continuous lines, respectively. The retention volume presented in this graph is relative to the column volume.

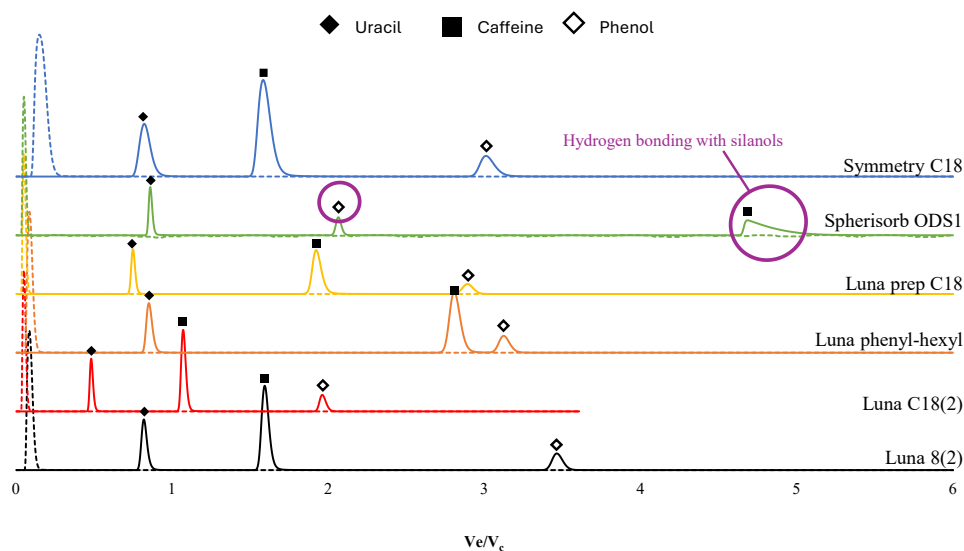


Figure 7.3 | Chromatograms from Run 2 of the Tanaka method. UV responses from before and after the column are represented by the dotted and continuous lines, respectively. The retention volume presented in this graph is relative to the column volume.

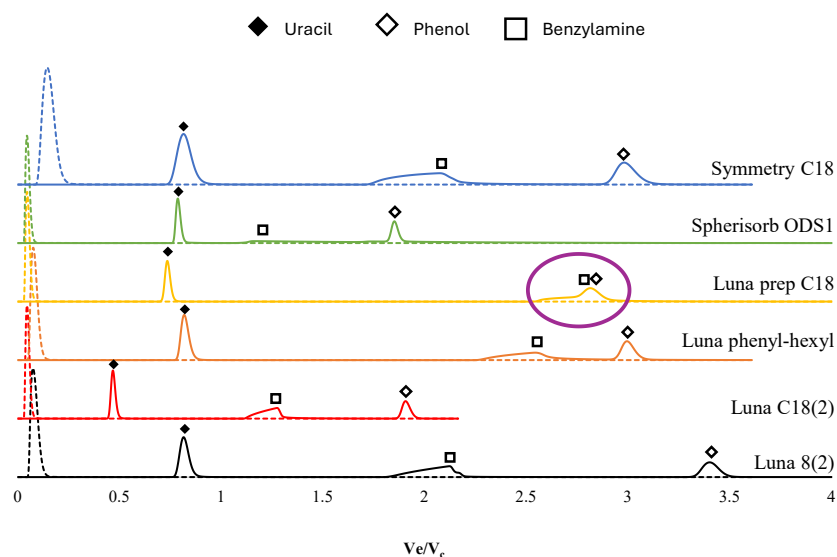


Figure 7.4 | Chromatograms from Run 3 of the Tanaka method. UV responses from before and after the column are represented by the dotted and continuous lines, respectively. The retention volume presented in this graph is relative to the column volume.

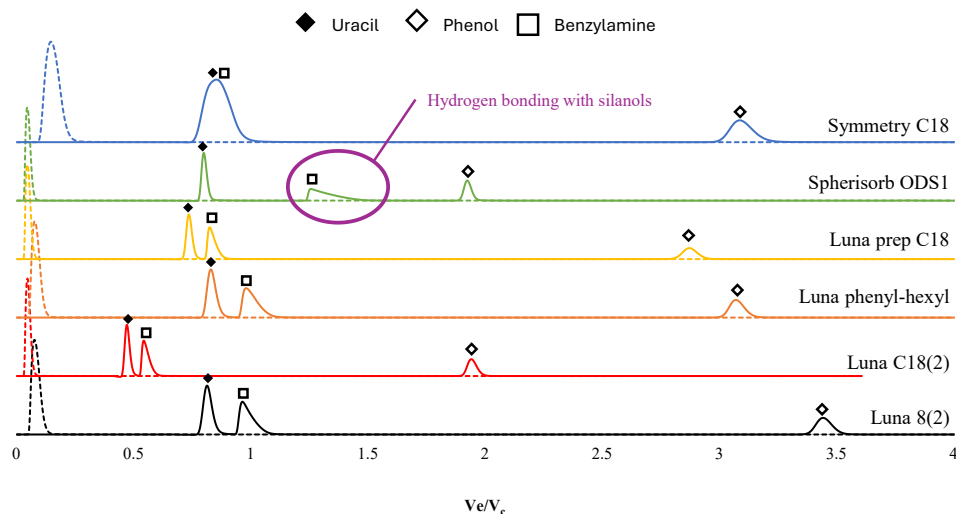


Figure 7.5 | Chromatograms from Run 4 of the Tanaka method. UV responses from before and after the column are represented by the dotted and continuous lines, respectively. The retention volume presented in this graph is relative to the column volume.

Although this method was designed to characterise RPLC columns, it is essential to highlight that only the parameter α_{CH_2} provides clear information on the hydrophobic selectivity of the resin. Therefore, future work on this method includes expanding the number of data points for this

parameter by determining propyl- and hexyl-benzene retention. As shown in Figure 7.6, the α_{CH_2} values vary depending on the probes used; however, this difference was only around 0.1-0.4 units for the columns tested in this work. It is unclear if the slight variation follows a trend, as each column seems to have different responses as the amount of carbon in the probes increases. Since very little difference is observed, this expansion does not show significant relevancy compared to the traditional measurement. Still, it confirms the robustness of the assumption that the selectivity of a methylene group can be analysed through these probes.

The Tanaka method could also benefit from expanding the estimation of the steric selectivity to compounds with different numbers of chained benzene groups. Although this was excluded from this work, it remains an open opportunity for future development.

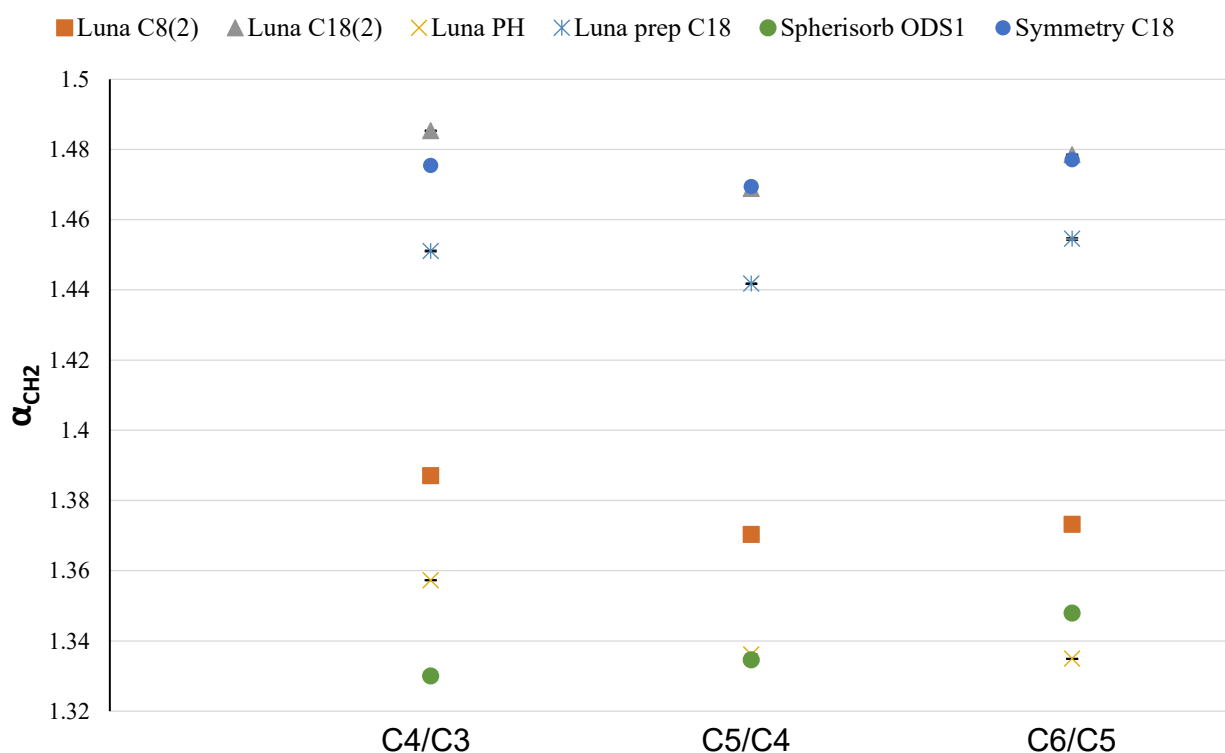


Figure 7.6 | Expansion to the Tanaka method for α_{CH_2} values. The α_{CH_2} values are calculated from the selectivity ratios of butyl-/propyl-benzene (C4/C3), pentyl-/butyl-benzene (C5/C4), and hexyl-/pentyl-benzene (C6/C5).

7.2 Mapping of chromatographic behaviour of model peptides at different conditions

A design-of-experiment (DoE) method was performed as a preliminary test to assess the effect of mobile phase composition and instrument parameters in chromatographic peptide purification. The Krokkin standards (Figure 6.1, Table 6.1) were injected into the Phenomenex C18(2) column selected from the Tanaka method. Samples P1, P5, and P6 were excluded from this experiment to maintain structural similarity and solubility. The effects and factors considered in the DoE can be found in Table 6.6.

The effect screening was performed using the standard least squares procedure to assess the fitting of a factor-crossed model. Results from the analysis of variance (ANOVA) of the least squares fit are shown in Table 7.1.

Table 7.1 | MANOVA analysis of chromatographic response under DoE. Green ticks (✓) symbolise the correlation between the chromatographic response and the effect by statistical significance (threshold = 0.05). Single ticks denote a significance of $p < .001$, and asterisks represent different value ranges for the p -value. Single (*) and double (**) asterisks correspond to $0.001 < p \leq 0.015$ and $0.015 < p < 0.05$, respectively. V_r , retention volume; Area, area under the curve; Tailing F, tailing factor. The percentages in width rows correspond to the width relative to the percentage of the peak height.

Response	Factor				
	Peptide	%ACN	Flow rate	Ion-Pairing Agent	Temperature
V_r	✓*	✓	✗	✓	✓*
Area	✓	✓**	✓	✓	✗
Height	✗	✗	✗	✓	✓*
Tailing F	✓*	✓*	✗	✓	✗
Tailing F (10%)	✗	✗	✗	✓	✗
Width (50%)	✓*	✗	✓	✓*	✓
Width (10%)	✗	✗	✓	✓	✗
Width (5%)	✗	✓**	✓*	✓	✓*

7.2.1 Effect of peptide sequence and organic content

The chromatographic responses significant for peptide solutes were V_r , area, tailing factor and width (50%). These results are expected as these parameters are unique to each analyte. Since the hydrophobicity of each peptide is significantly different (see Table 6.1), the V_r changes in accordance with the concentration of the organic modifier in the mobile phase. The sensitivity

of the separation can be described by the Linear Solvent-Strength (LSS) model. This model is addressed later in Chapter 9.

The mass of the solute injected determines the area and height of the peak. It was observed that the choice of peptide injected does not affect the chromatographic height as this response is correlated to the UV absorbance of the analyte. Since all injections had the same size and came from identical sample vials, the concentration (UV absorbance) was constant, and the changes in the refractive index of mobile phases composed of H₂O and ACN are known to be minimal[140], the height of the peak is not affected. The area is, however, affected as the peak of each peptide deforms (tailing factor) and broadens (width) at different rates based on dispersion in the tubing.

7.2.2 Effect of flow rate

It is important to note that the flow rate does not show any statistically significant effect on the V_r or height of the peaks, suggesting the absence of any dynamic effect during peptide adsorption. The use of V_r over t_r is essential in the analysis of chromatographic behaviour as it removes the dependency of t_r on the mobile phase flow rate. For example, if an analyte is injected into an LC at flow rates F and $0.5F$, the difference in retention between the first and second cases in terms of time will be double. However, if the retention time is normalised to the mobile phase flow rates and expressed in terms of volume, the value is nearly identical.

Surprisingly, the flow rate shows a statistical effect on the area and width of the peaks despite the opposite being true for the tailing factor. This behaviour is possibly caused by multicollinearity between the responses, but further testing with an extended set of analytes, including small molecules with different longitudinal diffusion rates, would be required to draw conclusions.

7.2.3 Effect of ion-pairing agent

All chromatographic parameters show a clear dependence on the presence of an ion-pairing agent in the mobile phase. Ion-pairing agents like TFA and FA reduce the interactions of the peptide/analyte with the active silanol groups on the resin's surface. As discussed in Section 3.2.3, residual surface silanols can interact with an analyte through electrostatic interactions when active (for example, negatively charged) and through hydrogen bonding with the hydroxyl group. The mobile phase at a pH of 7.00 was maintained using a potassium buffer (PB) to account for the effect of pH variance caused by the ion-pairing agents and reduce any impact on the quality of the DI water used in these experiments. The concentration of 20 mM PB was selected as it was found to be the minimum concentration required to produce repeatable results without causing desalting/precipitation issues (Figure 7.7).

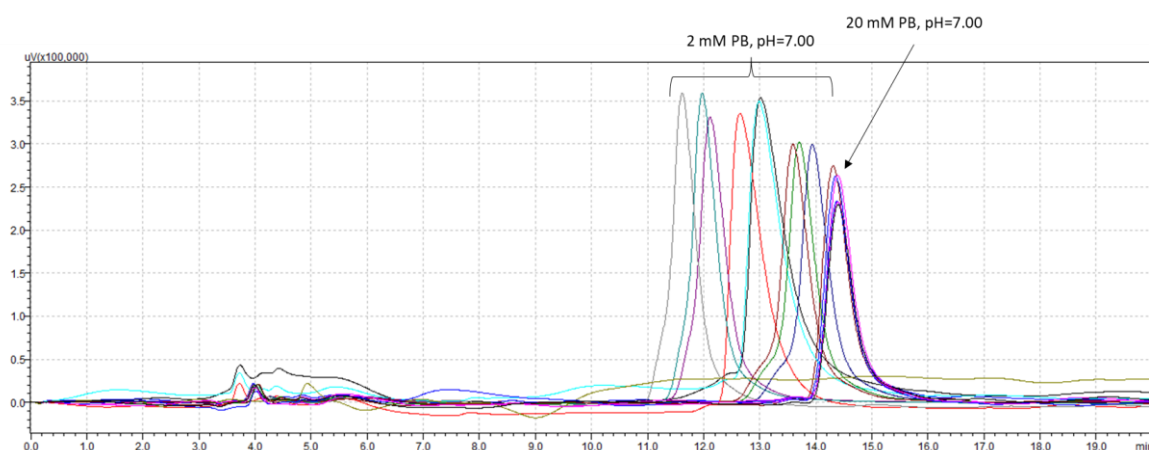


Figure 7.7 | Effect of buffer concentration on retention.

The presence of the ion-pairing agents or buffers is known to affect the chromatographic area and height, as the refractivity and, thus, the maximum relative intensity of the mobile phase is known to be significantly lower than the signal from pure water or H₂O/ACN mixtures due to a masking effect (Figure 7.8).

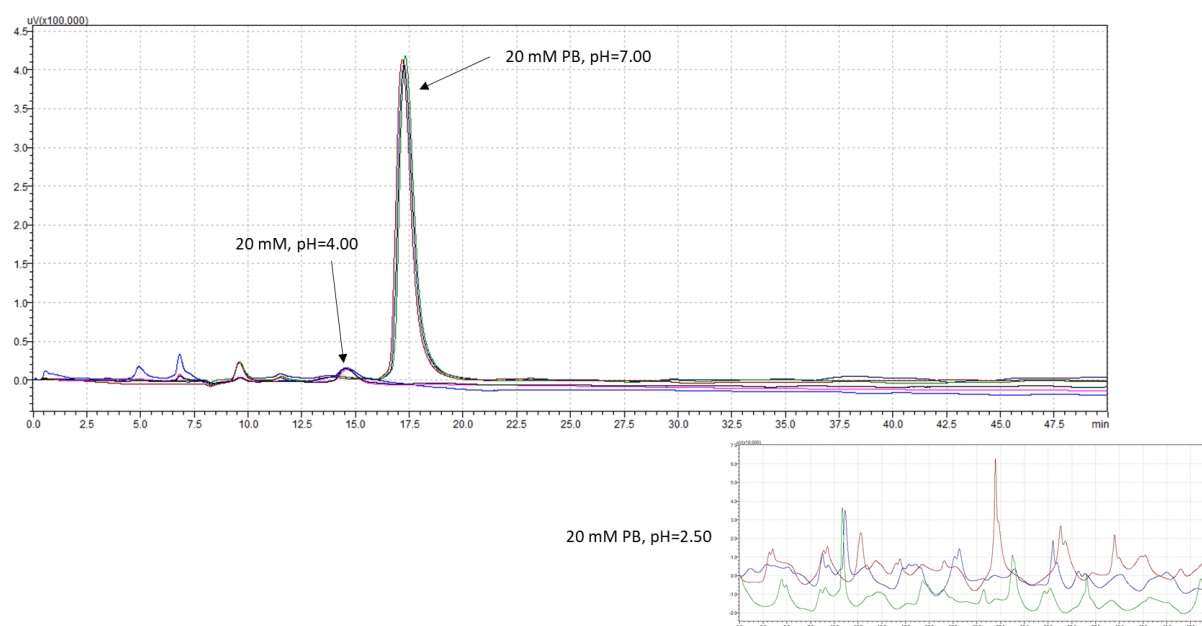


Figure 7.8 | UV signal masking effect caused by salt buffers. A 0.2 mg P2 mass was injected into the RPLC column. UV detection was performed at 202 nm.

7.2.4 Effect of temperature

Temperature is well-known to affect the retention of analytes. Temperature directly affects these interfacial interactions, leading to changes in retention time, peak shape, and separation efficiency. As the temperature increases, the retention of analytes typically decreases. This decrease is because higher temperatures enhance the solubility of analytes in the mobile phase, making them less likely to interact with the hydrophobic stationary phase. In other words, the analytes are more soluble in the mobile phase at elevated temperatures, leading to faster elution from the column and shorter retention times[141-143]. This effect is especially pronounced for larger, more hydrophobic molecules like peptides and proteins, which interact strongly with the stationary phase at lower temperatures[144, 145].

In addition to reducing retention times, temperature can improve peak shape and resolution in RPLC. Higher temperatures reduce the viscosity of the mobile phase, which in turn reduces back-pressure in the system. This allows for more efficient mass transfer of analytes between the stationary and mobile phases, leading to sharper peaks and improved separation efficiency. Temperature can also help reduce peak tailing, a common issue in peptide separations due to strong hydrophobic interactions between the analytes and the stationary phase. Increasing the

temperature weakens these interactions, leading to more symmetrical peaks and better chromatographic performance.

However, the effect of temperature on retention is not always straightforward. In some cases, increasing the temperature can lead to changes in the selectivity of the separation. This change is because temperature affects different analytes in different ways, depending on their size, structure, and interaction with the stationary phase. For example, two analytes that coelute at one temperature may be better separated at a higher or lower temperature due to changes in their relative retention times. Therefore, column temperature optimisation is an integral part of method development in RPLC, particularly when working with complex mixtures of peptides or proteins.

Moreover, temperature can influence the stability of analytes, especially for temperature-sensitive peptides and proteins. Excessive heat can cause denaturation, degradation, or conformational changes in biomolecules, which may affect their chromatographic behaviour and lead to poor reproducibility[146, 147]. As such, it is important to find a balance between using a temperature that optimises chromatographic performance and maintaining the chemical and biological integrity of the analytes.

7.2.5 Principal Component Analysis

Due to the large number of variables tested simultaneously in RPLC experiments, the DoE proposed here produces a large, complex, and multidimensional data set. The data was simplified using principal component analysis (PCA) to deepen the analysis. PCA transforms the original variables into a smaller set of uncorrelated principal components, which capture the maximum variance in the system[148, 149] This reduction potentially makes interpreting the data more manageable and helps us understand the key drivers of variability in the experiment. In this case, the variable 'peptide' was removed from the PCA as it was found to skew the data significantly. This is reasonable as most separation results depend highly on the analyte and resin. In fact, if multiple columns had been tested with this DoE, the PCA results would also be skewed by the column contribution. Instead, HETP and NTP were added to the analysis.

As shown in Figure 7.9, the PCA results can be divided into three distinct groups: red, black, and green. According to the loading plot, the first component is directly contributed to (located on or very close to the vertical axis) by the tailing factor (+), %ACN (+) and V_r (-). And the height (+), temperature (-), and flow rate (-) directly contribute (located on or very close to the vertical axis) to the second component. The negative and positive symbols between brackets after each variable are the direction of the contribution to its respective component.

Other variables showed mixed contributions to both components. Factors that affect the general shape of the peak (area and width) positively contribute to the second component to a greater extent and negatively contribute to the second component to a lesser extent. Temperature contributes negatively to both components but has a higher effect on component 1. HETP and NTP cancel each other contribution on both components (angle between vectors of 180° and equal magnitude), which is surprising but logical as each is inversely proportional to the other. Perhaps a higher contribution disparity would have been observed if columns of different lengths had been analysed.

Based on the analysis of the components, the groups can be described as follows:

- Red: Data points whose variance primarily depends on the retention volume and the general shape of the peak.
- Black: Data point whose variance primarily depends on the peak shape, mobile phase content, and temperature but not so much on the flow rate.
- Green: Data points whose variance primarily depends on the flow rate and mobile phase temperature.

Based on the number of data points in the principal components plot, the majority of the data variance is represented by groups black and green. Therefore, the retention of peptides in isocratic elution is mainly affected by the mobile phase content, temperature, and flow rate. These results are consistent with the MANOVA results. It is important to note that the two PCA components represent only 49.9% of the data variance. Therefore, conclusions based exclusively on PCA are not final; however, they provide relevant information when complementing the MANOVA of a DoE.

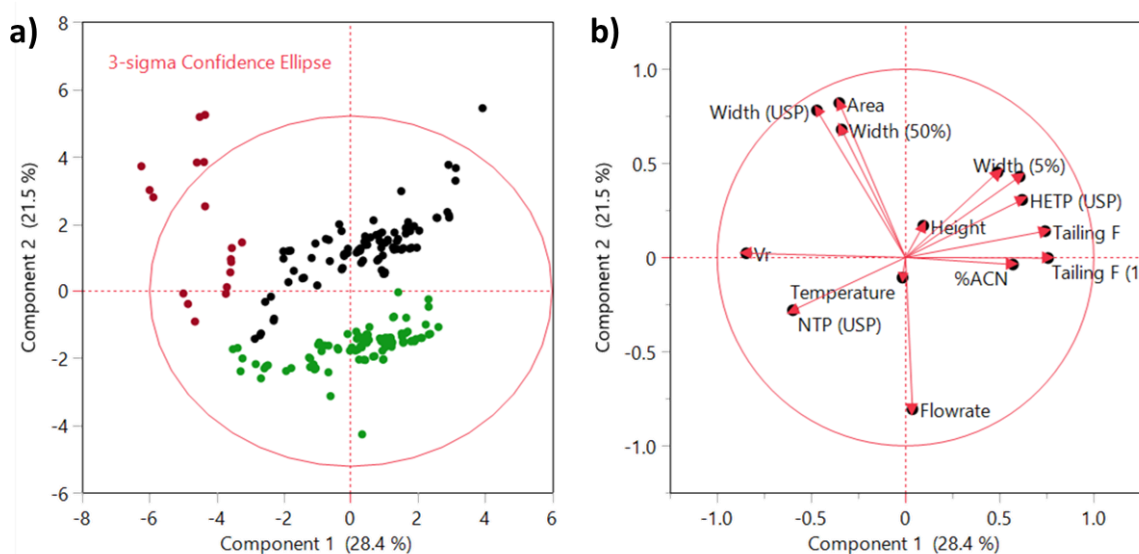


Figure 7.9 | Principal Component Analysis of chromatographic response under DOE. Principal components (a) and the loading plot (b).

7.3 Conclusions

The experiments conducted in this Chapter provided valuable initial insights into the chromatographic behaviour of peptides under various chromatographic conditions, particularly in the context of RPLC. The Tanaka method was successfully employed to quantify the suitability of commercial RPLC columns, demonstrating that this technique is both robust and capable of highlighting variations between different commercial column types, as well as the effects of column ageing and column use history. A pre-screening method was proposed as a time-saving approach for column selection when multiple columns must be evaluated. Minor inconsistencies between data sets highlighted the relevance of factors like batch variations and packing efficiency in commercial columns.

The design-of-experiment (DoE) methodology effectively elucidated the influence of key factors such as mobile phase composition, temperature, flow rate, and ion-pairing agents on peptide separation using RPLC. The findings emphasise the significance of these variables in achieving consistent and reliable chromatographic results. Specifically, the interaction between temperature and retention time was consistent with literature findings, and it was confirmed that temperature optimisation is crucial for enhancing separation efficiency and peak shape. The application of principal component analysis (PCA) helped in simplifying the multidimensional data

set, revealing that most of the variance in peptide retention is attributable to mobile phase composition, temperature, and flow rate. Although PCA alone accounted for less than half of the data variance, when combined with the MANOVA results, it provided a comprehensive understanding of the main contributors to peptide retention behaviour.

In conclusion, the experiments verified the effectiveness of both the Tanaka method and the DoE approach for finding relevant chromatographic parameters. These methods offered a strong foundation for studies to refine the understanding of the fundamentals of peptide separation through RPLC throughout the following Chapters.

Chapter 8

RPLC Adsorption of Small Molecules

I extend my appreciation to Jinhong Peng and Yuxin Sun, who supported this Chapter by executing the experiments and contributing to research conclusions during their Chemical Engineering third-year research project at Imperial College London.

This Chapter compares the adsorption measurements of caffeine as a model small molecule for adsorption onto three different RPLC columns using the peak maxima (PM) and frontal analysis (FA) methods. It provides insights into the adsorption behaviour and sets the groundwork for method selection for more complex molecules.

8.1 Determination of V_M and τ

RPLC columns Luna C8(2), Luna C18(2) and Luna Phenyl-Hexyl with a particle size of 5 μm and column dimensions of 4.6x150 mm (ID x L) as listed in Table 6.3 were selected as they showed good performance in Chapter 7 and through recommendation from industrial collaborators. These columns were used to measure the adsorption of caffeine as a small-molecule probe to understand adsorption. Caffeine was selected due to its small size, straightforward UV detectability, and standard molecule in the Tanaka method. The small size of the probe is essential as we are trying to measure the adsorption of a probe that will have access to all the available pore volumes in the RPCL resin.

As discussed in Section 4.1.1, the V_M of a column represents the approximate volume of liquid inside a column during chromatographic conditions. In other words, it describes the minimum

amount of liquid required to fill the empty spaces in the column, including the pore and the interstitial volumes. Estimates of V_M are experimentally performed by measuring the minimum volume of mobile phase required to elute an unretained molecule, customarily referred to as dead volume (V_0). Under ideal conditions, V_0 and V_M are identical (only the kinetic portion of V_M); thus, they are frequently used interchangeably in the literature. However, reports have shown that no completely unretained marker exists for RPLC columns due to the complex surfaces of these silica particles[102]. Variations of V_0 are attributed to the mobile phase composition and marker molecules properties (the thermodynamic portion of V_M).

Alternatively, assuming that the adsorption of the mobile phase follows the Gibbs excess isotherm, V_M can be estimated thermodynamically for organic/water mixtures in contact with an adsorbent through the minor disruption method (MDM) defined by equation 6.44. The Gibbs excess isotherm predicts how the presence of different substances in the mobile phase affects the overall properties of the liquid in contact with the surface of the chromatographic resin.

The V_M values measured in descending order for columns C8, Phenyl-Hexyl, and C18 are 2.21, 2.17, and 2.11 mL, respectively (Figure 8.1). Since the mobile phase was composed of acetonitrile and water, these values are only valid for ACN/H₂O mixtures. Chapter 11 explores this topic further, including other organic modifiers.

The order of the V_M values observed are not surprising since the apparent size of the ligand is proportional to the length of the ligand when fully extended; thus, the V_M values are inversely proportional. In other words, since C18 occupies more space than Phenyl-Hexyl and C8, the volume available to the mobile phase reduces. It is important to note that this inference is only valid when the ligand surface coverage and the packing density are similar between columns.

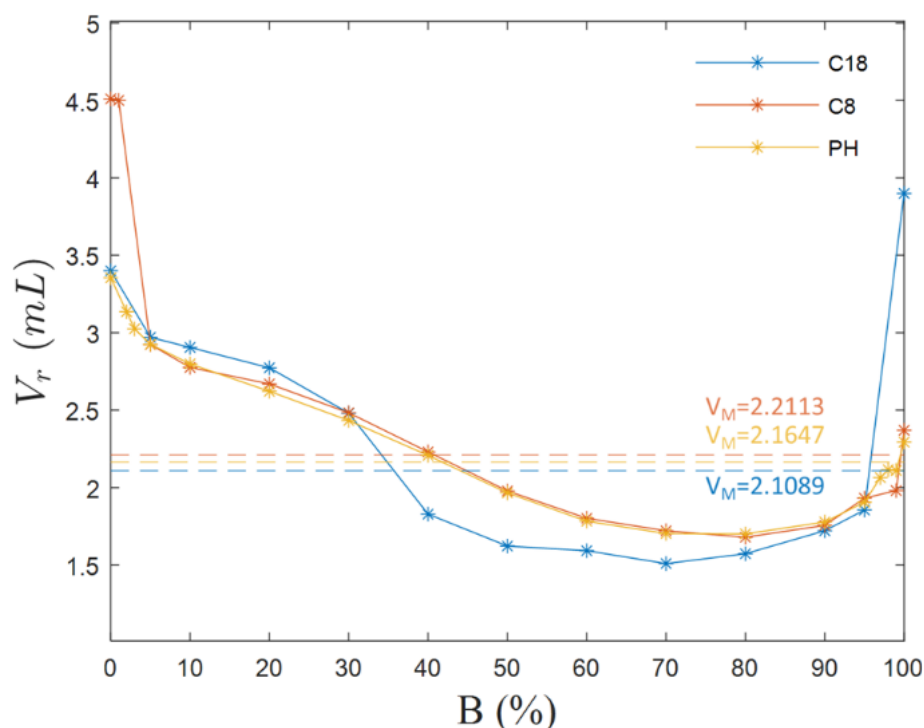


Figure 8.1 | Estimation of V_M values in ACN/ H_2O mixtures for 150-mm length RPLC columns. The organic modifier in the mobile phase, acetonitrile, is represented as B. Percentages are shown as volume/volume ratios. V_M values were calculated using equation 4.30.

From the linear region of the excess isotherms (Figure 8.2), it was found that the C_{ads} value of column the Phenyl-Hexyl column is lower ($19.88 \mu\text{mol.g}_{\text{resin}}/\text{m}^2$) than those of C18 and C8 (22.02 and $22.41 \mu\text{mol.g}_{\text{resin}}/\text{m}^2$, respectively), meaning that each molecule of ACN would occupy a surface area of adsorption of around 7.54 , 7.41 , and $8.35 \text{ \AA}^2$ per gram of resin onto C18, C8 and Phenyl-Hexyl.

By implementing equation 4.35, the thickness of a finite interface layer (τ) of ACN/ H_2O onto the surface of adsorbents within the columns for C18, C8, and Phenyl-Hexyl could be calculated, which approximately 11.54 , 11.75 and $10.42 \text{ \AA.g}_{\text{resin}}$, respectively. Accordingly, the maximum volume of adsorbed mobile phase molecules (V_{ads}) is 0.462 , 0.470 and 0.417 mL , respectively. This accessible interface layer is a stagnant surface film and, therefore, is part of the strict definition of the 'stationary phase'. This implies that of the available V_M , around 22% (v/v) is available for adsorption.

Additionally, I have determined the theoretical diameter of an adsorbed ACN molecule based on the surface areas for adsorption and assuming a molecular spherical structure. The

corresponding molecular diameter based on pure ACN molecules is 3.10, 3.07 and 3.26 Å, respectively. These results represent a 50% reduction in estimated adsorption diameter compared to Monte Carlo simulations based on a single molecular model[150], implying that the interfacial multilayer between an RPLC column and an ACN/H₂O mobile phase is formed by the equivalent of 4 ACN, compared to the simulation. Therefore, it implies that the ACN molecules are densely adsorbed at the RPLC adsorbent/mobile phase interface.

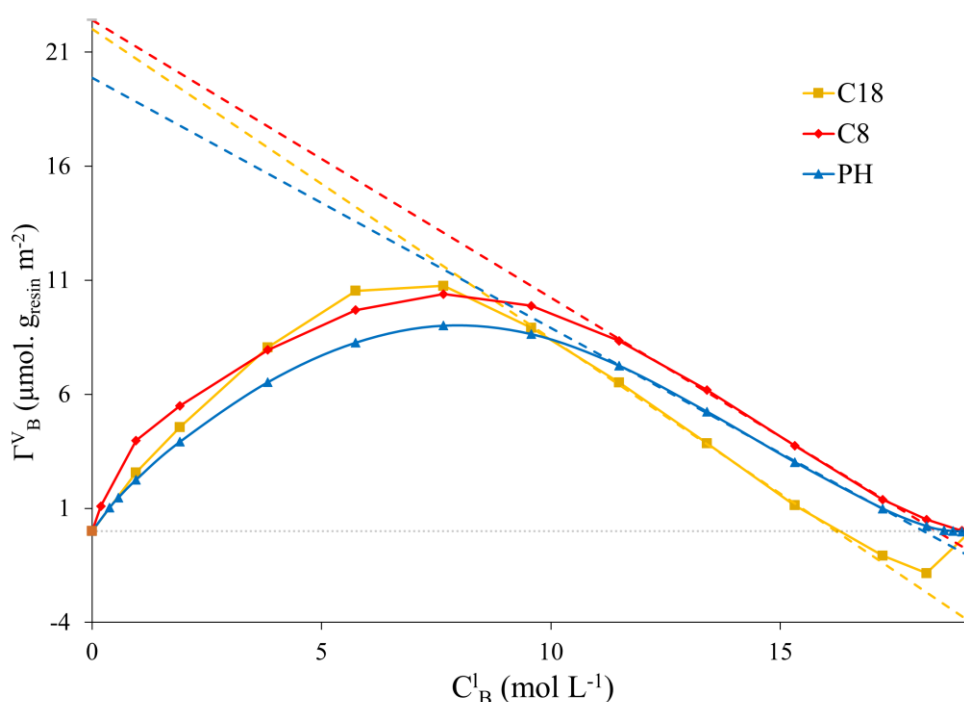


Figure 8.2 | Excess isotherm of ACN/H₂O mobile phase onto 150-mm length RPLC columns.

8.2 Breakthrough curve and adsorption isotherm by Frontal Analysis

As per the principle of frontal analysis (FA), adsorption saturation of column adsorbents is achieved to determine the overall loading capacity at equilibrium; thus, a plateau is formed once the column is fully saturated in that specific concentration. If the analyte follows an ideal adsorption mechanism with only a single adsorption site, a single frontal chromatographic plateau would be expected[151]. However, all columns tested showed the appearance of a second plateau, as shown in Figures 8.3-5, implying the presence of a second adsorption mechanism.

The formation of this second plateau can be explained by (1) the formation of second or multiple layers on the stationary phase or (2) the existence of a completely different adsorption site as a result of co-existing analyte-surface interactions such as hydrophobic interactions at the resin ligand, or hydrogen bonding at the silanols, or electrostatic interactions at active/protonated silanols, or π - π interactions between aromatic rings. The existence of a lower plateau during desorption observed at the rear part of the breakthrough curves supports (1). However, the breakthrough curves alone do not provide enough information to support one of the two assumptions; further justification is carried out in Section 8.4.

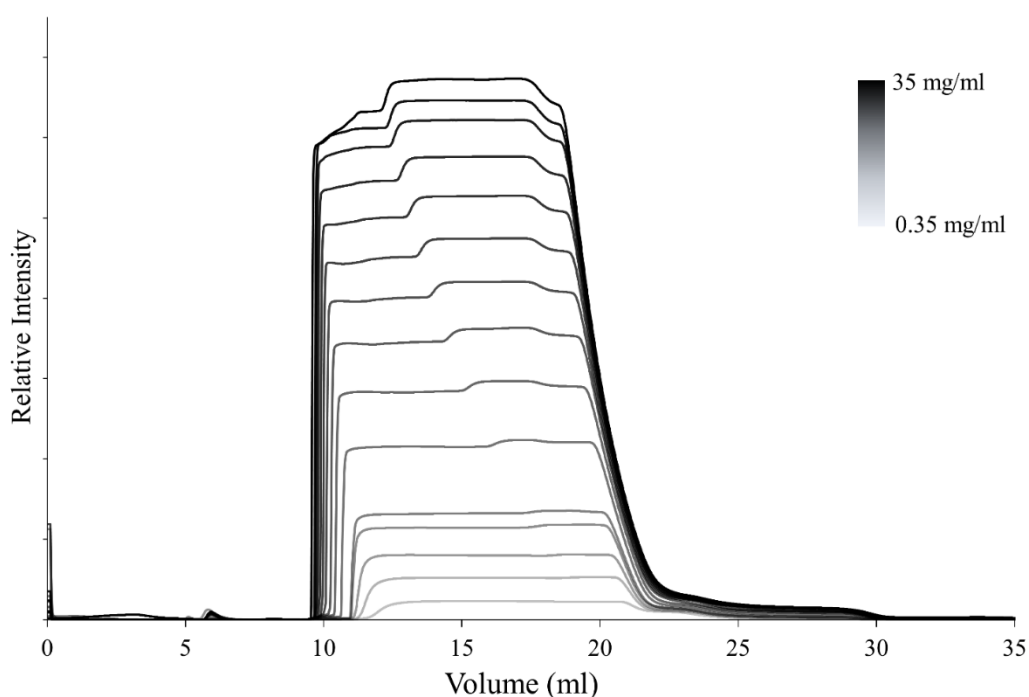


Figure 8.3 | Breakthrough curve of caffeine for C18 column. Recorded by UV detector at $\lambda = 308\text{nm}$.

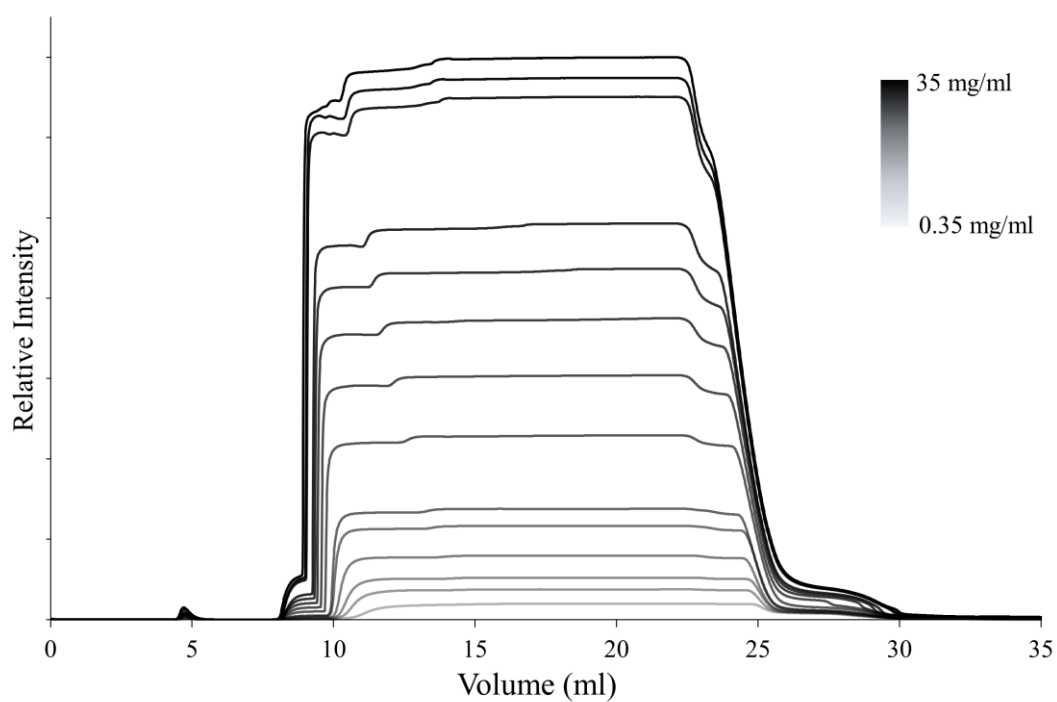


Figure 8.4 | Breakthrough curve of caffeine for C8 column. Recorded by UV detector at $\lambda = 308\text{nm}$.

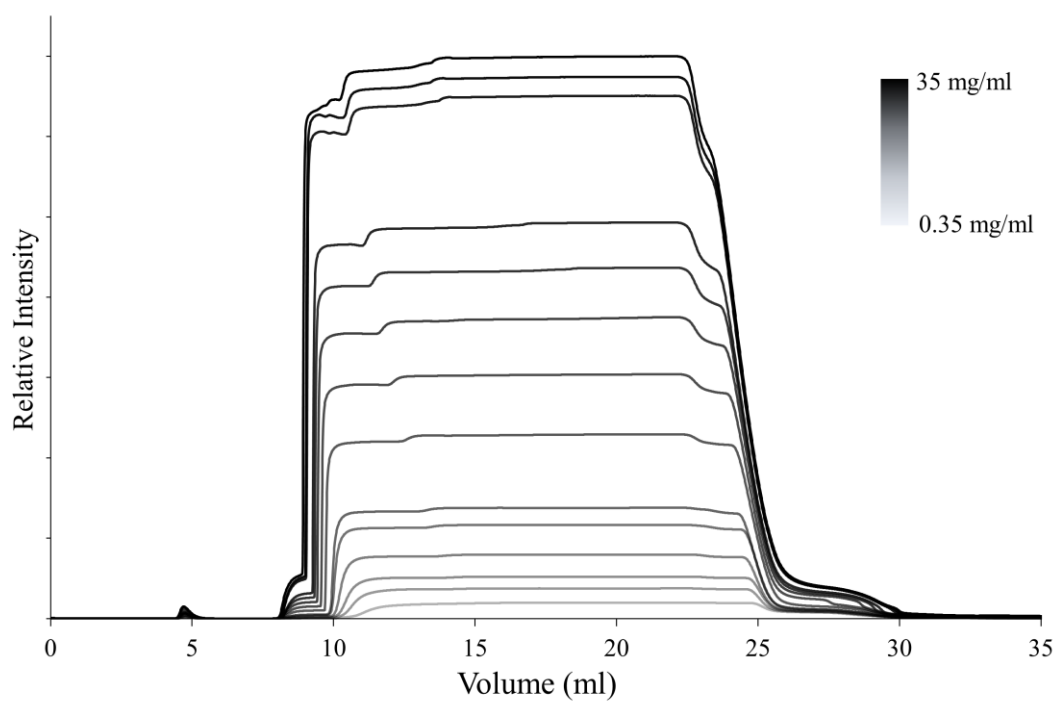


Figure 8.5 | Breakthrough curve of caffeine for Phenyl-Hexyl column. Recorded by UV detector at $\lambda = 308\text{nm}$.

The adsorption isotherms based on the FA data have also been calculated and plotted in Figure 8.6. These calculations were performed using equations 4.22 and 4.23. However, these equations were adjusted, accounting for the recording of the precolumn detector in the analytical LC system. Example chromatograms can be seen in Figure 8.7. The equations are adjusted to calculate the area between curves for adsorption and desorption. This approach provides more accurate results than the standard method, where the FA method has to be run twice, first with the column connected to the system and then without.

Another advantage of the precolumn detector is that it monitors the correct injection front shape. This confirms the height of the plateau and the origin of any perturbation observed by the post-column detector. Using this pre-column data, we confirmed that the appearance of the second plateau is caused by the adsorption phenomena and not by an error in the delivery of the mobile phase.

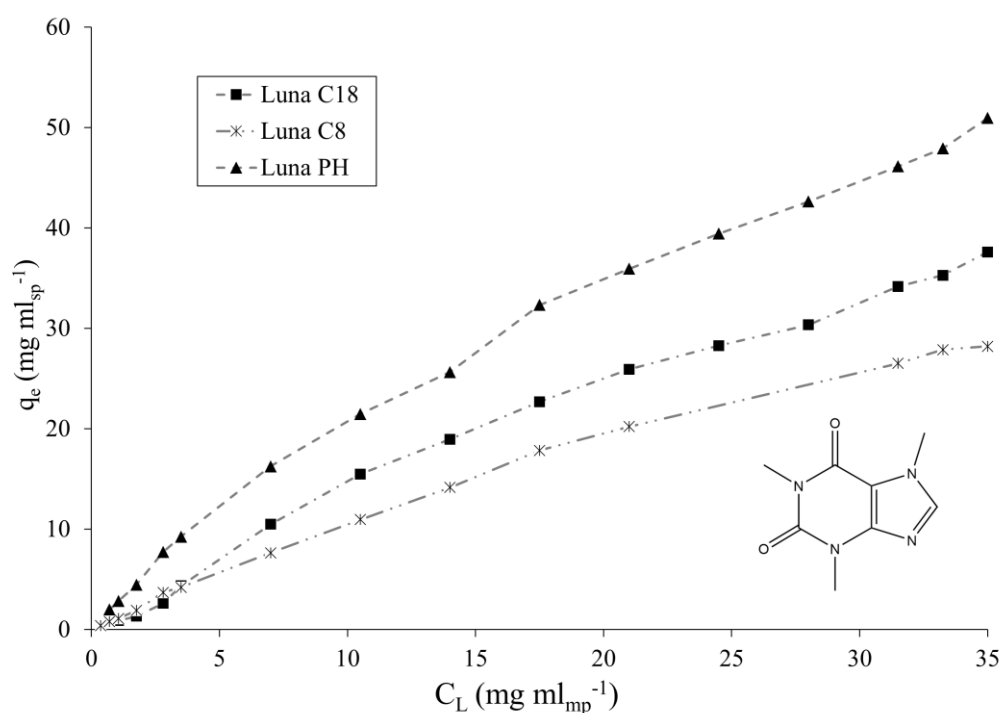


Figure 8.6 | Adsorption isotherms of caffeine for RPLC column by FA method. PH refers to Phenyl-Hexyl.

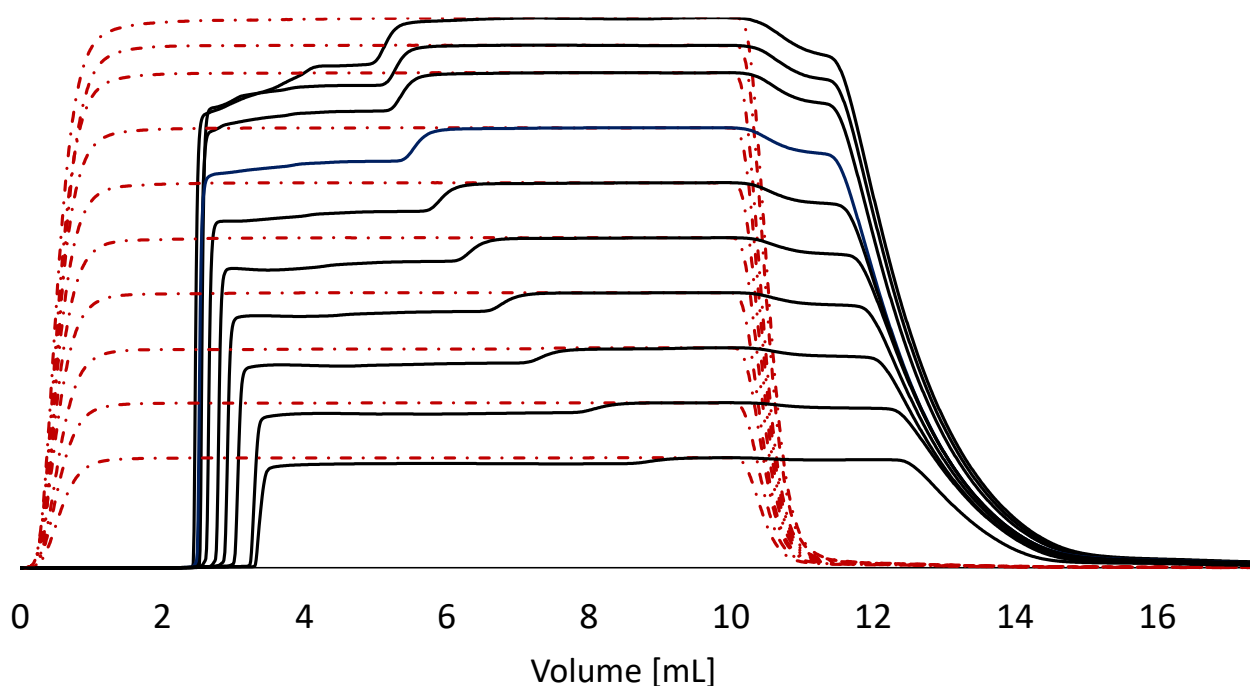


Figure 8.7 | Example of FA data adjustment based on pre- and post-column detection. Red dotted lines are the recorded chromatograms from the pre-column detector. Black lines represent the chromatograms recorded with the post-column detector. Volume is also adjusted according to the response of the first detector.

Since columns C18 and C8 have alkyl chains with different chain lengths covalently attached to the silica surface of the resin, similar hydrophobic and electrostatic interactions between the analytes and stationary phase would be expected, implying that the distribution of analytes and loading capacity in high-energy sites would also be similar. The experimental adsorption data shows a good agreement with this theoretical statement; the equilibrium concentration of caffeine in column C18 shows the same trend and comparable adsorption at low concentrations of the analyte. In the higher range of concentration of analyte molecules, column C18 has a longer alkyl chain, which provides more high-energy sites for analytes to be adsorbed and a relatively higher overall loading capacity. The size exclusion effect is not considered since caffeine is small enough to access all the pore volume within the adsorbent particles[152].

The caffeine isotherm for column Phenyl-Hexyl shows the highest equilibrium concentration among all three columns across the entire concentration range. This result can be explained by the potential π - π interaction between analytes, caffeine, and the aromatic ring (hexyl group) at the head of the ligand on the silica surface. This proposed stronger attraction from the stationary

phase contributes to greater loading capacity on the high-energy site and the overall capacity of adsorption.

8.3 Adsorption isotherm by Peak Maximum

The peak maximum (PM) method was also used to analyse the adsorption mechanism of the three tested columns. As per the principle of PM, the analyte should be injected in an infinite dilution adsorption region, i.e., only a small amount of analyte should be injected into the system. As the analyte molecules are adsorbed onto and desorbed from the stationary phase, the final peak shape is the overall representation of the adsorption and desorption phenomenon occurring within all the infinitesimal spaces along the length of the column. The literature only discusses the application of a sample at a minimal injection volume (1 μL max) in order to produce infinite dilution adsorption conditions. However, there is no fundamental background behind this practice. Suppose the adsorption of analytes is as fast as we expect it to be in RPLC. In that case, the peak shape will only be determined by the injection mass, regardless of the sample concentration or injection volume.

To test our last statement, peak properties with varying injection parameters were compared to expand the concept of infinite dilution. Figure 8.8 shows that the analyte peak is higher and broader, and its maxima appear earlier with increased injection volume for the same sample concentration. This type of peak shape change is typical of a concave isotherm (see section 4.5.1). The same behaviour was also found with an increase in concentration and the same injection volume, as shown in Figure 8.9. This result indicates that in the adsorption process, as more mass/molecules of analytes is/are injected into the column, the elution would result in the same chromatographic peak. Thus, the peak shape is only dependent on the mass of the injection. This statement is confirmed in Figure 8.10 as the same peak shape is observed ($V_r = 4.16 \pm 0.03$ mL; Peak height = $152.9 \times 10^3 \pm 6.66 \times 10^3$ mV; $W_{0.5h} = 0.58 \pm 0.01$ mL) for the same mass of injection regardless of the volume of injection and sample concentration.

Based on these results, it can be affirmed that PM methods that calculate d from peaks of different injection volumes will provide the same information as those where the injection volume was kept constant and the concentration of the sample was adjusted before each injection. This conclusion is particularly useful as the first approach reduces the number of samples that have to be previously prepared, reducing time, the quantity of samples required, and handling errors.

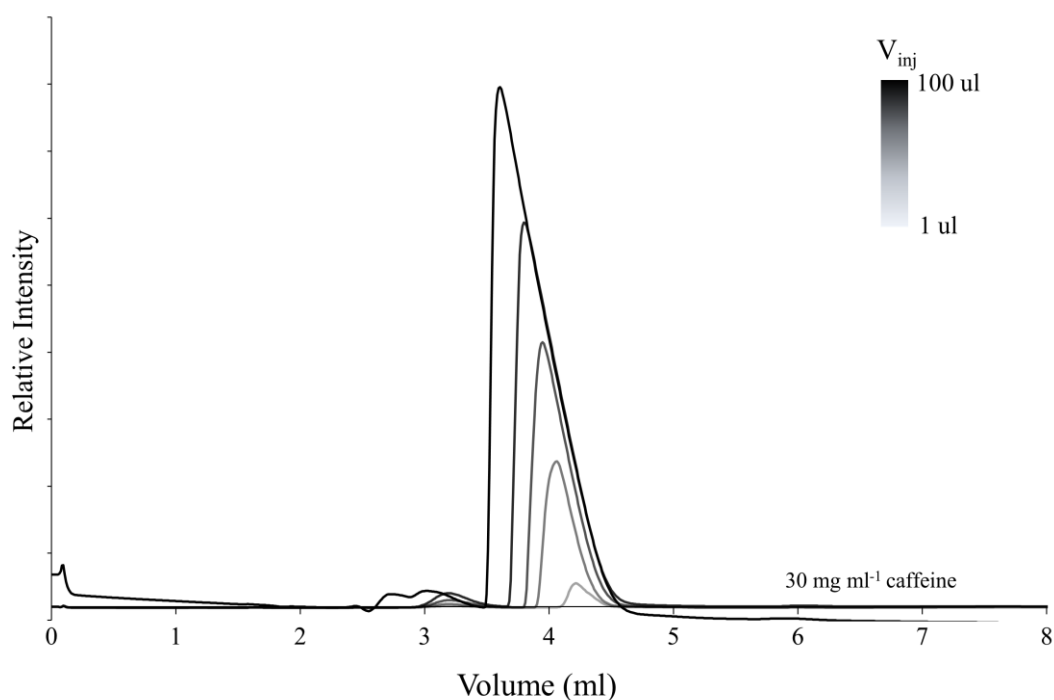


Figure 8.8 | Peak signal response for different injection volumes

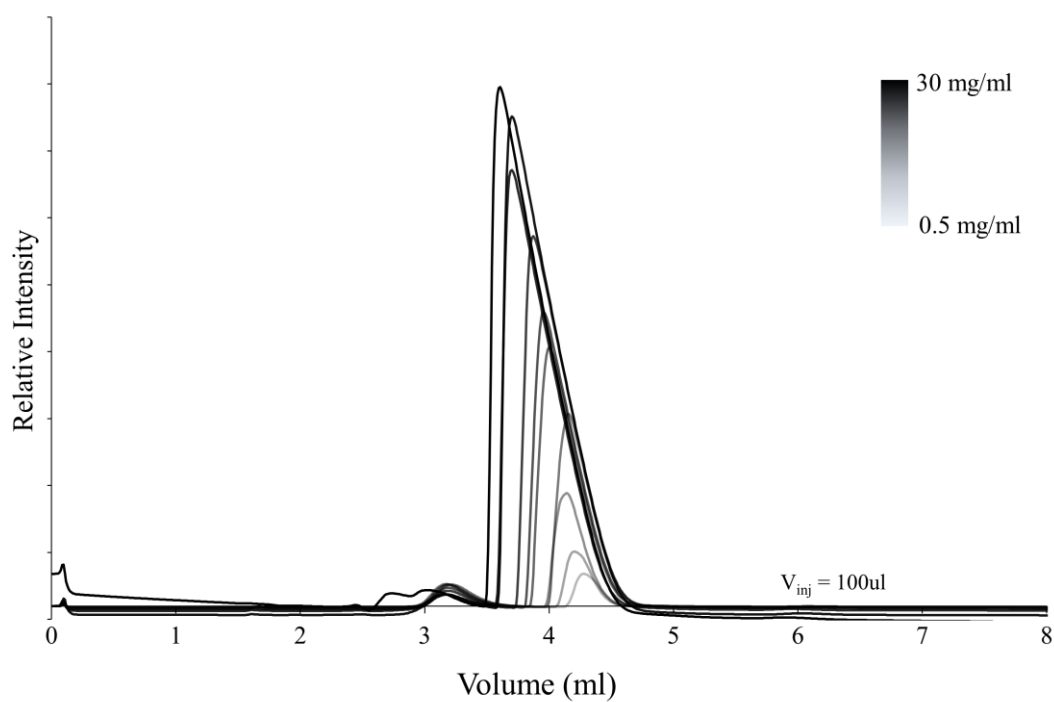


Figure 8.9 | Peak signal response for different sample concentrations.

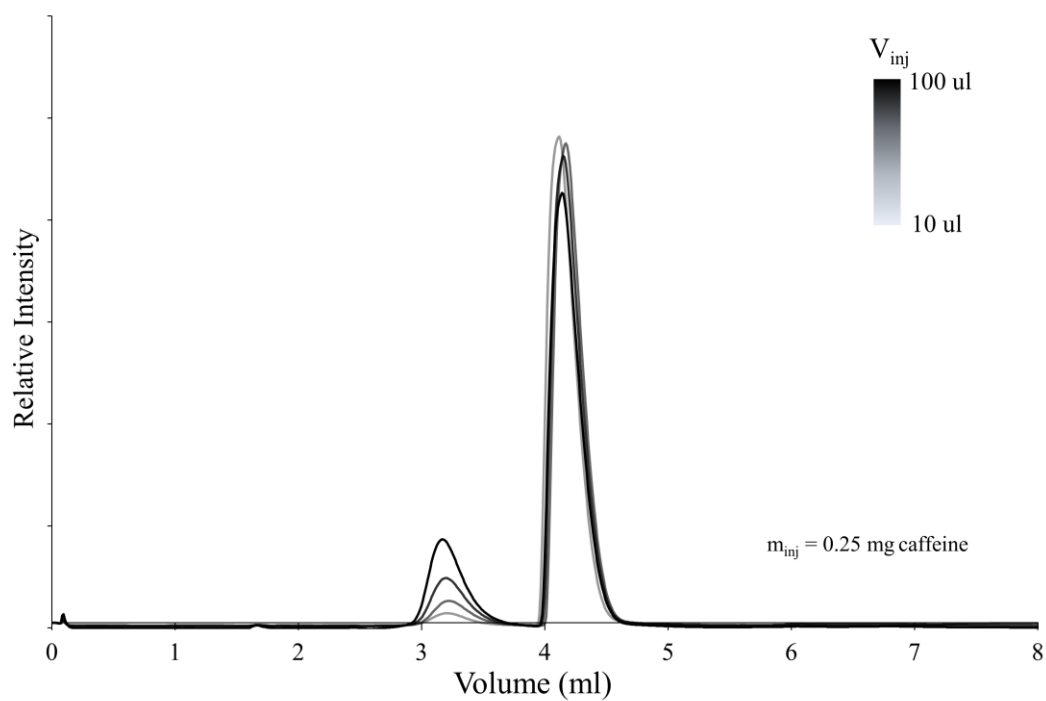


Figure 8.10 | Peak signal response of constant injection mass at different volumes.

It is important to note that this approach should only be applied to PM methods. Other methods, such as ECP, do require a highly concentrated sample injected with a small V_{inj} . Figure 8.11 shows the difference between adsorption isotherms computed using ECP from the peaks of injected samples from the same vial at varying volumes of injections. The isotherm curves downwards as the V_{inj} increases. I speculate that this behaviour occurs because the error of calculating isotherm from a single point, like in ECP and FACP, is higher than in other techniques like PM and FA. Slight variations in the injection will result in higher deviations in the peak shape and, thus, the estimation of the isotherm.

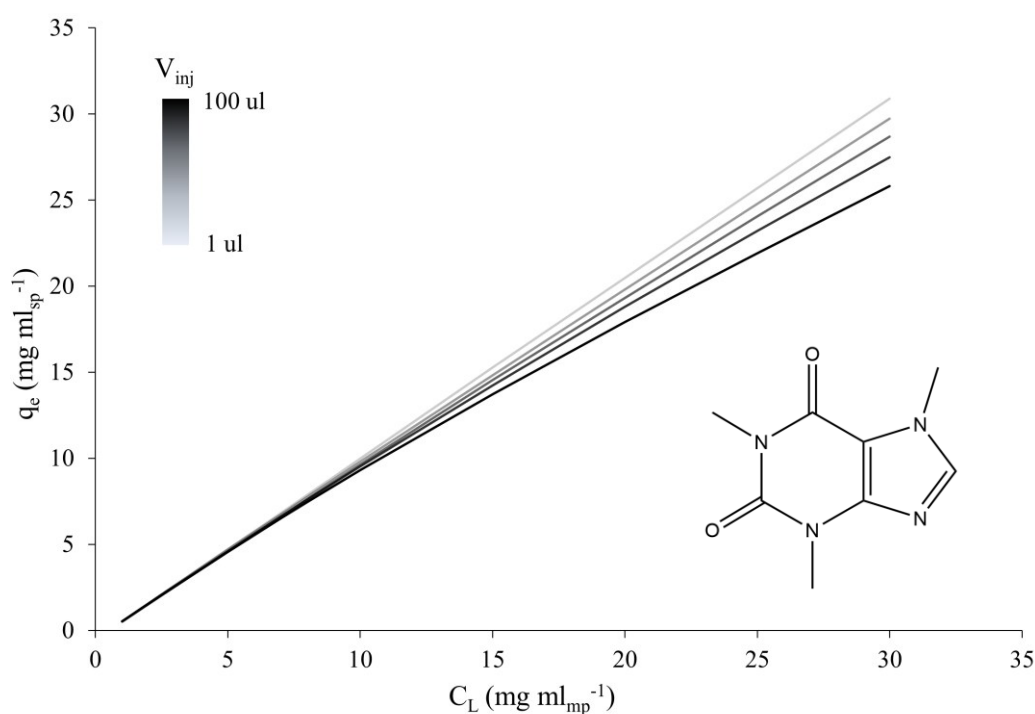


Figure 8.11 | Adsorption isotherms computed using ECP for Luna C18 at different injection volumes. Sample concentration of 15 mg mL^{-1} caffeine.

The adsorption isotherms of caffeine onto columns Luna C18, C8 and Phenyl-Hexyl (Figure 8.12) were computed from the set of chromatographic peaks shown in Figure 8.8 and Figure 8.9 for C18, and the same injections were performed in the other columns. The PM isotherms follow the same trend as those measured through FA, but the adsorption values differ. Additionally, the shape of the PM isotherms is smoother, a common observation for calculations performed by

numerical integration. Since equation 6.38 is cumulative of all experimental points, variations are minimal.

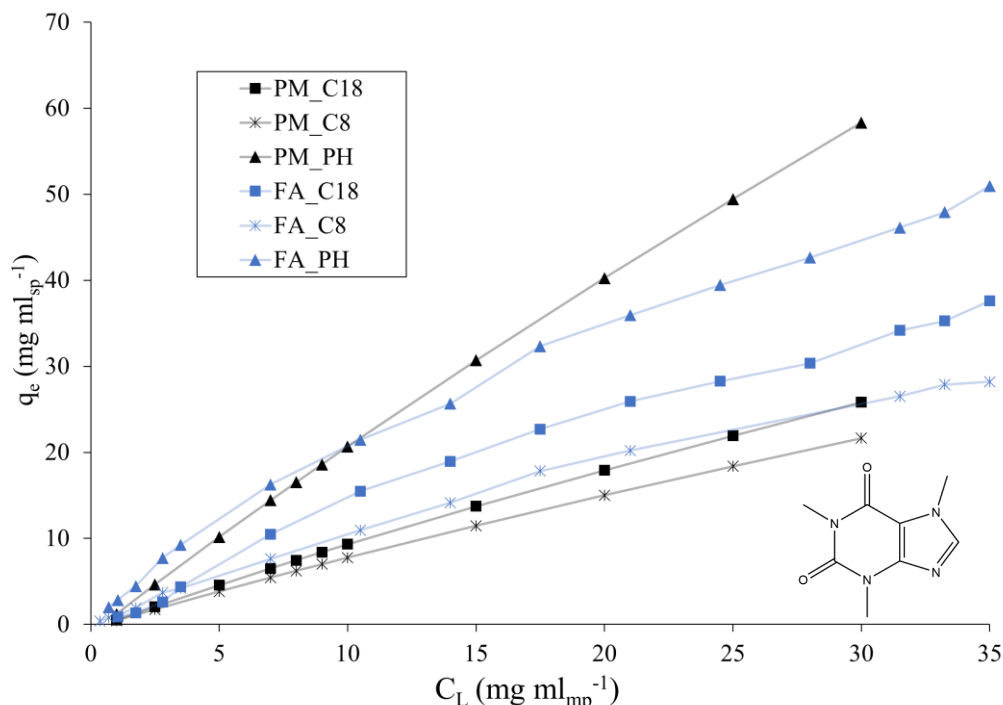


Figure 8.12 | Comparison of caffeine adsorption isotherms calculated using PM and FA. Blue and red plots are the PM and FA isotherms, respectively. PH stands for Phenyl-Hexyl.

8.4 Comparison between isotherm estimation methods

The adsorption isotherms determined by PM have shown that column Phenyl-Hexyl has the highest equilibrium concentration and C8 has the lowest, consistent with the FA isotherms and supports the enhanced adsorption by the π - π interactions between adsorbent and adsorbate. When comparing both methods, PM gives more linear-like adsorption behaviours and intersects their FA counterparts. However, an overestimation of the equilibrium concentration produced by the PM method is expected because the mass of analytes injected ($m_{inj,max}=3$ mg) is significantly lower than that for FA ($m_{inj,max} = 25$ mg). The analytes adsorbed onto the stationary phase in the latter do not saturate the entire column; hence, the equilibrium adsorption concentration is never achieved.

Since all the isotherms had a concave shape, the Langmuir model (equation 6.41) was chosen to fit the adsorption data and obtain isotherm parameters. The results are presented in Table 8.1.

The major difference between methods is in parameter q_m representing the maximum adsorption capacity of the column for PM. FA-based calculation has proved the overestimation of equilibrium concentration compared with PM.

Table 8.1 | Langmuir isotherm parameters for caffeine adsorption onto RPLC columns.

	FA			PM		
	C18	C8	Phenyl-Hexyl	C18	C8	Phenyl-Hexyl
q_m [mg.mL ⁻¹]	108.31	82.25	110.24	258.89	228.53	748.74
$K_L \times 10^3$ [mL.MP. mg ⁻¹]	15.20	15.03	23.00	4.05	3.00	3.02
SSE	10.43	1.12	8.43	0.33	0.20	1.46
q_m [μmol.mL ⁻¹]	610.04	423.56	567.68	1332.80	1191.60	3855.22
$K_L \times 10^3$ [mL.MP. μmol ⁻¹]	2.52	2.95	4.50	7.19	6.70	5.49
SSE	6.19	0.30	2.23	0.09	0.054	0.387

The BET model (equation 6.42) was selected to validate the formation of a second adsorption layer mentioned in Section 8.2, due to its derivation from a multi-layer physical adsorption mechanism. When the model is fitted, the parameter K_{BET2} nullifies, meaning that the model becomes Langmuir. Hence, there is no more loading capacity apart from the first layer, which denies the upper layer hypothesis.

Gritti and Guiochon concluded that the organic-rich layer that forms has a thickness of 4 Å based on the excess isotherm of MeOH/H₂O mixtures calculations. This thickness is insufficient to allow the caffeine molecules to create an additional layer or any form of layer stacking. However, they proved that the organic-rich layer formed from ACN/H₂O mixtures measures 13 Å, which provides enough space to create a second layer and display BET adsorption behaviour[111]. Since the amount of resin packed into a commercial column is proprietary for each manufacturer, it is nearly impossible to compare our thickness layer results with those obtained from self-packed columns without destroying the column.

Similarly, when the Redlich-Peterson model was employed (equation 6.40), the exponential parameter g was equal to 1, and the model converted into the Langmuir model, which indicates macroscopic homogeneous adsorption. However, since the active surface silanols cannot be capped with current technologies[153]. There should always be at least two sites with different adsorption energy: alkyl ligands and silanols. Therefore, the surface chemistry of silica particles

directly contradicts the idea of a single layer with homogenous adsorption onto a single adsorption site.

One potential explanation is that, even though the silica material is theoretically heterogeneous, single-layer adsorption is possible when (1) the difference in adsorption energies between the analyte molecules and the two adsorption sites is so significant that the main interactions are either hydrophobic interactions or electrostatic but not both, (2) the desorption energy exceeds that of adsorption or (3) the distance between ligands is greater than the hygroscopic diameter of the analyte molecule.

In the case of my results, I suspect that (1) or (2) applies. I believe the gap between the adsorption energies of the two sites and the desorption energy changes depending on the organic content. This gap is relatively small at lower organic content levels, and the analyte can interact with all available sites. As the content increases, the gap widens until desorption energies exceed other energies and adsorption is no longer possible; thus, analytes elute instantly (close to V_M).

8.5 Conclusions

FA is a more suitable method for obtaining chromatographic adsorption data and estimating the maximum column capacity since complete column adsorption saturation is achieved. However, this method requires large amounts of analyte compared to PM. Although that is not of much concern for small molecules, it is very relevant for peptides and low-available/expensive drugs. It is also crucial to note that the experimental conditions of FA are limited to the solubility of the peptide because of the risk of gelation inside the column. PM is beneficial in these cases and was selected for the experiments in Chapter 9.

Even though the PM method cannot provide accurate adsorption parameters when compared to FA, it provides valuable information regarding isotherm trends to compare the adsorption of probes onto the same or different columns.

For future work, phenol can be used as a probe to compare its adsorption with caffeine, revealing the effects of hydrogen bonding on adsorption. This comparative analysis can also be expanded to all probes used in the Tanaka method, as it would provide valuable insight into the columns' adsorption behaviour.

Chapter 9

On Peptide Adsorption

I extend my appreciation to Jinhong Peng and Yuxin Sun, who supported this chapter by executing the experiments and contributing to research conclusions during their Chemical Engineering third-year research project at Imperial College London.

This Chapter explores the intricate world of peptide adsorption, with a specific focus on the complex interplay between peptide hydrophobicity, adsorption onto the stationary phase, mobile phase composition, interface/interphase structure, and subsequent analyte elution. Here, I have characterised the adsorption isotherm behaviour of standard peptides under both chromatographic 0adsorption and elution conditions, with a keen eye on the role of ligand coverage, peptide dimensions, and π - π and hydrophobic interactions. Our analysis of these complex interactions aims to provide significant insights into the field of adsorption chromatography and to push the boundaries of peptide purification process optimisation.

9.1 Peptide chromatographic retention during isocratic elution

The chromatographic elution order of peptides P1-P6 from an RPLC column (Figure 9.1) correlates with the increase in peptide hydrophobicity resulting from the replacement of two glycine residues (Gly, G), a neutral amino acid, with leucine (Leu, L), an amino acid characterised by its hydrophobic properties, as observed in the transition from P2 to P5. This identical effect is replicated through either the introduction of an additional leucine (P5 to P6) or the substitution of a glycine and a serine (Ser, S), which is a polar uncharged amino acid, with a hydrophobic phenylalanine (Phe, F) (P1 to P2). The increased hydrophobicity of these biomolecules results in strong adsorption onto RPLC stationary phases. Consequently, an optimal solvent strength at the mobile phase is necessary to induce peptide desorption. In other words, peptide solubility in the mobile phase overcomes the adsorption energy at this solvent strength, facilitating elution.

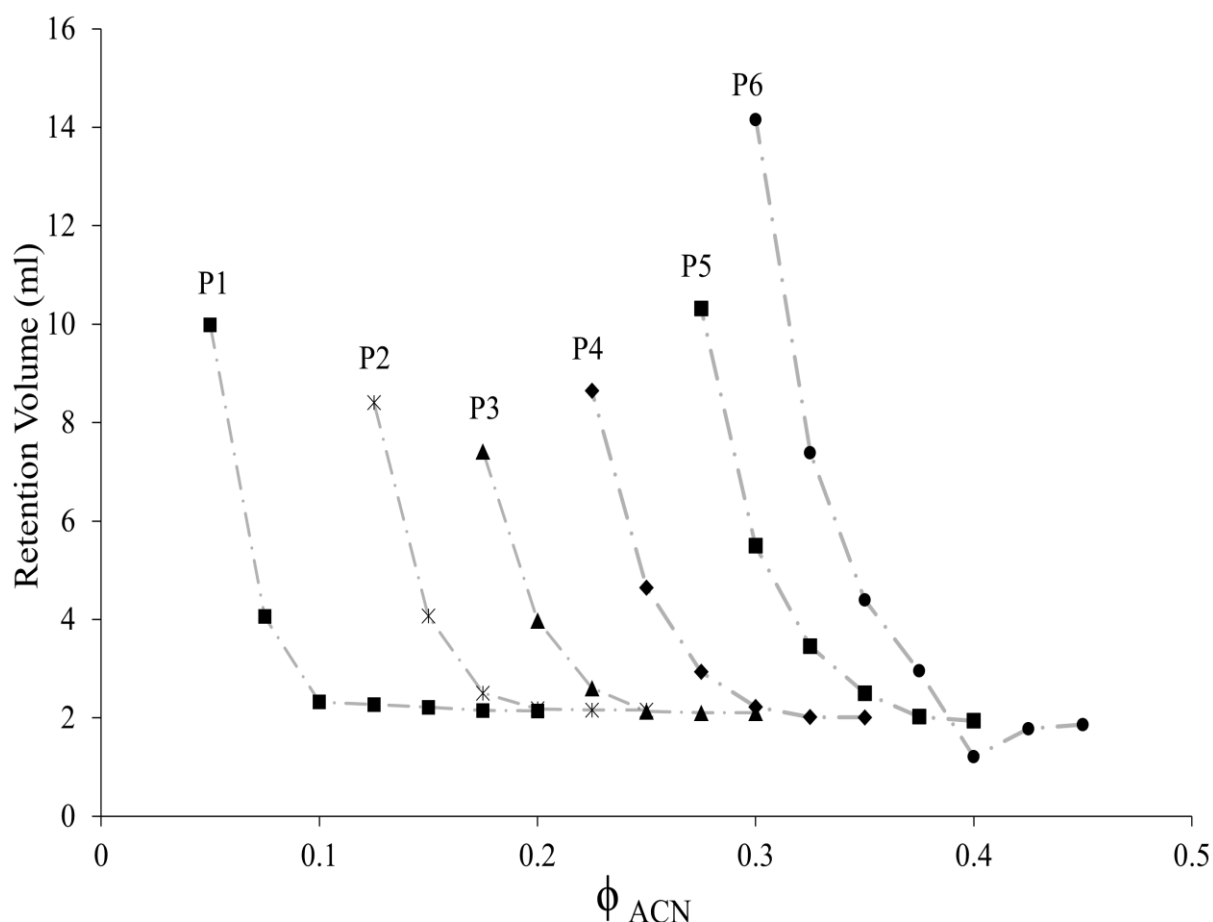


Figure 9.1 | Retention volumes of peptides for Luna C18(2) column. The ACN content (ϕ_{ACN}) is expressed as volume-to-volume ratio. The mobile phase was ACN/H₂O with 0.1% TFA.

It can be observed that the elution range of peptides, that is, the range of ϕ_B concentrations in isocratic elution where analytes elute, becomes narrower as their hydrophobicity increases, meaning that control over the organic content of the mobile phase becomes crucial to facilitate optimal adsorption and desorption for highly hydrophobic peptides. It is also important to notice the asymptote in retention volume as the solvent strength increases and each peptide elutes instantly. The asymptote is unique for each peptide, and values are higher than the mobile phase volume of the Luna C18 column ($V_M = 1.64$ mL), which were calculated from the minor disturbance method for acetonitrile and water with 0.1% TFA (Figure 9.2). This discrepancy indicates a reduction in the accessible volume experienced by the peptides, possibly attributable to size

exclusion effects or high ligand density on the surface of the resin. Therefore, it can be referred to as the apparent or adjusted void volume (V_0^*) unique for each analyte.

Size exclusion effects have been studied as the basis of SEC[154, 155] and have been observed in RPLC through the separation of proteins at elution conditions. While size, typically quantified through hydrodynamic radius, constitutes the primary factor driving size exclusion effects originating from pore blockage, potential repulsive interaction arising from opposing properties between the resin or ligand and the biomolecule can similarly lead to exclusion phenomena. Limited information is available in the literature concerning these opposing forces in chromatography.

The results show a volume reduction ranging from 10 to 30% ($[V_M - V_0^*]/V_M$). As expected, the largest peptide (P1) exhibits a more pronounced reduction in accessible volume compared to the shortest (P5), showing the effect of size exclusion effects. However, upon assessing the collective loss of accessible volume relative to peptide length, a constant reduction of 50 nL per residue was found for all three columns.

Another type of exclusion effect depends on the ligand density on the surface. Specifically, when the ligand density is too high, steric hindrance limits the interaction of the analyte/solute with the high-energy adsorption sites present within the ligand chains on the particle surface. The ligand density for C8 and Phenyl-Hexyl are approximately equivalent based on the Berendsen-de Galan equation (equation 3.2), at 3.48 and 3.37 $\mu\text{mol.m}^{-2}$, respectively. Unsurprisingly, that of C18 is comparatively lower, determined at 2.35 $\mu\text{mol.m}^{-2}$. While this calculation overlooks the presence of the end-capping group, a parallel qualitative comparison can be anticipated, given that the functionalisation efficiency is inversely proportional to the length and mobility of the ligand due to steric hindrance[156, 157].

This exclusion effect can also influence the excess isotherms (Figure 9.2). This type of isotherm measures the adsorption of the solvents in the mobile phase. Since adsorption is influenced by different factors, including the mobility of the bonded-phase group and its retraction or extension (ligand collapse)[158]. The analysis of mobile phase excess adsorption can be complicated. Chapter 11 discusses this topic in detail with additional experiments.

In this scenario, columns C18 and Phenyl-Hexyl have a lower V_M compared to column C8. Therefore, it can be inferred that the ligand C8 in ACN/ H_2O + 0.1% TFA is indeed shorter than the other

two. Notice that the addition of TFA as the acidic modifier reduces the value of V_M by 25% for all three columns (compared to Figure 8.1).

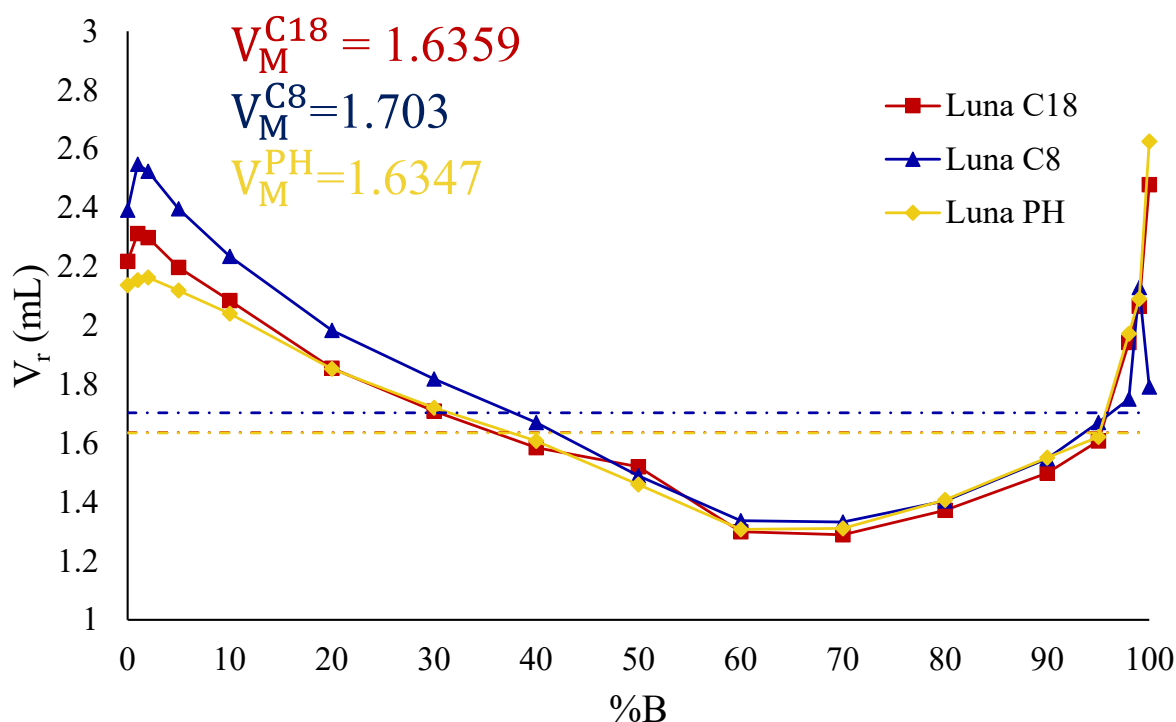


Figure 9.2 | Minor disturbance/perturbation method for 150-mm RPLC columns. Columns Luna C18, Luna C8, and Luna Phenyl-Hexyl (PH) for ACN/H₂O + 0.1% TFA.

9.2 Isotherm behaviour under adsorption conditions

The adsorption condition in the PM method represents an equilibrium state wherein the analyte undergoes both adsorption onto and desorption from the stationary phase at a rate conducive to elution, thereby yielding a specific value for retention time/volume. FA more comprehensively characterises true sorption measurements as discussed in the previous Chapter; nevertheless, the application of this PM technique to peptides and other biomolecules is limited by experimental factors such as low solubility, gelation, or economic considerations.

The PM data demonstrates similar adsorption isotherm behaviours for P1 onto C8 and C18, as depicted in Figure 9.3. Due to the absence of an aromatic ring in the molecule, the influence of π - π or cation- π interaction is not evident in the case of P1; instead, the adsorption is primarily

governed by hydrophobic interactions. For the same reason, low adsorption levels onto Phenyl-Hexyl are observed.

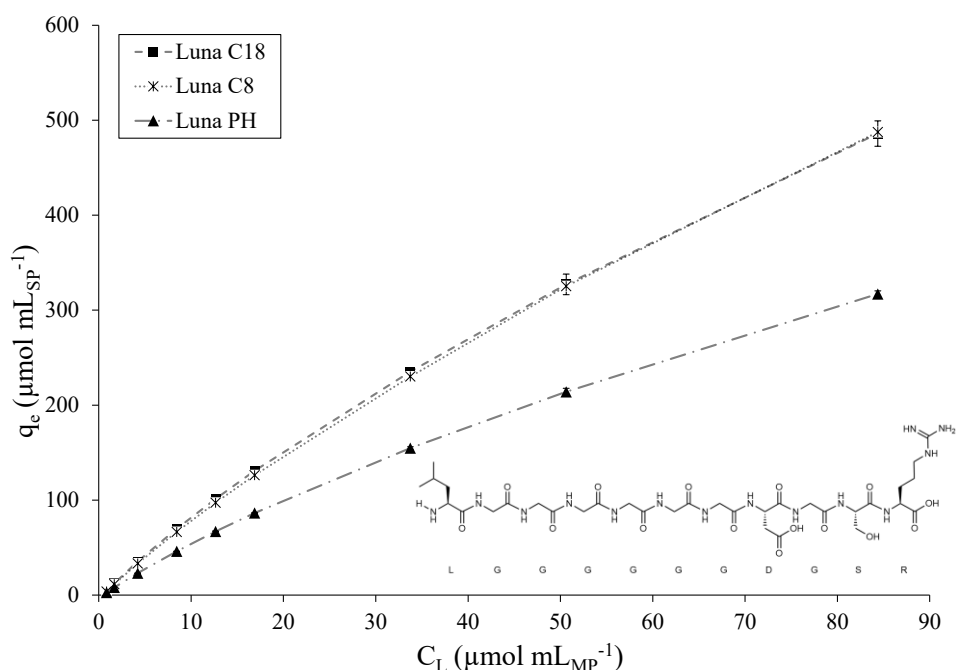


Figure 9.3 | Adsorption isotherm of peptide P1 for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 5% ACN/H₂O (v/v) + 0.1% TFA (v/v).

The synergy between π - π interaction and hydrophobic interaction between analytes and the stationary phase is anticipated as these interactions are known for their synergistic role in catalysing reactions[159], enhancing peptide stability and specificity[160], as well as facilitating the stabilisation of peptides within micelles. In contrast, for the adsorption of P2 to P6, where an aromatic ring is present at the serine side chain, a lower saturation capacity is evident in column Phenyl-Hexyl compared to columns C8 and C18. In fact, isotherms in Phenyl-Hexyl maintain a relatively uniform profile (similar values at the highest C_L) across all peptides despite the varying organic content (Figure 9.4 -Figure 9.8). This observation is contradictory, as the presence of both aromatic and hydrophobic groups on the surface would typically suggest a cumulative effect that should enhance stability and lead to an increase in adsorption capacity with more hydrophobic peptides.

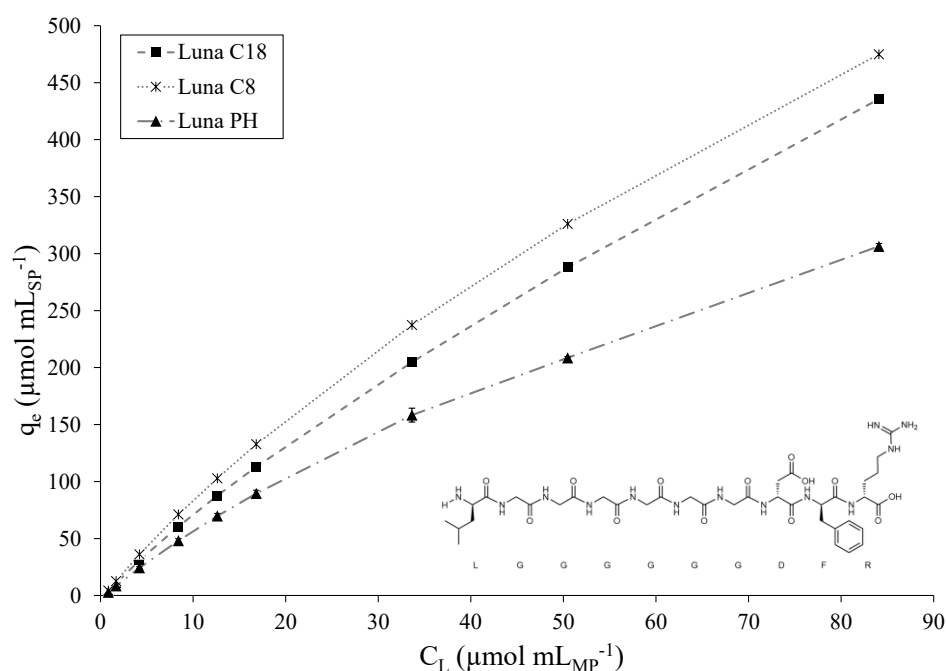


Figure 9.4 | Adsorption isotherm of peptide P2 for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 12.5% ACN/H₂O (v/v) + 0.1% TFA (v/v).

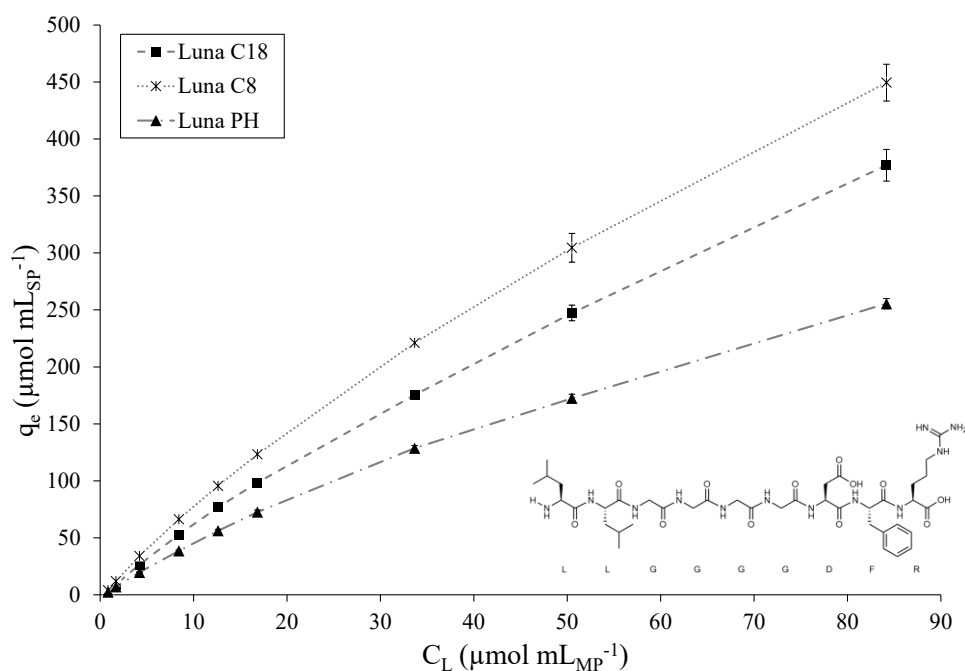


Figure 9.5 | Adsorption isotherm of peptide P3 for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 17.5% ACN/H₂O (v/v) + 0.1% TFA (v/v).

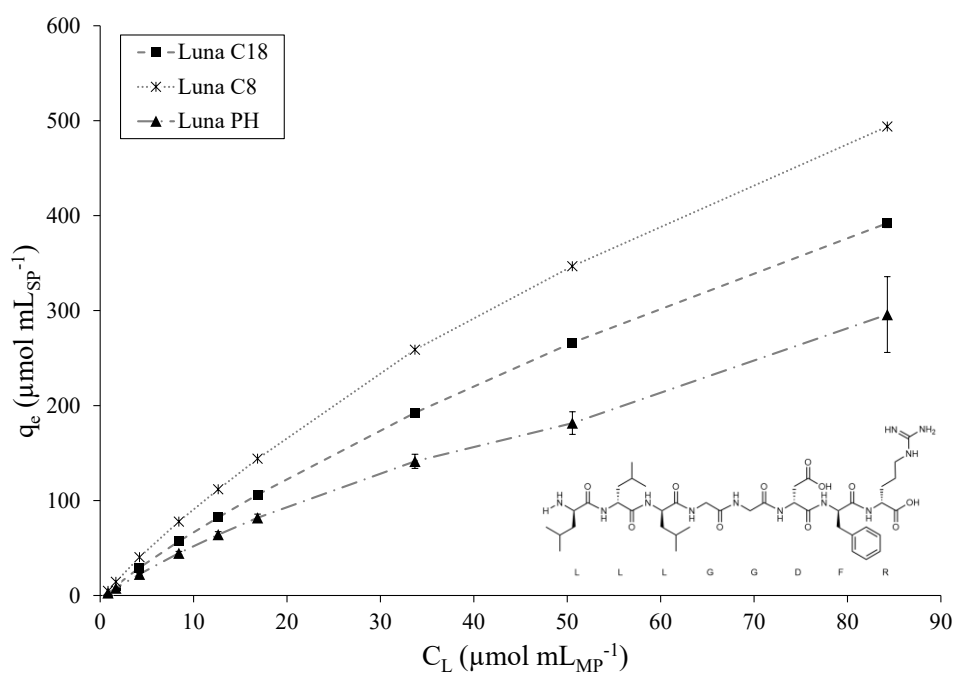


Figure 9.6 | Adsorption isotherm of peptide P4 for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 22.5% ACN/H₂O (v/v) + 0.1% TFA (v/v).

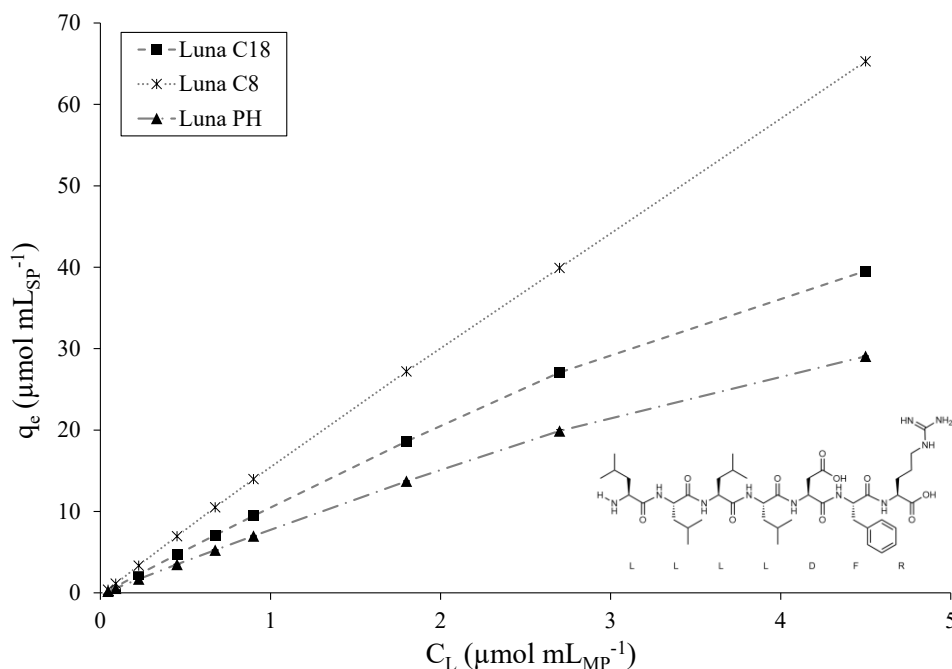


Figure 9.7 | Adsorption isotherm of peptide P5 for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 27.5% ACN/H₂O (v/v) + 0.1% TFA (v/v).

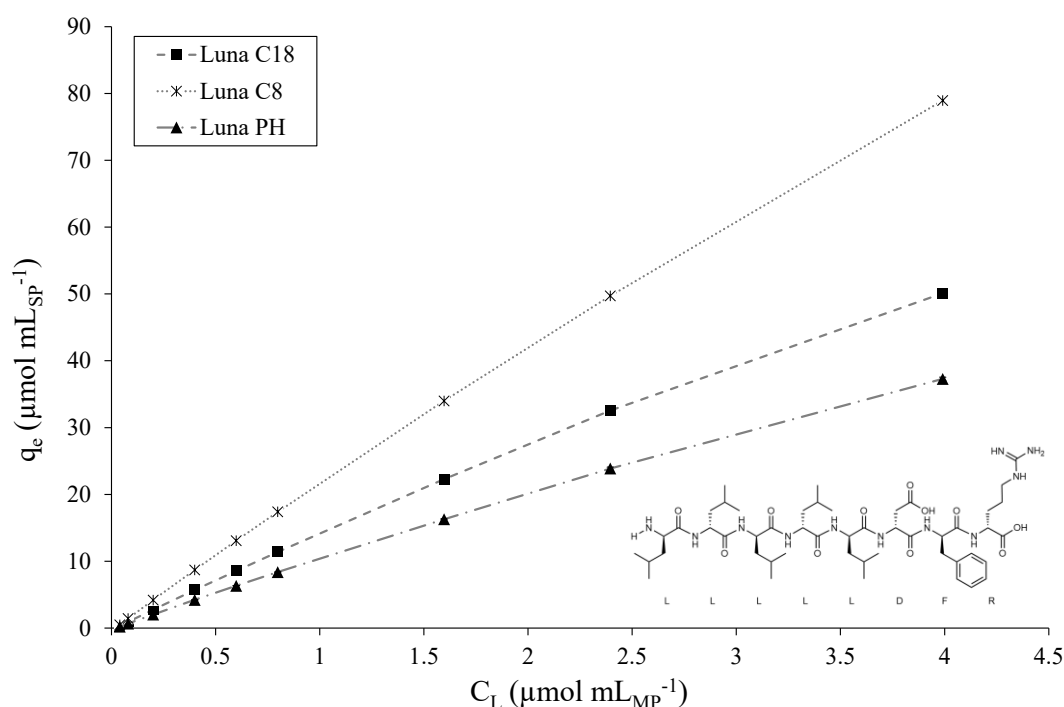


Figure 9.8 | Adsorption isotherm of peptide P6 for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 30% ACN/H₂O (v/v) + 0.1% TFA (v/v).

In the context of peptide adsorption, it can be hypothesised that the interfacial interactions are confined primarily to the low-energy sites situated at the ligand tails. Consequently, this confinement reduces access to high-energy sites adjacent to the ligands on the stationary phase. This restriction can be ascribed to the dimensions of the peptide and the spatial arrangement of ligands on the adsorbent surface.

In the case of Phenyl-Hexyl, a distinct difference in adsorption arises due to the presence of an aromatic group within the ligand tail as opposed to a hydrophobic group. This variance in tail composition results in a lower binding energy for π interactions compared to hydrophobic interactions. Consequently, the adsorption capacity of that column experiences a constant reduction. In other words, all peptides tested interacted only with the phenyl group but not with the alkyl motif (hexyl).

Another significant observation concerns the higher adsorption levels observed in column C8 in comparison to column C18 for P3. This difference becomes more pronounced when changing from P4 to P6. This observation can be rationalised by considering the inherent attributes of

both the peptides and the column resins. By considering the ligand coverage calculated in the previous Chapter according to the Berendsen-de Galan, the distance between each hydrocarbon surface ligand per metre squared of resin would be approximately 0.95, 0.78, and 0.79 nm for C18, Phenyl-Hexyl and C8, respectively. Suppose the coverage of ligands on the surface is low. In that case, the residual silanols and other end-capping species become more accessible to peptides, resulting in contributions from hydrophilic and ionic interactions during adsorption. So, a higher ligand coverage generates a relatively more homogenous surface, which enhances the hydrophobic interaction with peptides.

Considering the ligand coverage calculations presented in the previous section and assuming a homogeneous distribution, the approximate distances between each ligand per square meter of resin surface area are 0.95, 0.78, and 0.79 nm for C18, Phenyl-Hexyl, and C8, respectively. As the peptides used in this work form flexible random coils in water, it is possible to assume a linear structure as the most thermodynamically preferred structure during RPLC conditions. Given the minimum amino acid length of glycine, which is approximately 0.4 nm[161], P1-P6 or any other peptide larger than a triglycine would possess dimensions substantial enough to preclude interactions within the high-energy sites at the inter-ligand space. Therefore, the difference in ligand coverage and the adsorption layer thickness between columns C8 and C18 could explain higher adsorption for more hydrophobic peptides, as observed from P2 to P6.

In conclusion, when ligand coverage is relatively low, the residual silanols and other end-capping species tend to be more accessible to peptides. As a result, interactions originating from hydrophilic and ionic properties come into play during the adsorption process. Conversely, a higher ligand coverage leads to a more uniform and relatively homogenous surface, thereby enhancing the hydrophobic interaction with peptides. Consequently, the variation in coverage between columns C8 and C18 may facilitate the higher adsorption observed for more hydrophobic peptides.

An additional factor influencing the efficiency of adsorption involves the accessibility of pores. Considering P1 and P2, where both peptides exhibit comparable sizes (approximately 10 nm) yet differ in hydrophobicity, the similarity in the concentration of adsorbed peptide onto C8 and C18 suggests that the limiting factor of adsorption could be attributed to peptide size.

Given that all three resins possess an average pore size of 10 nm, a single molecule of P1 or P2 (being the largest of all) cannot effectively occupy the entire pore opening. Consequently, as the

theoretical occupancy reduces, the available surface area for adsorption does, too, thereby confirming the existence of a size exclusion effect. As the peptide size decreases, a more significant number of pores become accessible, thereby enhancing peptide adsorption. This rationale extends to the P3 to P6 range, where similar dynamics hold true.

There have been reports indicating that ligands positioned near the pore opening can constrict or even entirely block the pore opening[162]. Assuming a kinetic diameter of 0.07 nm per carbon[163], the presence of two C18 molecules at opposing positions within the pore aperture would effectively reduce the diameter to 7.5 nm. Shorter alkyl chains mitigate the potential for pore blockage on the surface, consequently leading to higher adsorption.

9.2.1 Adsorption behaviour and modelling

The chromatograms of all peptides exhibit the distinct characteristics of Type I or Langmuir-like adsorption behaviour, similar to those observed for caffeine (see Chapter 8). These include the observed receding apex of the chromatographic peak in terms of retention volume with increasing injection mass and extended tailing. Examples for P1 and P6 are shown in Figure 9.9.

All adsorption data of P1 to P6 were fitted with the Langmuir model (equation 6.38), and corresponding isotherm parameters are shown in Table 9.1. The Langmuir constant (K_L) represents the affinity of the adsorbate for the adsorbent expressed as a ratio between adsorption and desorption rates. Therefore, a higher value would be expected with increased peptide hydrophobicity, as seen in column C8. Parameter (q_m) represents the maximum adsorption capacity of the monolayer. This value is completely theoretical due to the peptide's constrained solubility at the interface. Reaching such levels can trigger peptide aggregation at the injection point and, in certain instances, even lead to gelation, especially during elution.

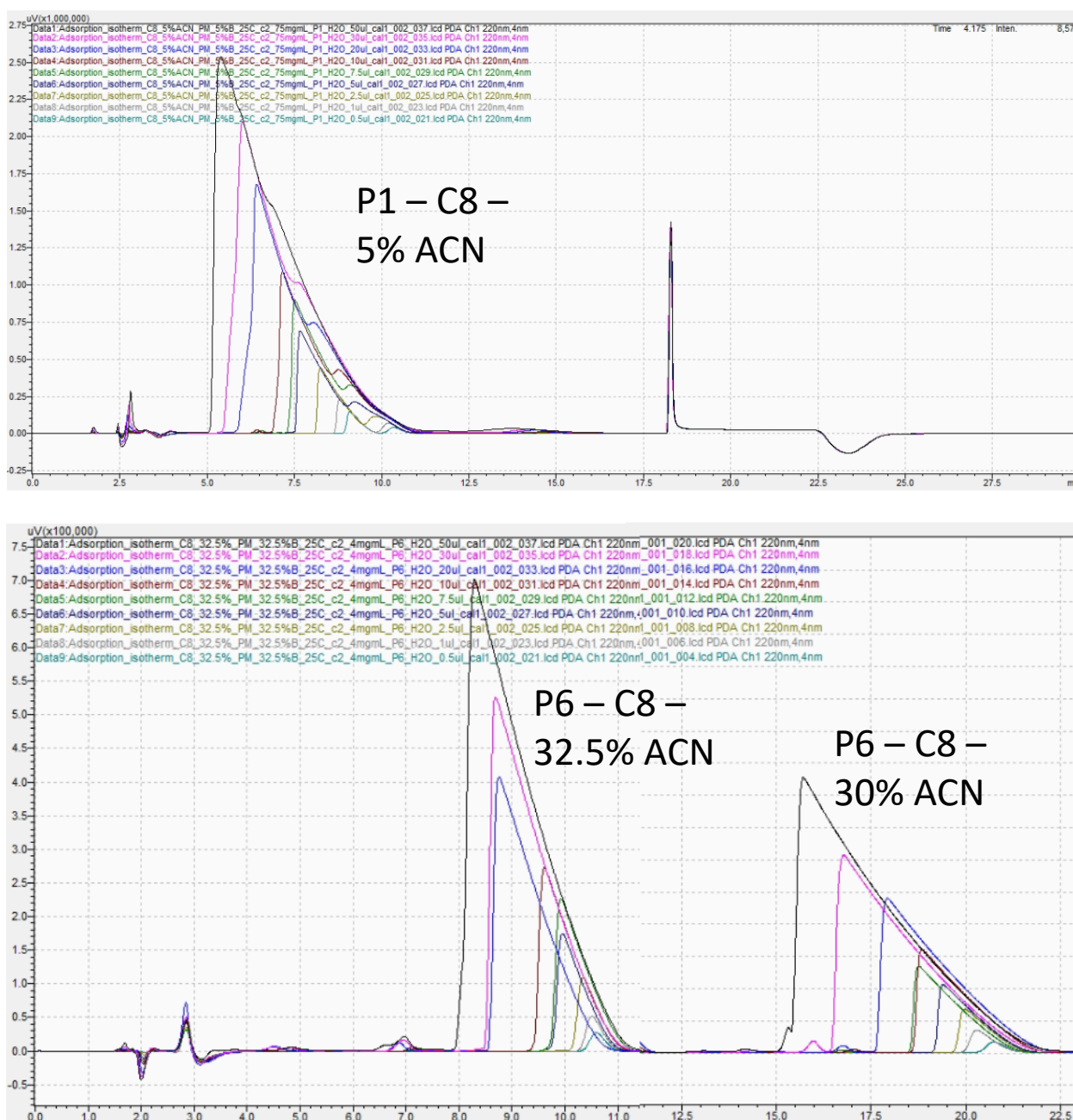


Figure 9.9 | Peak maxima (PM) method chromatograms. Peptides P1 (upper) and P6 (lower) were injected into the Luna C8 column.

Table 9.1 | Langmuir parameters for standard peptide isotherms.

Parameters		P1 – 5%B	P2 – 12.5%B	P3 – 17.5%B	P4 – 22.5%B	P5 – 27.5%B	P6 – 30%B
Luna C18	$q_m (\mu\text{mol.mL}_{\text{SF}}^{-1})$	1576.0	1607.9	1409.2	1222.1	170.7	312.45
	$K_L \times 10^3 (\text{mL}_{\text{MP}}.\mu\text{mol}^{-1})$	5.25	4.39	4.30	5.58	67.64	47.99
	RMSE	2.44	2.43	2.81	1.61	0.34	0.22
Luna C8	$q_m (\mu\text{mol.mL}_{\text{SF}}^{-1})$	1795.4	1371.0	1369.8	1254.8	1001.5	696.7
	$K_L \times 10^3 (\text{mL}_{\text{MP}}.\mu\text{mol}^{-1})$	4.40	6.27	5.76	7.67	15.48	32.04
	RMSE	1.90	2.10	2.36	1.88	0.18	0.19
Luna Phenyl-Hexyl	$q_m (\mu\text{mol.mL}_{\text{SF}}^{-1})$	994.1	798.2	721.8	1005.2	123.0	267.0
	$K_L \times 10^3 (\text{mL}_{\text{MP}}.\mu\text{mol}^{-1})$	5.51	7.31	6.43	4.81	69.4	40.75
	RMSE	1.70	2.62	1.96	6.20	0.24	0.11

The Langmuir model presupposes a homogeneous adsorption process, wherein the adsorbate is assumed to be adsorbed in a monolayer configuration. Despite the evident heterogeneity of the RPLC surface, peptides are primarily capable of interacting solely with the interface and the lig- and tails, as elucidated earlier. Consequently, additional electrostatic interactions with the acidic silanols or competitive hydrophobic interactions with the TMS end-capping are not anticipated. This, in turn, reduces the surface to an effective homogeneous adsorbate. In line with this notion, it must be possible to prove the formation of a monolayer by studying the structure of the interface before the presence of the analyte (adsorbate). This is possible by calculating the excess isotherm from the perturbation method presented in Figure 9.2. These results are shown in Figure 9.10.

From here, the theoretical thickness (τ) and the maximum volume (V_{ads}) occupied by the solvent-rich finite adsorbed interphase layer of ACN/H₂O/TFA can be estimated using equation 4.35. The V_{ads} values for C18, C8 and Phenyl-Hexyl are 0.25, 0.29, and 0.23 mL, respectively.

Similarly, their corresponding τ values are 6.23, 7.29 and 5.83 Å.g_{resin}. Note that although manufacturer information on the amount of resin in the column would be preferred, specific surface areas can still be used for comparative analysis.

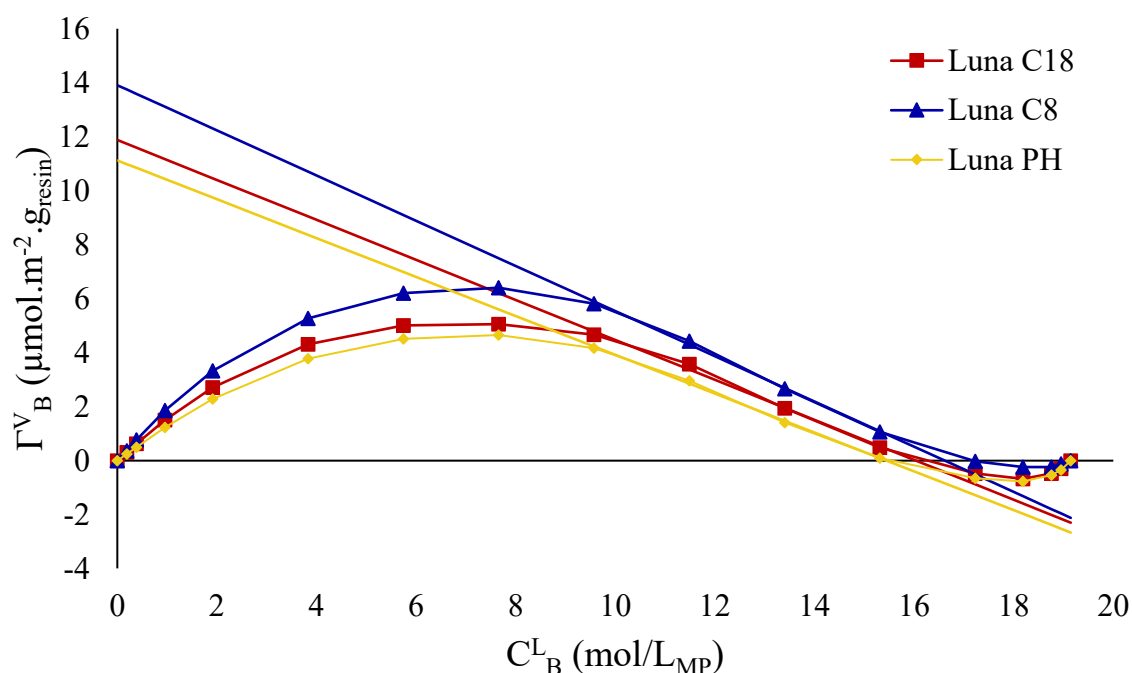


Figure 9.10 | Excess adsorption isotherm of ACN/H₂O + 0.1%TFA onto RPLC columns.

Figure 9.11 shows the total amount of eluent present in the finite interface between the bulk mobile phase and the adsorbent as a function of the bulk liquid composition according to Equation 4.35. It is my view that the assumption of a fixed interface/interphase layer is unrealistic. Instead, the interface thickness will expand at elevated organic content, mirroring the behaviour illustrated in the figure. Similarly, the ligands, starting in a constrained state when exposed to pure water, with a length of 7 Å[164], will undergo gradual relaxation and volume expansion as the organic content in the interface increases due to the changes in the mobile phase composition. Hence, the greatest extension of both the interface and the ligand layer will occur at 100% organic content, signifying that the interface/interphase layer will solely contain organic molecules, and C18 will assume its most extended configuration.

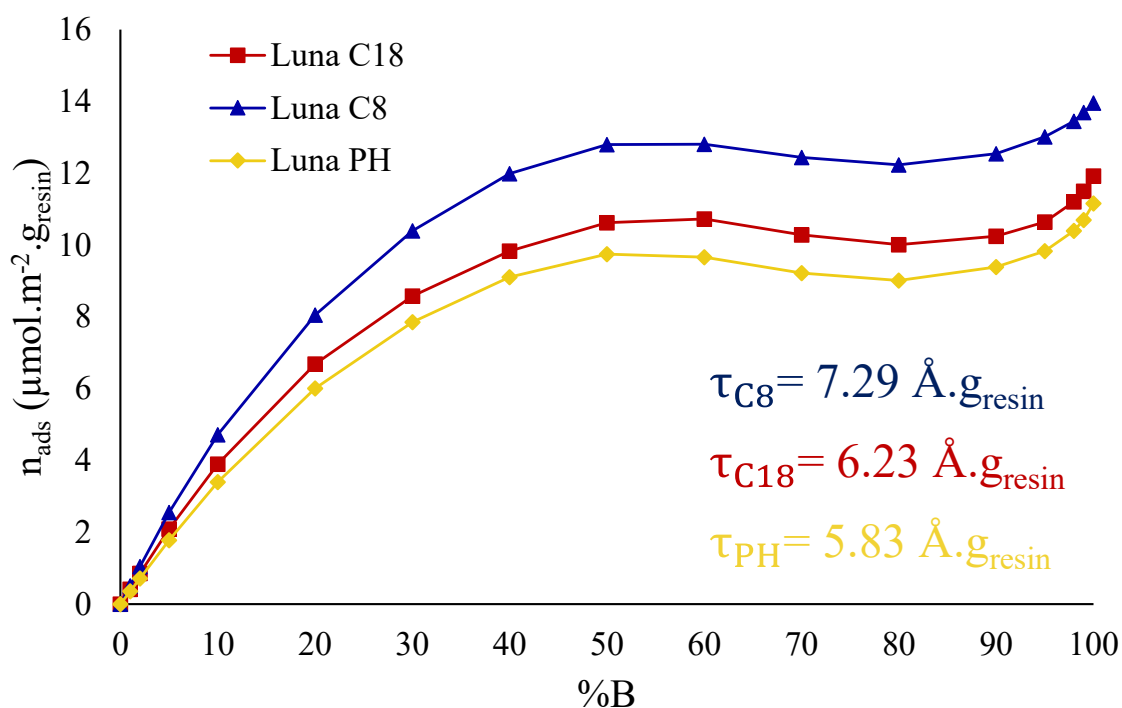


Figure 9.11 | Total mass of ACN/H₂O + 0.1%TFA eluent adsorbed onto RPLC with interfaces of theoretical finite thickness τ .

As shown in Section 4.6.1, during the initial lower organic content in the mobile phase, the interface/interphase composition rapidly increases due to high interactions between the bulk liquid and the surface of the resin until a plateau is reached. For all three columns tested in this work, that plateau at 50% organic composition in the mobile phase for all three columns. This observation is particularly interesting as it aligns with the optimal chromatographic elution range commonly observed for naturally occurring peptides during proteomics analyses[165, 166]. Exceptions to this pattern are rare and often pertain to specific cases, such as protected peptides, which may require higher organic content[167, 168]. This observation could suggest that, at 50% organic composition, the partitioning between the interphase and the mobile phase is underwhelmed by the susceptibility of peptides to solubilise into the bulk liquid.

This rapid change in interface structure and composition as it saturates with organic molecules reflects the strong dependence of isocratic runs on a fixed mobile phase composition selected and, hence, interface structure throughout the analysis. Any slight variation in the organic solvent content of the mobile phase can have a pronounced impact on the elution profiles of analytes. Since isocratic conditions maintain constant solvent strength, even a minor deviation in

the actual organic content from the intended composition can result in unexpected changes in the interaction between the analytes and the stationary phase.

Such variations would, in turn, lead to fluctuations in chromatographic retention times and elution patterns, as well as narrowing the elution profile of peptides, introducing variability and potentially compromising the accuracy and reproducibility of the chromatographic results. Therefore, meticulous attention to accurately determining and maintaining the mobile phase composition is essential to ensure reliable and consistent isocratic separations. Indeed, this concern is typically less prominent during gradient runs due to the inherent advantages of this chromatographic approach. During gradient, any minor errors in determining the initial organic content and the lack of a fully formed interface are often mitigated by the rapid change in solvent strength, which effectively focuses on the eluted peptides.

As the gradient elution process progresses, the organic content increases or decreases in the mobile phase, prompting a corresponding adjustment in the interaction between the peptides and the stationary phase. This adaptability helps compensate for inaccuracies in the initial mobile phase composition determination, minimising the impact of any retention time variability that might have arisen in the earlier stages of the gradient.

Overall, gradient elution enhances the robustness of the chromatographic method by providing a more forgiving environment for handling variations in solvent composition, which in turn leads to improved reproducibility and reliable separation outcomes. However, the potential for enhancing the column capacity for a specific analyte by initiating the gradient with an optimised interfacial structure remains in question.

9.2.2 Proposed peptide adsorption behaviour

All observations discussed so far are condensed and depicted into Figure 9.12. In this figure, the ligands are shown as **brown moving lines**, the interface layer surrounding the ligand on top of the packing material/adsorbent's surface is **yellow**, the bulk mobile phase is **blue/teal**, and the peptides are **grey lines**. The **red colour** in the peptide representation indicates the number of hydrophobic residues in the sequence, i.e. the most hydrophobic peptide (P6) is almost entirely red. In contrast, peptide P1 only has a red hydrophobic 'head'.

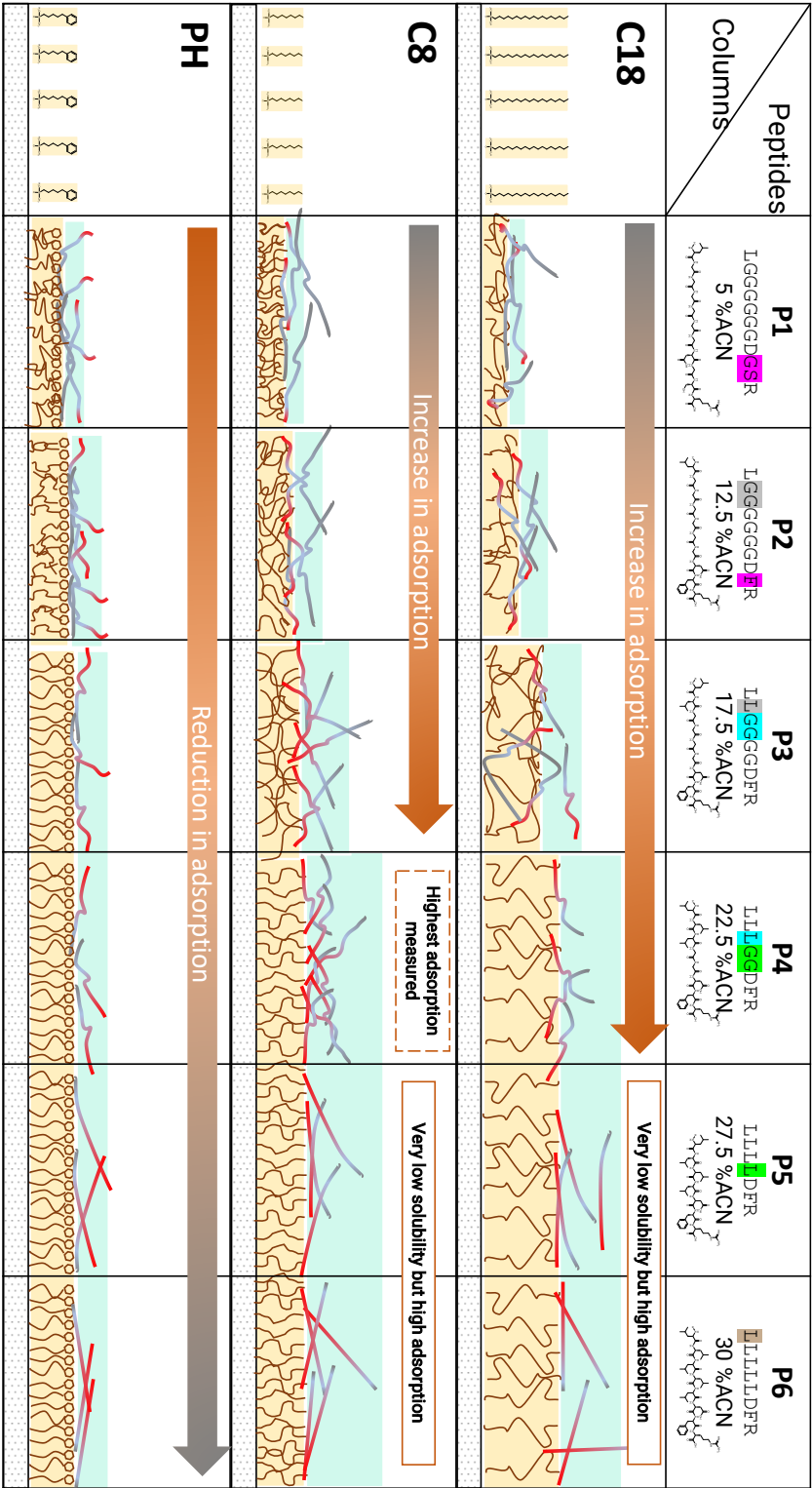


Figure 9.12 | Summary of adsorption behaviour of standard peptides on RPLC columns. The adsorption isotherm data of P1-6 determined by PM under adsorption conditions for three columns and corresponding peptide sequences. See the text for a detailed explanation.

As the organic content increases, the thickness of the interface layer increases as it fills with ACN/H₂O molecules. This dense interface is a rigid (stagnant) liquid interphase layer where adsorption occurs, and its volume and thickness determine the number of analyte/adsorbate/peptide molecules adsorbed. As injection occurs, the ligands extend slightly but never to their full extension. Peptides are instantly absorbed into the interphase layer if they are within their adsorption region, meaning that the overall adsorption energy is higher than the cumulative desorption energy and Gibbs energy that keeps the peptide in solution at the bulk (mobile phase) liquid. The desorption region will be covered in the next section of this Chapter.

The ligand packing density is also represented in the figure. Since the peptides are so large, the distance between each ligand (C8, C18 or Phenyl-Hexyl) is relatively small, and thus, the peptides cannot access the area between them. Instead, they only interact with the head of the ligands.

In the case of C8 and C18, the only possible interactions between the peptides and ligands are through hydrophobic interactions. Therefore, as the hydrophobicity of the peptide increases, so does the maximum adsorption capacity of the column. The Phenyl-Hexyl column exhibits a totally different behaviour. Since the peptides only have access to the phenyl groups at the head of the ligand, the only possible interactions are those of π -polar. Interestingly, since the adsorption capacity is relatively constant, it can be inferred that interactions occur with the arginine at the c-terminal of the peptide sequences.

9.3 Isotherm behaviour under elution conditions

Unlike adsorption conditions, the peptide elutes conditions represent a state where the analyte resides within the bulk liquid along the column. Subsequent isotherm measurements were conducted at an additional 2.5%B beyond the adsorption conditions in and at the final %B indicated for each peptide in Figure 9.1.

For all peptides, the adsorption isotherms gradually transition from Type I to linear and ultimately to no adsorption (complete elution) as the solvent strength increases, as shown in Figure 9.13 for P2 and P6 eluted from a Luna C18 column. The isotherms can be found in Appendix IV.

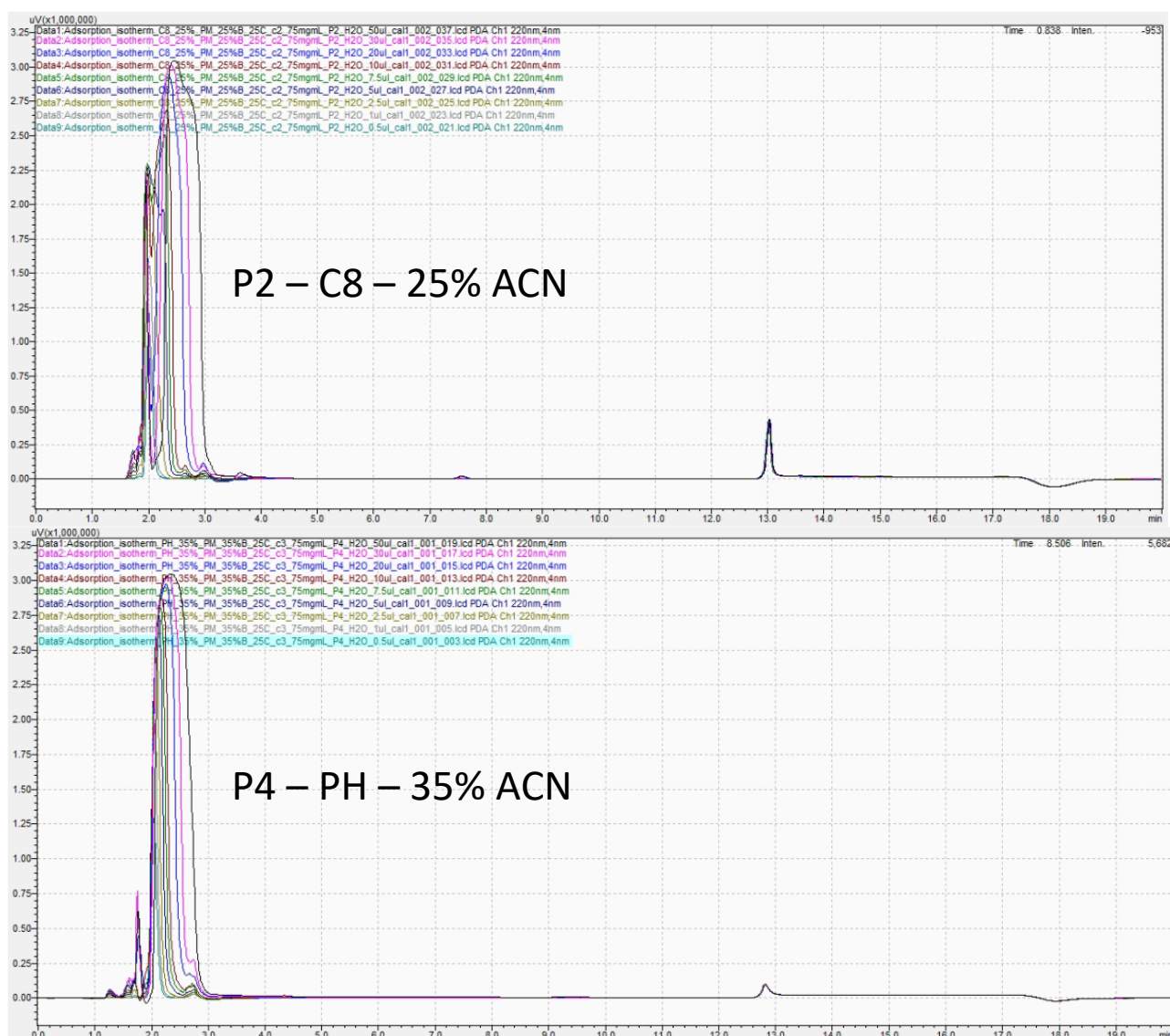


Figure 9.B | Chromatograms of standard peptides at eluting conditions.

In linear chromatography, elution conditions yield a peak at a consistent retention volume with increasing height and width as mass injected increases. However, the results under nonlinear conditions are not definitively elucidated. For most peptides studied in this work, peak splitting and retention of species can still be observed at high-mass injections, as shown in Figure 9.13. These observations are likely attributed to mass transfer phenomena, specifically expressed through surface diffusivity within the interface. As peptides are introduced into the column, they continue to experience diffusion within the interphase, leading to peak broadening and slight retention.

9.4 Conclusion

Peptide retention mechanisms and adsorption behaviour in RPLC conditions result from an interplay between peptide hydrophobicity, stationary phase characteristics, interface/interphase composition, and solvent conditions.

Our results revealed that the elution range of peptides tends to narrow as their hydrophobicity increases. This phenomenon emphasises the crucial role of controlling the organic content of the mobile phase to facilitate effective adsorption and desorption for peptides of varying analyte/solute hydrophobicities. Intriguingly, the accessible partitioning volume for peptides within the column decreased with peptide size and was influenced by ligand coverage. Ligand density significantly impacted adsorption behaviour, with a higher coverage leading to a more apparent homogenous surface and enhanced hydrophobic interactions during the adsorption process.

Furthermore, our investigation into adsorption isotherms revealed a gradual transition from Type I to linear and eventually to no adsorption as solvent strength increased. This behaviour was consistent across various peptides, highlighting the intricacies of the adsorption process. Despite the surface heterogeneity of the stationary phase, applying the Langmuir model revealed that peptide adsorption primarily occurs through interaction with the interfacial phase with the associated ligand tails, resulting in an effective homogeneous adsorption layer or phase.

By analysing the interface structure/composition prior to analyte introduction, I have demonstrated that the interface's thickness and the extension of ligand layers increase with elevated organic content. Our results provided valuable insights into the formation of interfacial layers/phases and the potential expansion of these interfacial layers/phases, indicating that the organic content of the mobile phase plays a pivotal role in the adsorption phenomena. By strategically tailoring the interfacial conditions at the outset of the chromatographic run, creating an environment that promotes favourable interactions between the analyte and the stationary phase might be possible. This approach could improve peak shape, increase retention, and enhance separation performance.

However, this approach requires a nuanced understanding of the interplay between the analyte's properties, the stationary phase characteristics, and the composition of the mobile phase at the beginning of the run. Subtle adjustments to the interfacial structure could influence how the

analyte initially interacts with the stationary phase, setting the stage for the rest of the separation process.

Chapter 10

SMART algorithm

This Chapter presents the SMART (Systematic, Multi-parametric, Adaptive, Responsive, Tailored) algorithm for optimising multi-isocratic elution methods in reversed-phase liquid chromatography (RPLC) peptide purification. The algorithm was tested on peptide standard samples as well as glucagon, demonstrating significant improvements in separation efficiency, selectivity, and purity compared to traditional gradient elution methods. The SMART algorithm accurately predicted and optimised elution conditions, achieving a final purity of 93.6% for glucagon from an initial crude purity of 27.5%. A traditional gradient method could only achieve 83.8% purity. This new method also allowed for a two-fold increase in column load capacity, enhancing process efficiency. The scalability of this approach was confirmed with consistent separation performance from analytical to semi-preparative scales. The multi-isocratic approach provided better resolution of closely eluting impurities and reduced concentration of aggregation and degradation products, making it advantageous for industrial applications. The study underscores the SMART algorithm's potential to revolutionise chromatographic separations, offering a robust, efficient, and scalable alternative to traditional methods. Future research should expand its application to more complex samples, refine its adaptability, and ensure long-term stability, establishing the SMART algorithm as a standard in industrial and academic chromatographic practice.

10.1 Retention theory

In chromatography, the retention of an analyte is represented by a band migrating through the column; therefore, retention volume V_R can be expressed from equation 4.12 to simply as:

$$V_R = V_M(1 + k) \quad (10.45)$$

where V_M is the column mobile-phase volume and k is the peak (band) retention factor. The fractional migration of the peak travels through at a speed determined in units per column-volume V_M of mobile phase and not per simply column volume (V_C). The latter is the volume of an empty column (no resin). It is important to note that V_M is marginally different to void volume (V_0) as it considers both interstitial and pore volumes (see Chapter 4.1.1). In some cases, depending on the estimation method, V_0 and V_M are the same.

The elution of the chromatographic peak through the column and the physicochemical factors that control the process are different for each column, solute and mobile phase system as multiple interactions coexist in the adsorption-partition phenomena of liquid chromatography. By 'system', I refer to the microenvironment constituted by the analytes, mobile phase, and stationary phase at the moving infinitesimal space within the column length. This infinitesimal space can be measured by volume (dV) or distance ($dx = dV/(V_R - V_M) = dV/V_M k$). The sum of all these fractional migrations with $\sum dx = 1$ when $\sum dV = V_R - V_M = V'_R$ can be expressed with the following equations:

$$\int_0^{V'_R} \frac{1}{V_M k_i} dV = \int_0^{t'_R} \frac{1}{t_M k_i} dt = 1 \quad (10.46)$$

10.1.1 Isocratic elution

In absolute RPLC, the retention of an analyte is a strong function of the mobile phase composition or solvent strength. The linear solvent strength (LSS) model conveys that ideality in the form:

$$\log k = \log k_w - S\phi_B \quad (10.47)$$

where k is a constant which is a function of the volume fraction of organic solvent ϕ_B (expressed in decimal form or 0.01 per cent B in aqueous solvent A), and the parameter S is a constant for a single analyte and fixed experimental conditions, it is also often referred to as solvent strength value and k_w is the extrapolated theoretical value of k at 100% aqueous conditions ($\phi_B = 0$).

Despite this wide use of the empirical LSS model, it is important to note that it ignores other effects, such as the chromatographic band profile dynamics and physicochemical interactions other than hydrophobic.

10.1.2 Gradient elution

When a linear gradient is applied to RPLC, the composition of the mobile in terms of solvent B leaving the column (ϕ_B) changes over time after the initiation of the gradient (t) and, assuming ideal mixing at the pump outlet. It can be described by:

$$\phi_B = \phi_{B,0} + \left(\frac{\Delta\phi_{B,g}}{t_g} \right) t = \phi_{B,0} + \left(\frac{\Delta\phi_{B,g}}{V_g} \right) V \quad (10.48)$$

where $\phi_{B,0}$ is the value of ϕ_B at the start of the gradient, $\Delta\phi_{B,g}$ is the change of ϕ_B during the gradient and t_g is the gradient time. The equivalent expression in terms of volume (V) is only valid when the volumetric flow rate is kept constant.

Because the values of k vary as the band migrates through the column during the gradient, a median or equivalent value of k is used in gradient elution (k^*). This median value is defined as the value of k when the analyte is at half the length of the column. Similarly, ϕ_B^* is the corresponding value of ϕ_B at that position. The LSS model can be applied to these values using different gradient slopes to estimate values of S and $\log k_w$:

$$k^* = \frac{V_g}{1.15V_m\Delta\phi_{B,g}S} \quad (10.49)$$

$$\phi_B^* = \phi_{B,0} + \frac{\Delta\phi_{B,g}}{V_g} \left(V_R - V_M - V_D - 0.3 \frac{t_M}{b} \right) \quad (10.50)$$

$$b = \frac{V_m\Delta\phi_{B,g}S}{V_g} = \frac{t_M\Delta\phi_{B,g}S}{t_g} \quad (10.51)$$

However, elution occurs once the peak completely exists from the column; the equations above and, in particular Eq. 6 can be transformed into the following:

$$\phi_{B,e} = \Delta\phi_{B,0} + \frac{\Delta\phi_{B,g}}{V_g}(V_R - V_M - V_D - V_{EC}) \quad (10.52)$$

where the elution is obtained from the retention volume measured at the apex of the peak. This formula can also be inferred logically as it is a simple translation of retention volume into mobile phase composition by considering the slope of the gradient and adjustments to additional volumes in the system other than the column.

As the values of mobile phase composition would vary depending on the efficiency of the solvent mobile phase mixing in the LC system, an equivalent isocratic composition (ϕ_{iso}) can be estimated from $\phi_{B,e}$ values using the following formula:

$$\phi_{iso} = \phi_{B,e} + \frac{\log\left(\frac{k^*}{6.4}\right)}{S} \quad (10.53)$$

10.2 Standard peptides – proof of concept

The efficiency of the SMART algorithm was tested using the model peptides P2, P3, and P4. These peptides were selected based on their subtle sequence differences, which produce measurable differential hydrophobic separation for both isocratic (Chapter 9) and gradient elution chromatographic conditions.

The elution order of the peptides, as anticipated, aligns with the computational log P values, a hydrophobicity metric, as depicted in Figure 10.1. Since adsorption occurs instantly and within the first millimetres of the column, according to observations in previous Chapters, a 5-minute delay at $\phi_B = 0.05$ was incorporated to ensure complete adsorption at the column's head.

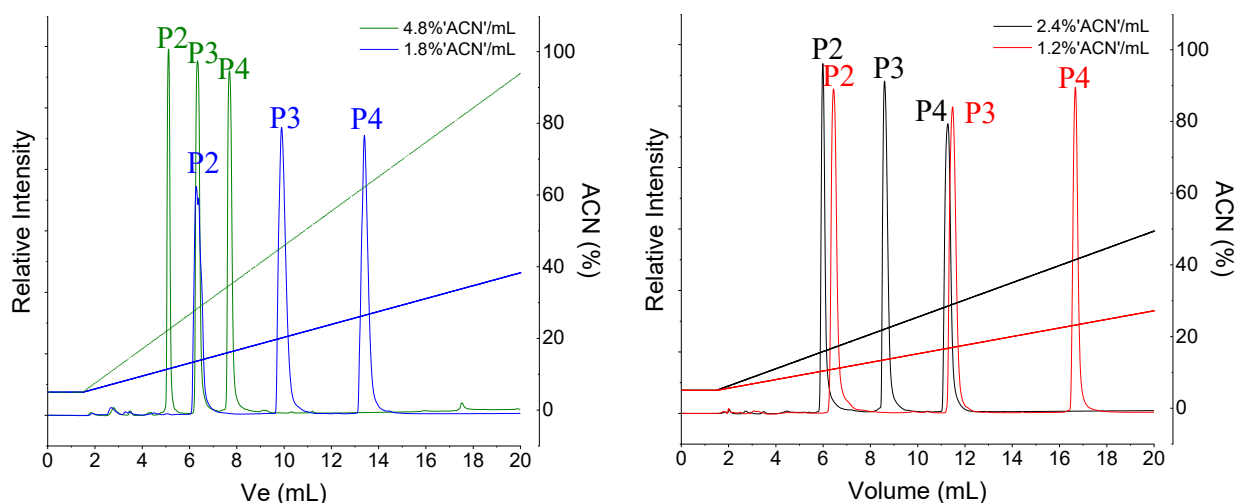


Figure 10.1 | Gradient elution order of P2, P3, and P4.

The gradient scouting step in the algorithm is designed to quickly ascertain the elution order of the tested peptide mixture for each column tested and identify the peaks of interest. A sufficiently slow gradient (0.5%B/mL or 1.25%B/CV) facilitates the understanding of the separation capabilities and constraints of the tested column as it resembles the conditions of an isocratic run. Generally, isocratic elution provides the maximum chromatographic separation/performance for an analyte-stationary-mobile phase tri-system. Once this 'maximum' threshold has been reached, separation cannot be improved unless one of the parts of the tri-system is changed.

To develop the SMART algorithm, the four chromatographic gradients shown in Figure 10.1 were used as the gradient scouting step. From these chromatograms, equation 9.46 was used to calculate an estimated ϕ_B range at which each standard peptide would elute. Isocratic runs were performed using those ϕ_B ranges for each peptide: 0.1 - 0.175 for P2, 0.175 - 0.275 for P3, and 0.275 - 0.375 for P4. Results are reported in Figure 10.2 (left). By applying the LSS model to the isocratic chromatographic run via a linear regression (Figure 10.2, right), then the values for S and $\log k_w$ (Table 10.1, Right) were calculated.

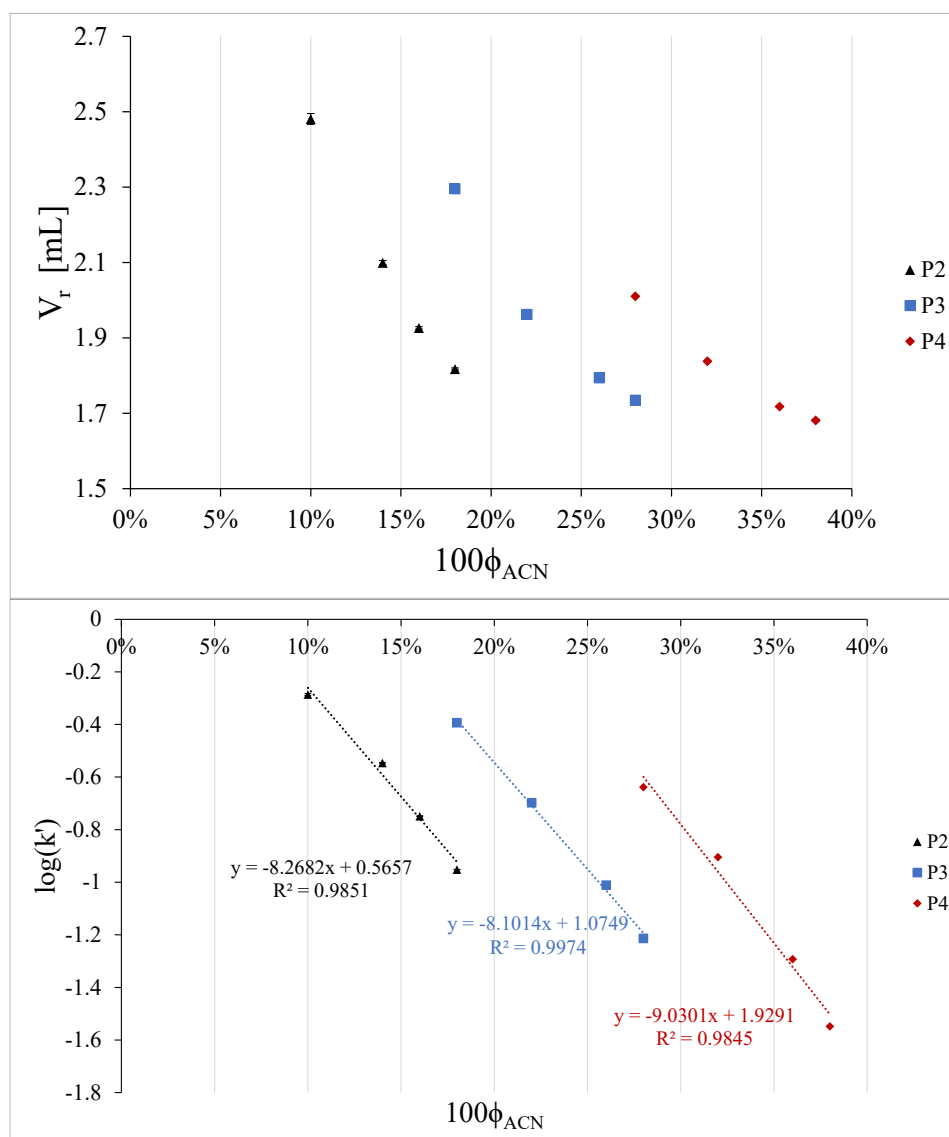


Figure 10.2 | Isocratic elution for P2, P3, and P4. Graphs of retention volume (V_R) and $\log(k)$ fitted to the LSS model. Mobile phase: A: water B: acetonitrile (ACN).

Alternatively, these values can also be estimated directly from the gradient data. Using equations 9.47 and 9.48, the values of ϕ_B^* and k^* are calculated as a first step, allowing for the estimation of new $\log k_w$ and S for each peptide across the four scouting gradient runs. The per cent errors between the isocratic and gradient methods were found to be $28.7 \pm 11.2\%$ for S and $82 \pm 47\%$ for $\log k_w$. From these values, the gradient data was used to estimate isocratic elution at ϕ_{iso} using equation 9.51 (Table 10.1, Left) of $7.24 \pm 0.06\%B$ for P2, $13.19 \pm 0.02\%B$ for P3, and $19.71 \pm 0.17\%B$ for P4.

It was found that calculating the ϕ_b range from ϕ_{iso} , provides more consistent results regardless of the gradient steepness and agrees with $\phi_{b,e}$ (equation 9.50) with slopes of 1.5-1.7%B/mL. Any shallower gradient data sets provided a reduced estimation of the elution behaviour, which is more similar to that of isocratic experiments.

Table 10.1 | Gradient analysis for P1, P2, and P3. (Left) Calculations of initial parameter for isocratic elution using gradient results from Figure 10.1 and equations 9.45-9.50. **(Right)** Calculations of the Linear Strength Model (equations 3) based on isocratic elution.

Gradient elution								Isocratic elution	
Peptide	Gradient slope ($100\phi_{B,g}/V_{g'}$ %ACN/mL)	Gradient retention factor (k')	$100\phi_B^*$ (%ACN)	S	$\log k_w$	$100\phi_{B,e}$ (%ACN)	$100\phi_{iso}$ (%ACN)	S	$\log k_w$
P2	2.40	2.01	9.20 ± 0.01	11.01	1.30	11.93 ± 0.04	7.365	8.27	0.57
	1.20	4.03	6.29 ± 0.01			9.01 ± 0.01	7.185		
	4.80	1.01	11.79 ± 0.01			14.48 ± 0.19	7.177		
P3	2.40	1.99	15.49 ± 0.01	11.14	1.99	18.18 ± 0.05	13.63	8.10	1.08
	1.20	3.98	12.37 ± 0.01			15.07 ± 0.02	13.21		
	4.80	0.99	17.73 ± 0.01			20.42 ± 0.13	13.16		
P4	2.40	2.11	9.20 ± 0.01	10.51	2.58	24.64 ± 0.03	20.05	9.03	1.93
	1.20	4.22	6.29 ± 0.01			21.30 ± 0.02	19.57		
	4.80	1.05	11.79 ± 0.01			26.96 ± 0.09	19.51		

Interestingly, S-values for biomolecules are typically higher than those for small molecule species[169]. While this result may not surprise separation scientists specialising in biomolecules, it contradicts much of the existing literature on small molecules, where recommended S-values usually range between 4-5[170]. This discrepancy arises from the limited data sets collected from biomolecule separations compared to the larger number of small molecule data sets available.

A large S-value might be considered a disadvantage as the solvent composition significantly affects the analyte. However, this result also implies that under more aqueous conditions, the analyte will be robustly adsorbed onto the stationary phase (higher $\log k_w$). This strong adsorption of biomolecules, coupled with a low longitudinal axial dispersion[171], creates an environment

where the analyte can be robustly adsorbed and desorbed with a degree of control. This is commonly referred to as the *on-and-off effect*.

The multi-isocratic method was developed in this Chapter is based on the above-obtained LSS parameters. This separation aimed to elute the peptides from the head of the column with minimal separation (close to V_m values) to reduce both separation time and solvent consumption. The method was thus developed based on the values where two peptides are eluted together, as shown in Figure 10.3 (upper). An initial delay at a lower concentration of B was incorporated to drive adsorption at the column inlet. If the on-and-off effect holds true, then the elution volume of each peptide remains the same as in the isocratic runs after changing the concentration. In essence, provided the adsorption is sufficiently robust such that the retention factor (k) is virtually zero, the peptide will remain stationary on the chromatographic media until mobile phase conditions are changed to allow migration. If this change in conditions is sudden, the peptide will migrate at a rate equivalent to that if injected at the exact moment of change. This nuanced understanding of peptide behaviour under varying elution conditions provides valuable insights for developing efficient separation methods.

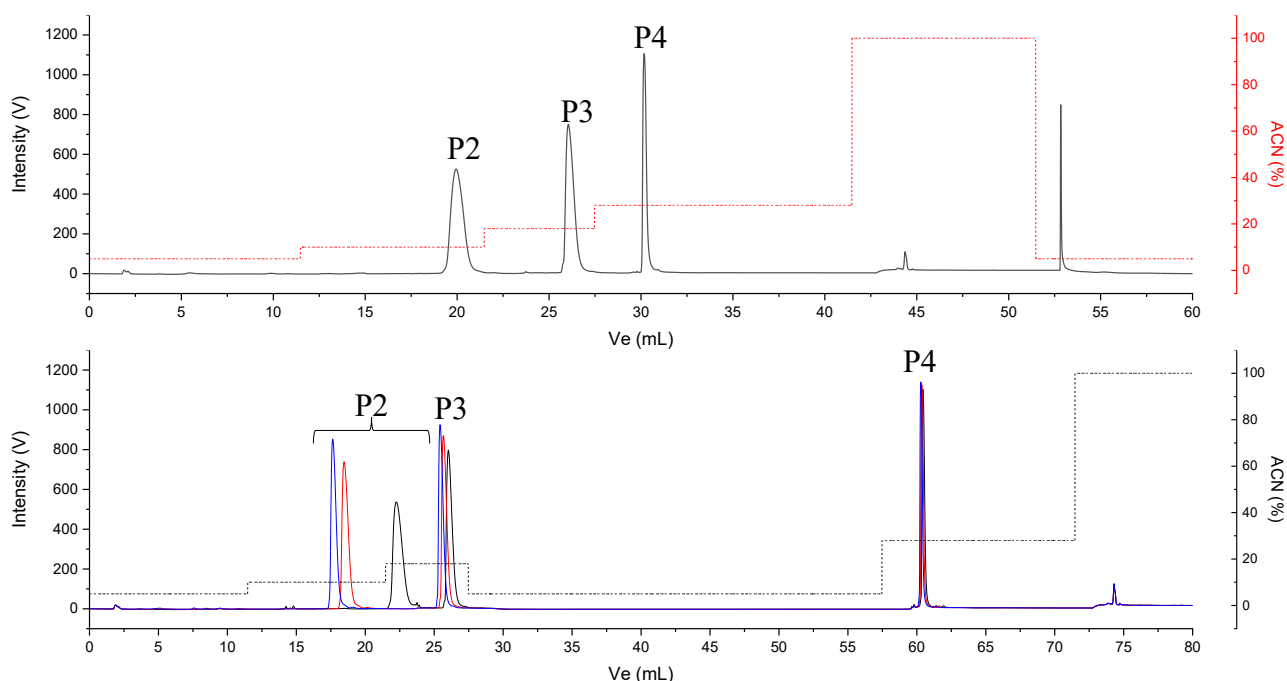


Figure 10.3 | Multi-isocratic run for P2, P3, and P4. Final multi-isocratic design to increase the chromatographic separation between peptides while keeping rapid elution (close to V_M). Mobile phase: A: water B: acetonitrile (ACN). Elution profile: 0-11.5 mL at 5%ACN; 11.5-20 at 10%ACN; 20-26mL at 18%; 26-40 m at 28%ACN (upper). A hold at 5% for 30 mL was added to the lower profile.

To further investigate the on-and-off effect in peptide adsorption/desorption phenomena, an extra hold step was added to the multi-isocratic profile between P3 and P4. The mobile phase was abruptly decreased to the initial delay composition. After a hold period of 30 mL, P4 was eluted at 28%ACN. Interestingly, no retention volume or peak width differences were found (Figure 10.3, lower). The same results were found after a 24-hour solute hold-up time, reinforcing the on-and-off effect theory (data not shown).

10.3 Glucagon – analytical scale RPLC separation

The native peptide glucagon plays an important physiological role in glucose-dependent insulin secretion when it binds to its receptor. Consequently, its amino acid sequence has been optimised for a class of synthetic peptides known as glucagon receptor agonists (GRAs), which are used in the treatment of type 2 diabetes[172-174]. Despite the current limited commercial advantage of native glucagon, it serves as a representative model peptide for GRAs when testing new purification strategies.

The gradient scouting experiments, shown in Figure 10.4, reveal that the crude glucagon sample contains a middle peak corresponding to the primary compound of interest. The separation selectivity of this peak relative to the two adjacent peaks improves as the gradient slope decreases. Moreover, the resolution of the C8 column converges faster than the other two columns, indicating faster retention volumes. The combined effects of a reduced base shift caused by rapid solvent mixing and slower mass transfer during shallow gradients result in a reduction of peak height and an increase in peak width.

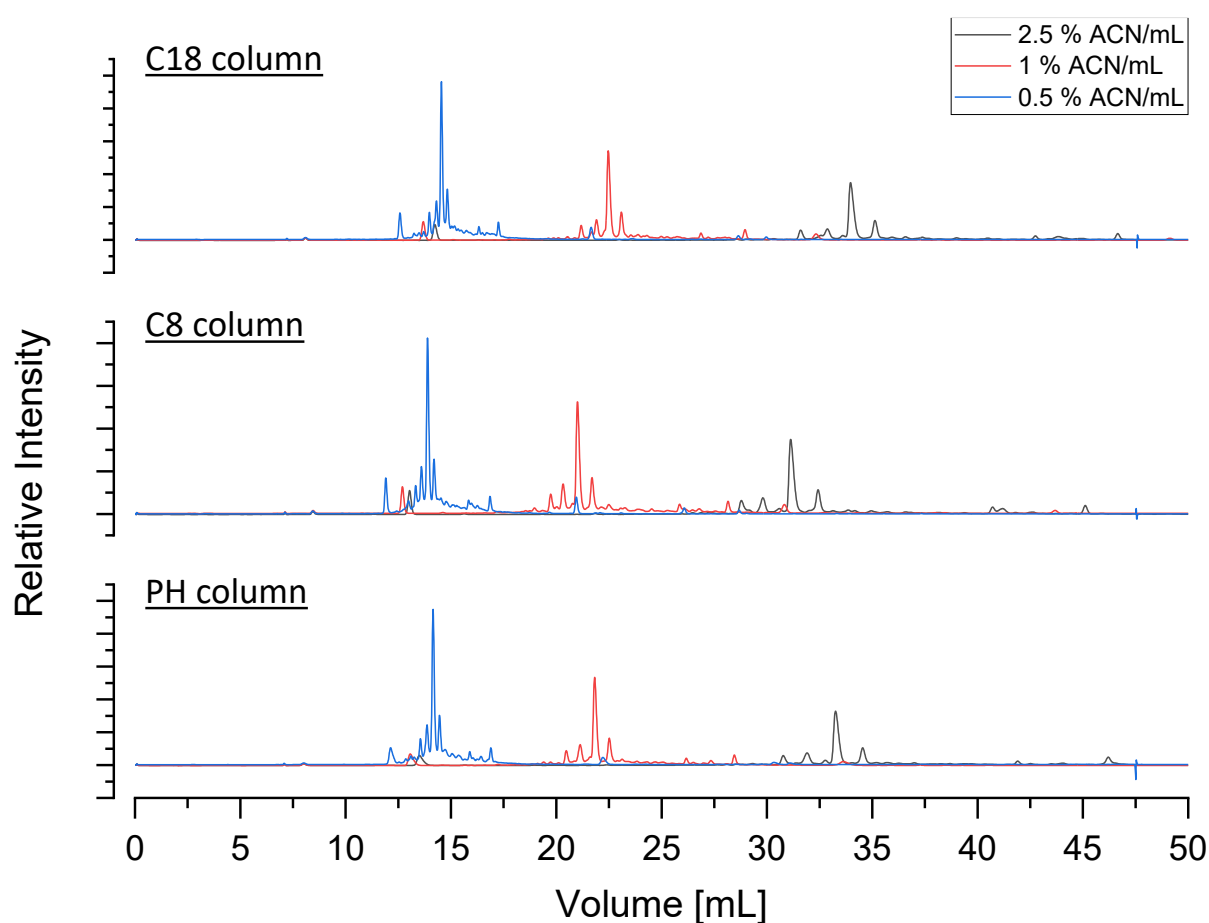


Figure 10.4 | Gradient elution of crude glucagon. m_{inj} = 0.1 mg. Flow rate: 1 mL/min. A: Water + 0.1%TFA. B: Acetonitrile+ 0.1%TFA. PH stands for Phenyl-Hexyl.

Following the same procedure as with the peptide standards, I have estimated the values of ϕ_{iso} for separating the main peak in each column using equations 9.46-50. Isocratic runs in the range of mobile phase composition of 0.29 - 0.305 for ϕ_B were conducted to assess the effect of the solvent strength on the separation performance of glucagon. By plotting $\log k$ versus ϕ_B for the main peak of interest, a comparative graph for column selection can be plotted (Figure 10.5). Notably, the S-values for glucagon are higher than 10, providing insight into the scale of the effect of the mobile phase composition on the RPLC separation despite being specific for the systems tested in this work.

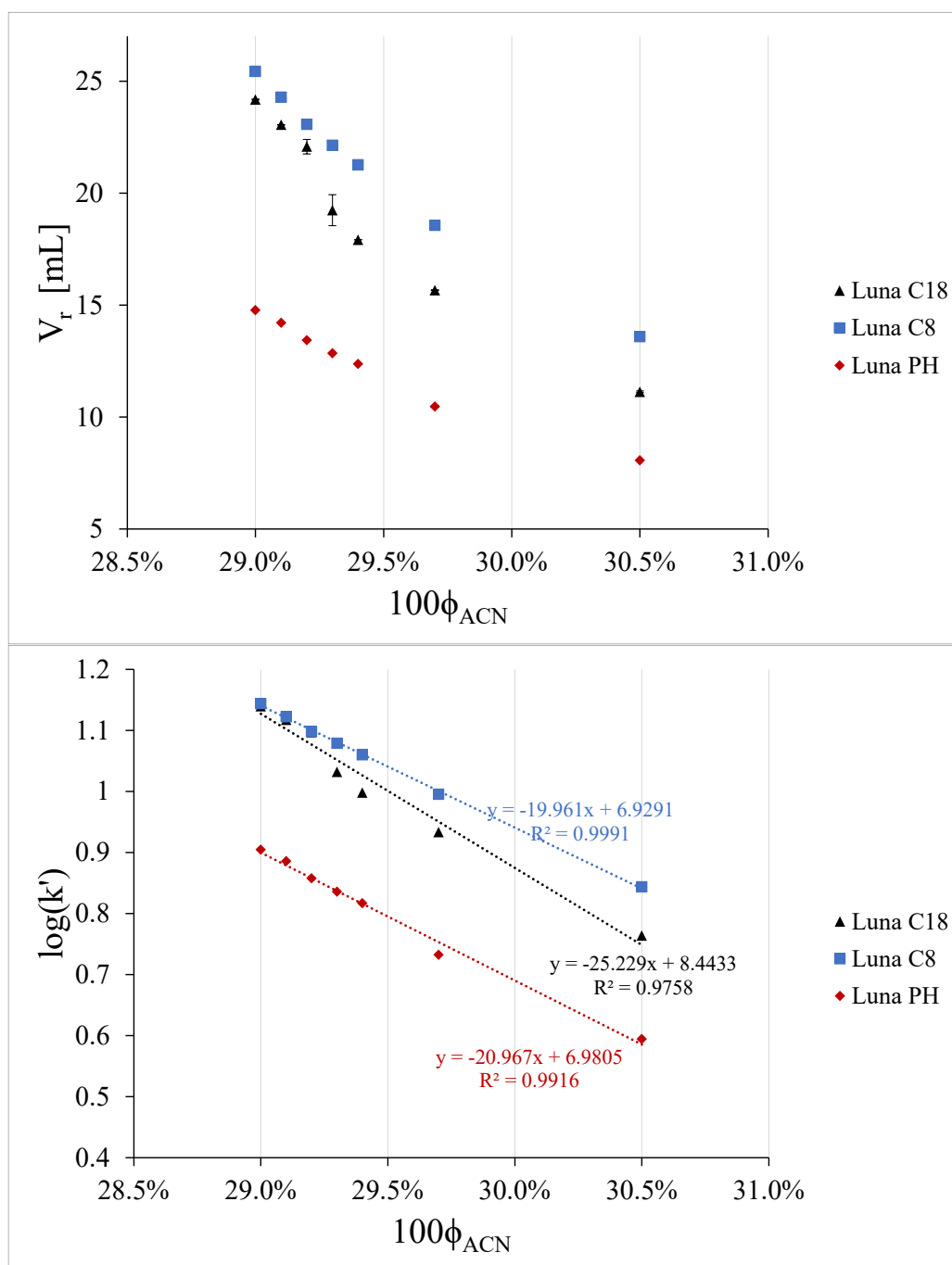


Figure 10.5 | Isocratic experiments for glucagon. Isocratic elution in terms of retention volume (V_R) and $\log(k)$ fitted to the run

Considering the known chemistry of glucagon, the surface properties of the resin, and the mobile phase contents, it is reasonable to assume that the adsorption-partition phenomena are governed by more types of chemical interactions other than hydrophobicity, such as electrostatic interactions and size exclusion effects, as well as dynamics such as multipath diffusion and mass

transfer. In fact, without adding the organic modifier TFA to the mobile phase, glucagon does not elute from the chromatography columns tested in this work.

Despite the complex partitioning process, the isocratic chromatographic-derived partitioning data for glucagon adheres to the LSS model ($r^2 > 0.95$). This behaviour is beneficial as it can assist with column selection, depending on the operation separation requirements. For instance, if a faster separation is required, a Phenyl-Hexyl column would be preferred as its $\log k_w$ is lower than other columns, and its importance over the control of solvent mixing is similar to C8 and less relevant than C18 ($S_{\text{Phenyl-Hexyl}} \approx S_{\text{C8}} < S_{\text{C18}}$). This correlates with findings from the excess adsorption and peptide adsorption analysis reported in Chapter 9. Faster desorption from the Phenyl-Hexyl column would be expected as the peptide only interacts with the phenyl groups.

From a process method design perspective, the S-value is also a measure of how carefully solvent mixing must occur at the pump delivery system. Considering an S-value of 15, the retention factor k reduces by approximately 40% by increasing 1%ACN of organic solvent. The parameter $\log k_w$ is an indicator of the affinity of the analyte to the surface of the chromatographic resin. Lower values indicate low adsorption affinity, thereby reducing the amount of solvent required to elute the analyte.

Another useful application of the LSS model is identifying if the peaks of selected solutes closer (critical peaks) to the main analyte of interest are of related or unrelated peptide structures. When the analytes are highly related chemical structures, the plots of $\log k$ vs ϕ_B of the critical peaks lie parallel to the main analyte, meaning the order of elution remains the same during the entire ϕ_B range. Critical peaks in a crude peptide sample are expected to be highly related structurally (i.e., homologues), meaning they are most likely to follow LSS behaviour.

In accordance with Figure 10.6-Figure 10.8, the analyte of interest and the critical peaks follow the LSS model, and the elution order does not change within the ϕ_B range. Here, I can identify that the most challenging separation is with any critical peaks which appear before the main peak for all columns, as the $\log k_w$ values for that peak and the main peak of interest are too close to each other. Additionally, even though the use of PH columns reduces the amount of mobile phase required, separation at low ϕ_B can become difficult as the resolution between the critical peak after the main peak and the main peak worsens at lower solvent strength ($S_{\text{MAIN}+1} <$

S_{MAIN}). This type of analysis can be executed on all peaks in the chromatogram; however, applying it to critical peaks is more convenient.

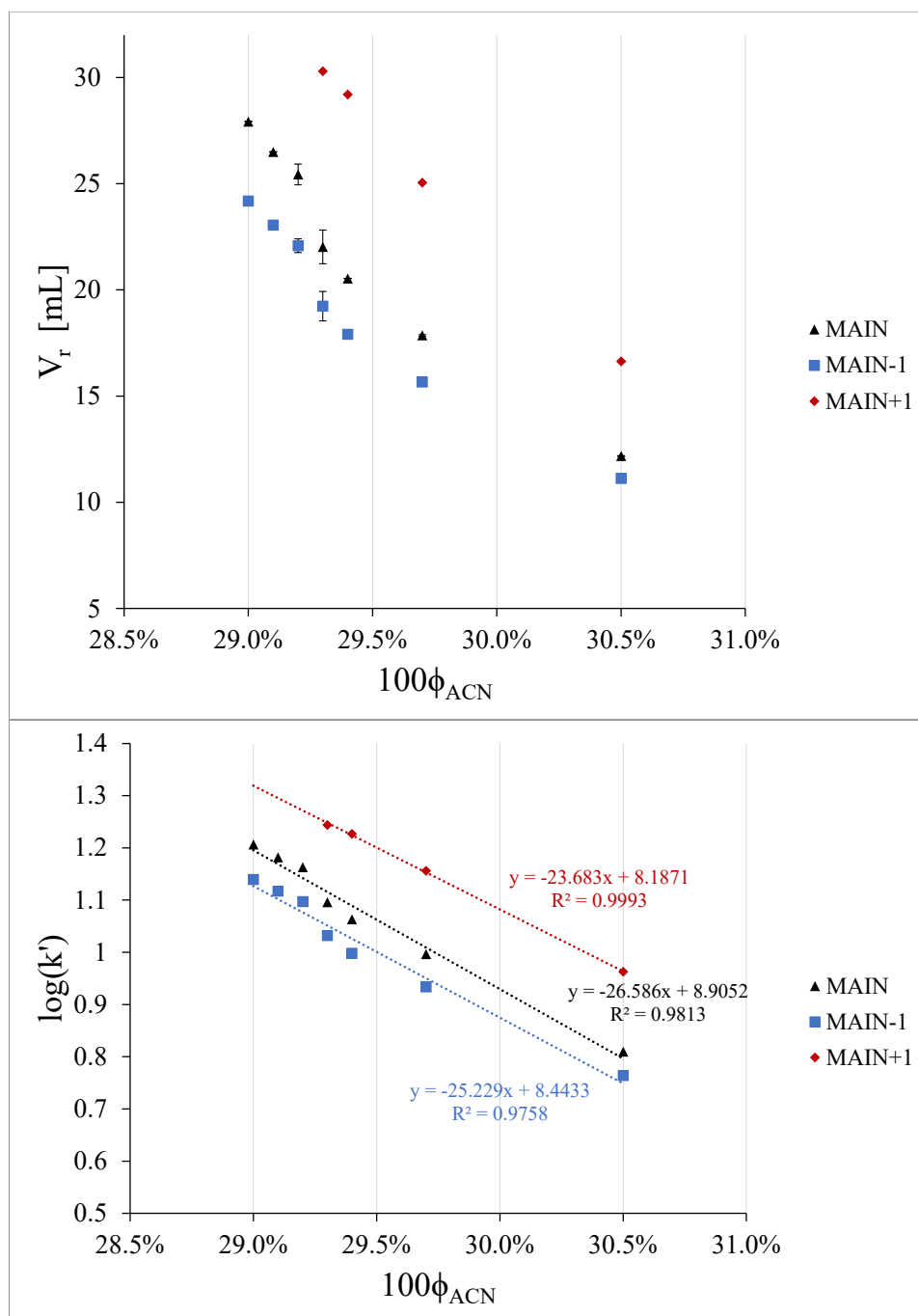


Figure 10.6 | LSS model applied to main peak and critical peaks in a Luna C18 column.

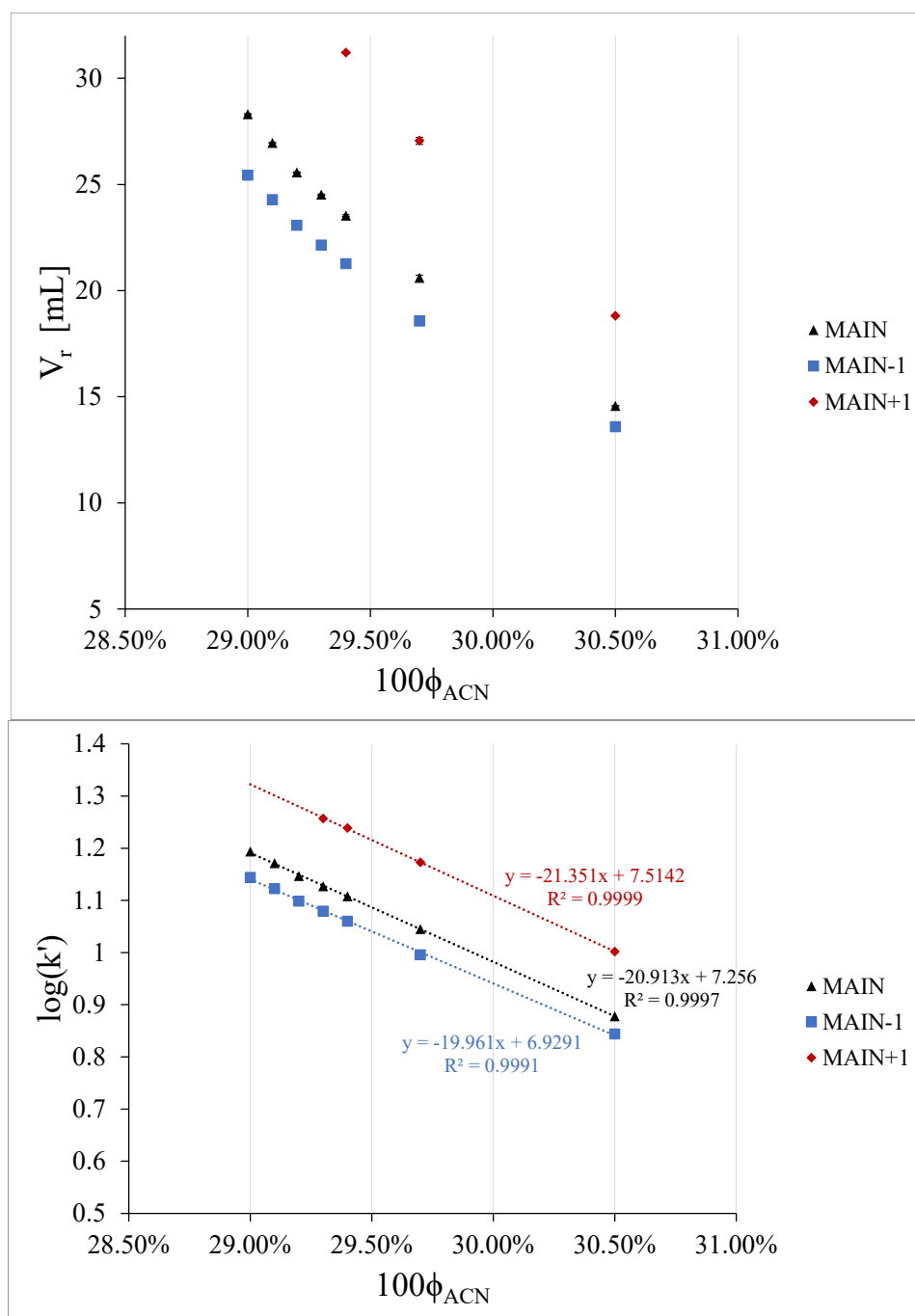


Figure 10.7 | LSS model applied to main peak and critical peaks in a Luna C8 column.

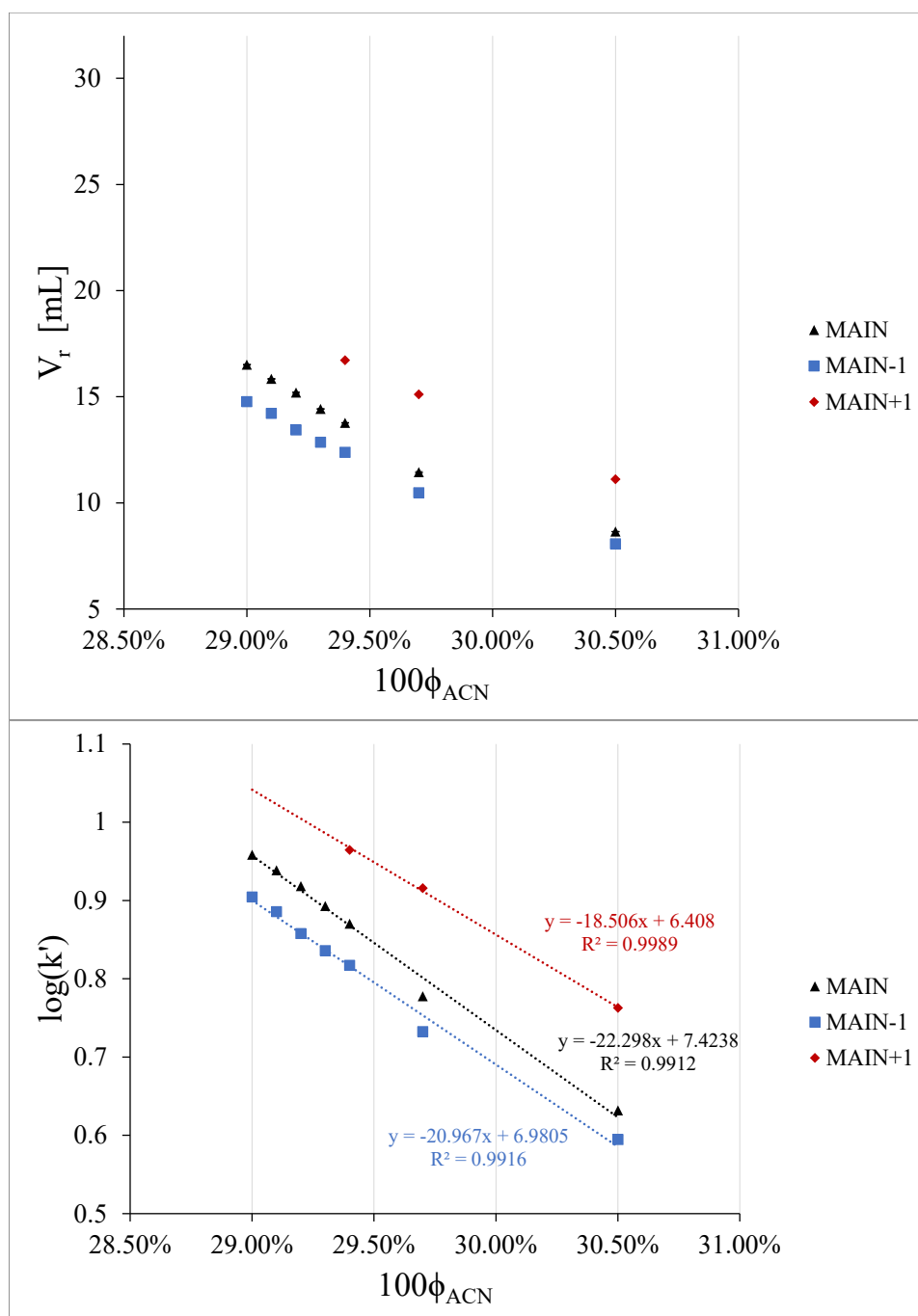


Figure 10.8 | LSS model applied to main peak and critical peaks in a Luna Phenyl-Hexyl column.

Although method design based on the LSS model is useful for some separations, but more detailed analysis beyond simply $\log k$ versus ϕ_B plots must be performed for more challenging separation problems. The LSS model only considers the first moment of the chromatographic peak (i.e., retention volume at the peak maximum or apex), so it does not analyse peak resolution or selectivity (risk of peak overlapping) aspects of the separation. Therefore, attention should be

paid to these additional chromatogram parameters, along with the LSS plots, to avoid overlapping peaks. The effect of chromatogram fronting and tailing is what determines the maximum separation possible for the RPLC systems.

Based on the findings from the LSS model, a final multi-isocratic elution profile was developed. In order to reduce the complexity of the separation, no specific separation goal was set. Instead, a single comparative method was developed in order to assess the performance difference between the three standard RPLC columns used in this Thesis. Results are shown in Figure 10.9. A clear distinction of the main peak can be observed, which confirms the efficacy of this type of rational method design.

The efficiency of purification by using a standard gradient elution method and comparing it with the newly designed method described in Figure 10.9 was assessed using the analytical C18 column with the crude glucagon samples. The maximum purity of glucagon recovered was observed to be less than 85% when the crude sample was purified using the gradient elution method. This suboptimal purification is ascribed to the inadequate resolution of a co-eluted shoulder and other closely associated impurities, requiring the ‘shaving’ of the primary peak of interest. Additionally, common process aggregation and degradation products of glucagon were found as early peaks[175, 176]. Conversely, the new multi-isocratic purification method yielded a purity approaching 94%. There was a substantial reduction in sample impurities as well as aggregation and degradation products. The chromatograms can be examined in Figure 10.10 and Figure 10.11. A summary of the purification analysis is presented in Table 10.2.

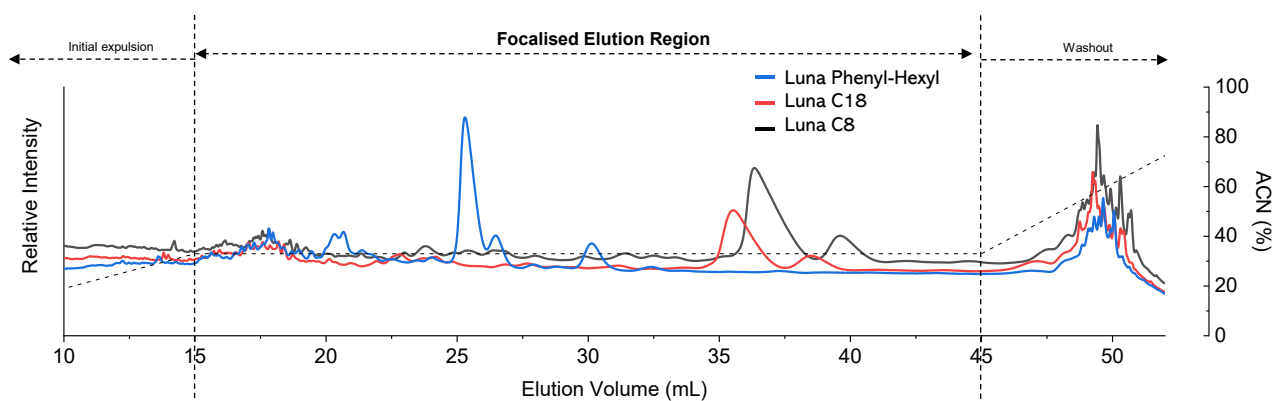


Figure 10.9 | Final method design for glucagon in three analytical columns. Flow rate of 0-5mL at 25%B, 5-20mL at 30.5%B, 12-13 mL at 25%B, 13-40 mL at 29.6%B, 40-45 mL at 50%B, and 45-55 mL at 25%B.

Table 10.2 | Purification analysis of crude glucagon reaction mixture in C18 analytical column.

Process	Mass injected	Main peak purity	Purity improvement	Sample impurities	Aggregation and degradation products
Crude	0.3 mg	27.0 ± 0.73%	N/A	~72.90%	N/A
Gradient 2.5%ACN/CV	2.1 mg (7 batches of 0.3 mg)	83.5 ± 1.56%	209%	~6%	~8%
SMART algorithm	1.5 (2.5 batches of 0.6 mg)	93.6 ± 0.82%	246%	<0.1%	~5%

It is important to mention that the new multi-isocratic method allows for a two-fold increase in the load capacity of the column compared to the gradient method. This advantage is attributed to the method's intrinsic format, designed to elute the primary peak of interest as a single peak, thereby eliminating the necessity for peak shaving. If the same loading conditions were applied to the gradient method, it would merge closely related peaks, leading to a substantial reduction in primary peptide recovery. This highlights a significant advantage of the multi-isocratic method, as it necessitates fewer batches to achieve a specified yield.

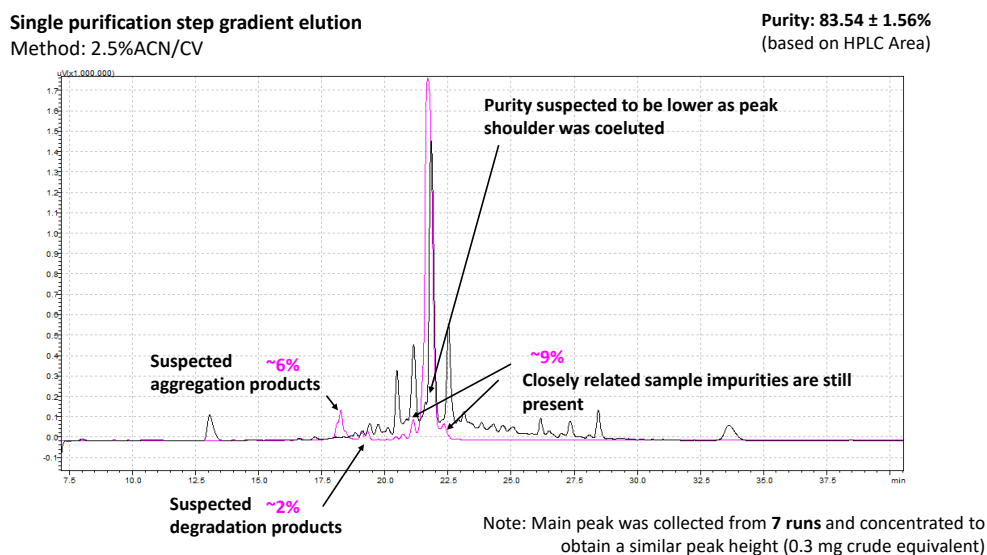
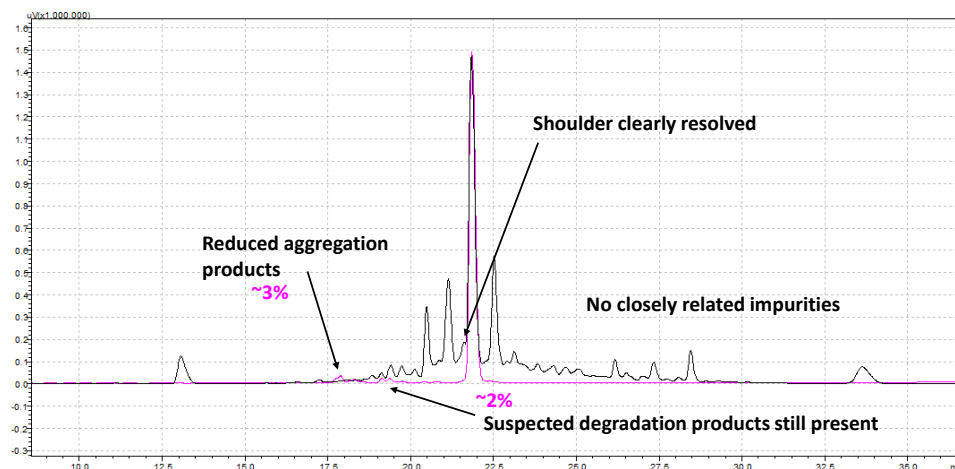


Figure 10.10 | Comparative chromatograms of crude glucagon gradient purification. The black chromatogram in both graphs represents an injection of crude glucagon sample. The pink chromatograms are the concentrated purified samples from a standard linear gradient method (upper) and the SMART method (lower).

Single purification step SMART method
 Method: main peak completely resolved in 9CV

Purity: $93.58 \pm 0.82\%$
 (based on HPLC Area)



Note: Main peak was collected from **2.5 runs** and concentrated to obtain a similar peak height

Figure 10.11 | Comparative chromatograms of crude glucagon SMART purification. The black chromatogram in both graphs represents an injection of crude glucagon sample. The pink chromatograms are the concentrated purified samples from a standard linear gradient method (upper) and the SMART method (lower).

10.4 Glucagon – scaling up

A scale-up column was integrated into both the analytical and semi-preparative systems to evaluate the scalability of the method. These columns, while identical in length and resin, differ in their internal diameters, resulting in scale-up ratios of 4.73 ($V_{C,\text{Semi-prep}}/V_{C,\text{analytical}}$) and 4.32 ($V_{M,\text{Semi-prep}}/V_{M,\text{analytical}}$).

One of the primary objectives at the manufacturing scale is to reduce solvent consumption. To this end, isocratic runs for the semi-preparative column loaded in the analytical system were conducted within the ϕ_B range of 30.0-35.0%. This range overlaps at the higher end with the analytical column experiments. If the scale-up performance is direct, considering only the change of V_M , this would ensure a $\log k$ value of less than or equal to 1.

The results, as depicted in Figure 10.12, reveal significant disparities in retention values between the analytical and semi-preparative columns in relation to the required mobile phase volume. However, when specific V_M values are considered and the logarithm of the capacity factor ($\log k$) is calculated, the results display a favourable overlap. This observation remains consistent even when the retention volume is converted to the number of column volumes.

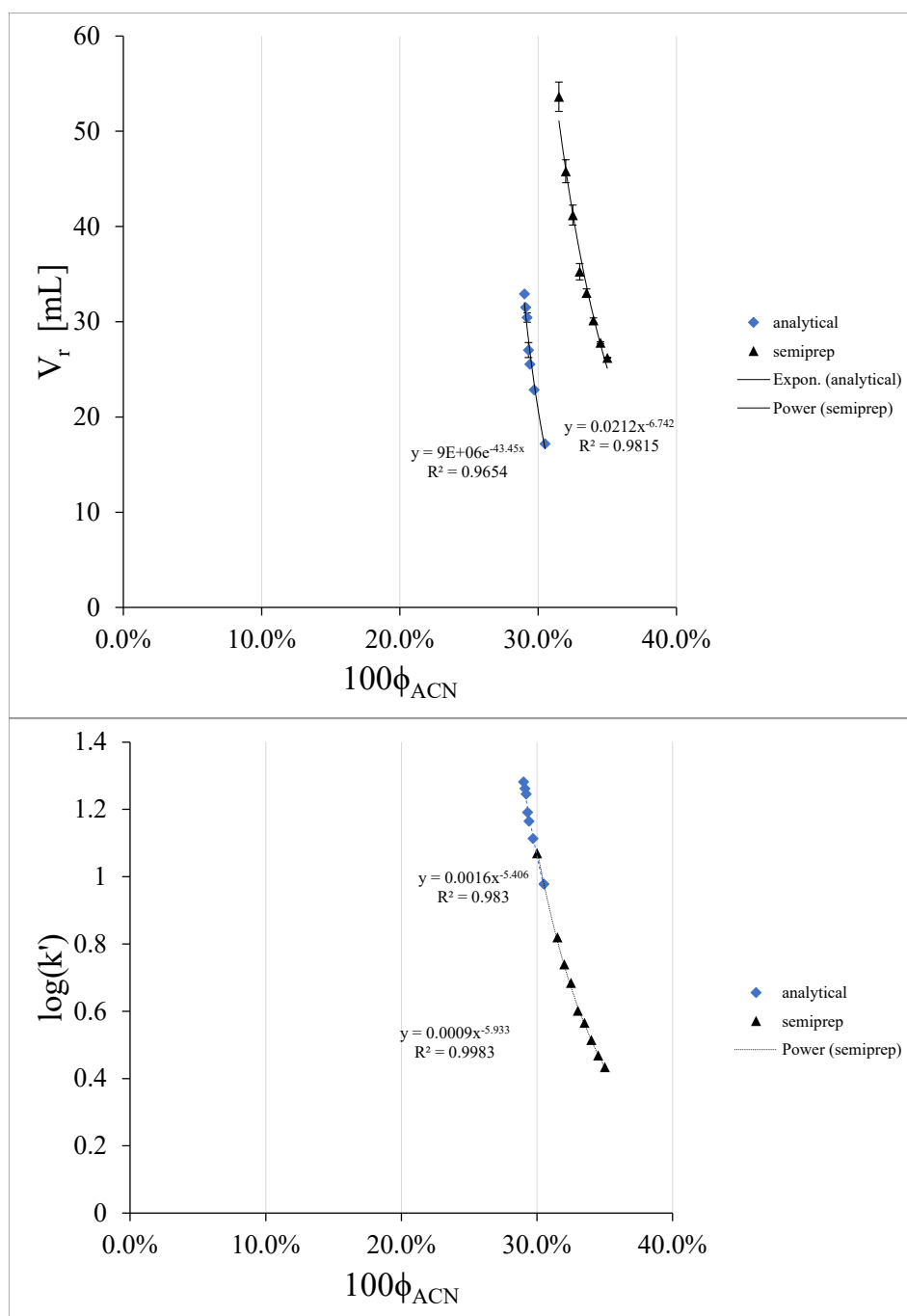


Figure 10.12 | Analytical versus semi-prep columns in analytical LC system. The analytical and semi-prep C18 columns were both connected to the analytical LC (Shimadzu). The same experiments were conducted, adjusting methods to the relevant scale-up factor.

While the use of V_C (column volume) is a common practice, it assumes identical column packing. By contrast, employing V_M provides a more reliable process scaling factor indicator in such cases, as shown in Figure 10.13. These types of graphs are invaluable for assessing solvent usage and the scalability of the process. Therefore, it is crucial to consider factors beyond $\log k$ vs ϕ_B plots

in the analysis. Although the LSS model only considers the first moment of the chromatographic peak (i.e., retention volume at the apex), it does not include either chromatogram express resolution or selectivity (risk of peak overlapping) in its analysis. Hence, these parameters should be given due consideration alongside the plots.

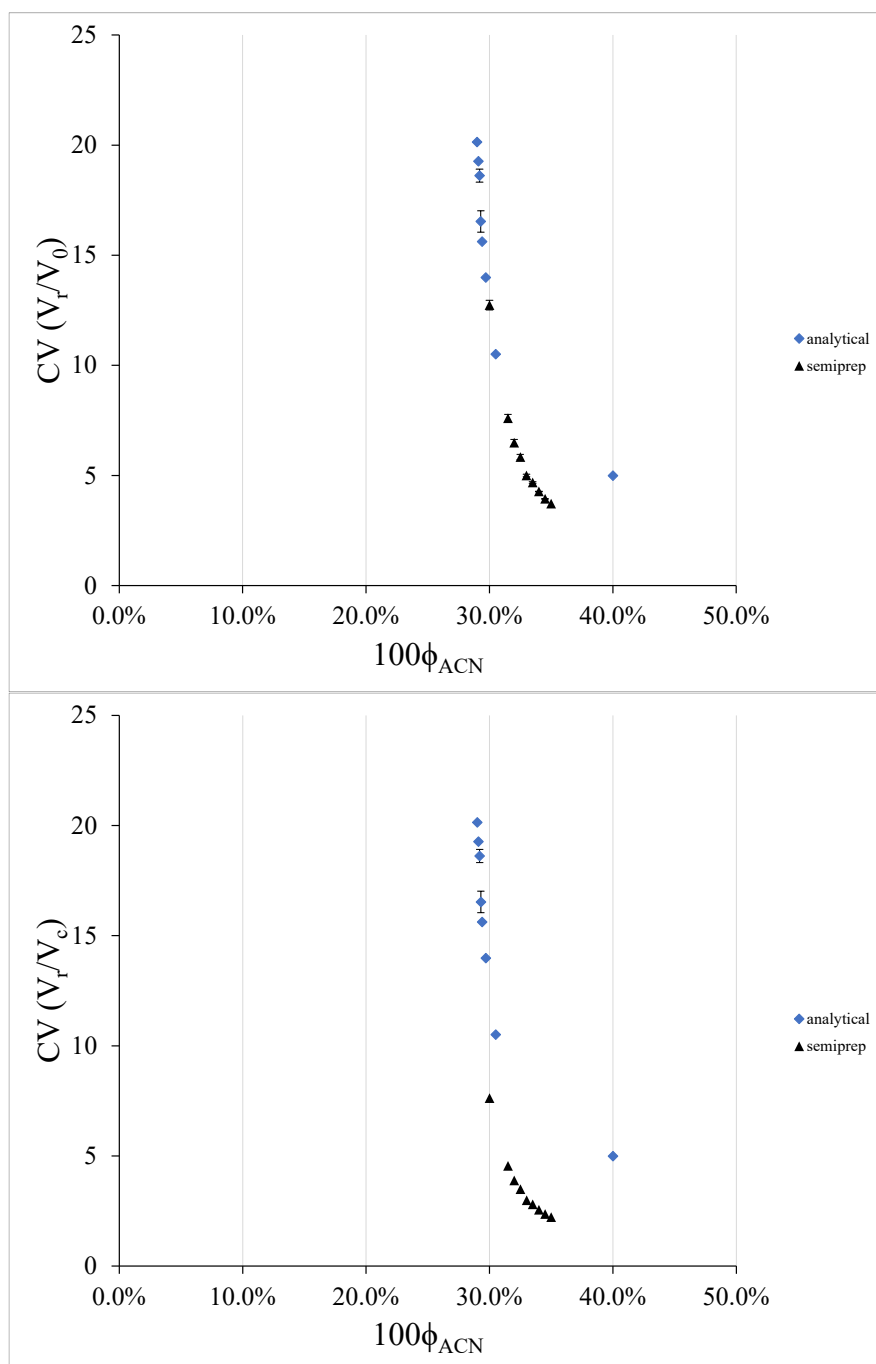


Figure 10.13 | Analytical versus semi-prep columns in analytical LC system. The analytical and semi-prep C18 columns were both connected to the analytical LC (Shimadzu). The same experiments were conducted, adjusting methods to the relevant scale-up factor.

When transferring methodology across chromatographic systems, an additional disadvantage arises when using V_C . This disadvantage is attributed to the variations in extra column volumes (V_{EC}) within the system, which can be due to differences in tubing size, the mixing chamber, or the pump heads.

In an experiment shown in Figure 10.14, the semi-prep column was installed in the AKTA (semi-prep) LC, and an isocratic range identical to the one previously described was executed. Following the adjustment of V_R values to align with the correct V_{EC} of the semi-prep system and V_M of the column, it was observed that the units yielded comparable results in terms of mobile phase consumption.

It is proposed that the minor variations at certain points in these plots are attributable to the inadequacy of the mixing chamber in the semi-prep system. More significant variations, particularly at lower ϕ_B , are believed to be caused by the different pump types employed by the two chromatographic units. The analytical LC utilises a parallel-type pump (quaternary and binary), whereas the semi-prep unit has dual pumps and a gradient mixer. These features in the semi-prep system are ideally suited for gradient purification as they minimise mobile phase pulsation and, therefore, improve the accuracy in blending solvents during isocratic runs. However, it is postulated that the higher ratio of tubing volume to mixing chamber volume in the semi-prep system, coupled with less precise flow rate control, could have significant implications if regular mobile phase composition calibrations are not performed. This point highlights the importance of routine mobile phase composition calibration in ensuring the accuracy and reliability of the system.

Although these results highlight the potential for scaling up peptide purification through the SMART method, further experiments are required to draw clearer conclusions. A broader range of peptide mixtures and purification conditions are needed to validate the scalability and robustness of these techniques. Specifically, additional investigations into the effects of varying column types, gradient profiles, and peptide structural diversity will help ensure that the proposed SMART method can be applied to more complex and diverse purification challenges.

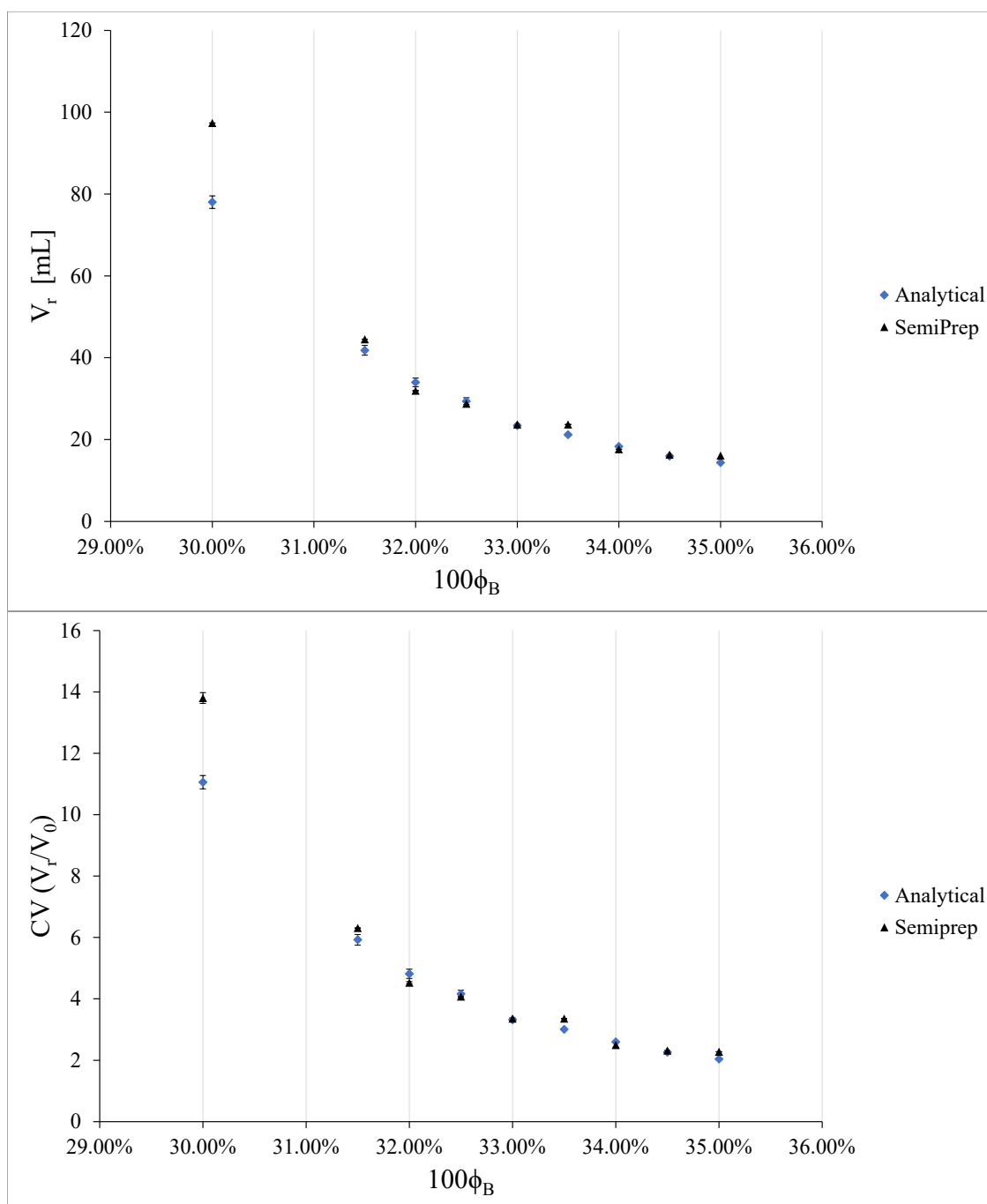


Figure 10.14 | Semi-preparative and analytical LCs comparison. The C18 semi-preparative column was connected to the analytical and preparative systems. Same isocratic experiments were conducted in both systems.

10.5 Conclusions

In this Chapter, a comprehensive analysis and application of the SMART (Systematic, Multi-parametric, Adaptive, Responsive, Tailored) algorithm for optimising multi-isocratic elution methods in RPLC has been introduced.

This study focused on the separation and purification of complex reaction mixtures, specifically peptides, and demonstrated the efficacy and advantages of the SMART algorithm in enhancing chromatographic separation performance.

The SMART algorithm was found to be highly effective in optimising the separation of peptide standards and glucagon, resulting in significant improvements in the purity and resolution of the target compounds. For the peptide standards P2, P3, and P4, the algorithm facilitated better control over peak elution and spacing during multi-isocratic runs, leading to higher selectivity and reduced separation times compared to traditional gradient methods.

In the case of glucagon, the SMART algorithm enabled the development of a multi-isocratic method that achieved a purity of 93.6%, markedly higher than the 83.5% purity obtained using a gradient elution method. This method also allowed for a two-fold increase in column loading capacity, reducing the number of required purification batches and enhancing overall process efficiency. These results highlight the algorithm's ability to manage complex separations effectively, offering substantial advantages over conventional gradient elution methods in terms of purity and throughput.

The potential scalability of the SMART algorithm was demonstrated by successfully transitioning from analytical to semi-preparative scales while maintaining consistent performance and solvent usage. This robustness underscores the algorithm's suitability for various scales of operation, making it a valuable tool for both research and industrial applications where scalability is crucial.

The study also highlighted several advantages of the multi-isocratic approach over traditional gradient methods. The SMART algorithm provided better resolution of closely eluting impurities and significantly reduced the presence of aggregation and degradation products. Additionally, the efficiency in solvent usage and the ability to handle higher load capacities make this

approach particularly advantageous for industrial applications where resource management is a critical factor.

In conclusion, the SMART algorithm represents a significant advancement in chromatographic separation and purification processes for complex biomolecules. Its application can enhance efficiency, selectivity, and scalability, offering a promising alternative to traditional gradient methods. Future research should aim to expand the application of the SMART algorithm to a broader range of biomolecules and complex samples, further refine its adaptability, and investigate the long-term stability and reproducibility of the developed multi-isocratic methods. These efforts will contribute to establishing the SMART algorithm as a standard approach in both analytical and preparative chromatography.

Chapter 11

Liquid-solid surface adsorption in RPLC

This Chapter explores the complex mechanisms of liquid-solid surface adsorption of solutes in reversed-phase liquid chromatography (RPLC), a critical process that underpins the separation and purification of analytes. Developing a fundamental understanding of how various factors, such as adsorbent surface chemistry, including surface ligands, solvent composition, and particle size, influence adsorption at the liquid-solid interface is crucial for optimising chromatographic performance. This Chapter examines the fundamental principles governing these surface physicochemical interactions using a range of experimental methods, including the minor disturbance method (MDM) for characterising column surfaces, and analyses how excess adsorption isotherms provide insight into the distribution of organic and aqueous molecules on the surface of the stationary adsorbent phase. This Chapter aims to comprehensively understand solute adsorption phenomena from the mobile phase that occur during RPLC and their implications for improving separation efficiency and selectivity in both analytical and preparative chromatography.

11.1 Effect of adsorbent surface ligand chemistry

RPLC columns are known for the hydrophobic nature of the adsorbent stationary phases. When hydrophobic interactions are the primary mode of adsorption, the hydrophobic ligand molecular length and its chemical structure directly affect the RPLC adsorption mechanism in the column. In the case of alkyl chains, like in C1, C8 and C18 (Figure 3.3), the length of the carbon chain increases the hydrophobicity of the ligand. Therefore, stronger hydrophobicity-driven adsorption with similarly hydrophobic analytes/adsorbates would be expected. As the aqueous/organic mobile phase percolates through the column, the solvent molecules distribute throughout

the surface of the adsorbent packing material. From a solute partitioning perspective, this solvent molecule distribution can be used to understand the retention mechanism for RPLC and the complementary hydrophilic interaction chromatography (HILIC), as both techniques have similar partitioning fundamentals.

For HILIC, it has been observed from molecular dynamic (MD) simulations that when an ACN/H₂O mixture is in contact with a silica surface, the water molecules form a 1.5 Å-wide region next to and overlap with the surface silanol groups. Adjacent to, but not overlapping with the water layer is a 4 Å-wide region containing acetonitrile molecules. The thickness of these layers is independent of the solvent mixture composition[177]. In the case of RPLC, MD simulations are more complex due to the double-end capping (ligand/bonded-phase and end-capping group) on the surface of RPLC silica packing materials. Here, surface diffusion is connected to the enrichment of ACN molecules, called ACN ditch, between the ends of the bonded-phase chains and the bulk liquid caused by the hydrophobic effect[178, 179]. Refer to Figure 11.1 for visual representation.

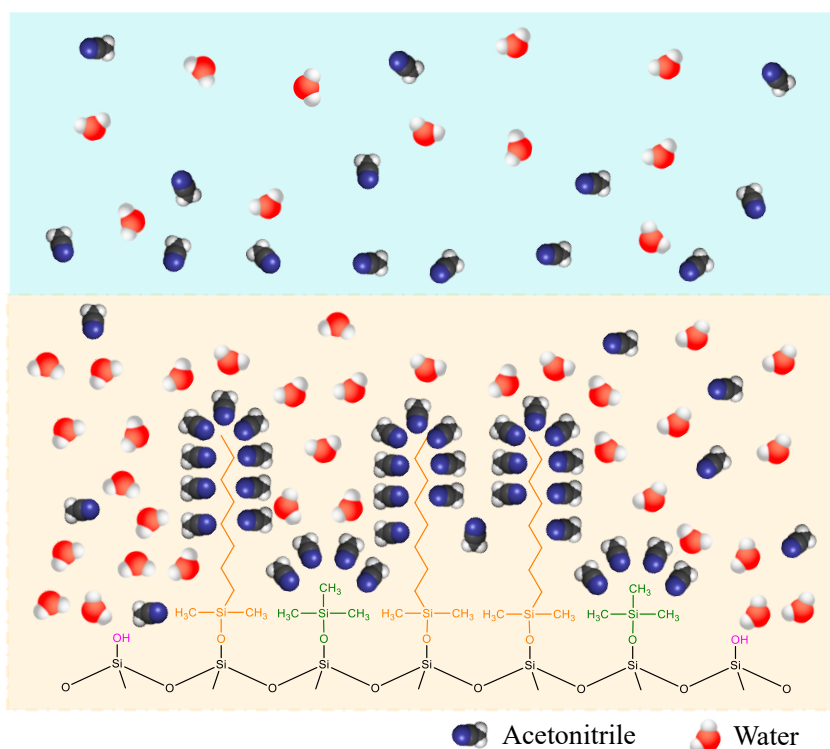


Figure 11.1 | Migration of acetonitrile and water molecules in an RPLC system. The yellow area is the stationary phase composed of the surface of the resin and the adsorption layer. The teal area is the bulk liquid of the moving mobile phase.

From an experimental perspective, as discussed in Section 4.6.1, excess adsorption isotherms are used to describe this adsorption distribution phenomenon of the binary organic/water mobile phase onto the stationary phase. Based on the competitive interaction of both solvent components with the different adsorption sites, the ratio of adsorbed organic to aqueous solvent on the resin surface characterises the surface hydrophobicity and the thickness of the interface layer.

The excess isotherm of ten columns was experimentally obtained through a perturbation approach called the minor disturbance method (MDM). This experimental method was first reported by Kazakevich and McNair[126] and followed by other authors[111, 180, 181]. The surface heterogeneity and the interface layer thickness were calculated based on the work of Yeung[182] and Gritti[111]. The full experimental methodology can be found in Section 6.3.5.

As seen in Figure 11.3, the retention volume of the disturbance peaks from the MDM method varies depending on the content of the organic solvent in the mobile phase. The retention volume increases at both concentration extremes ($B\% = 0$ and 100) due to the strong interactions between pure solvent molecules. In the case of water molecules, those interactions are hydrogen bonding, whereas dipole-dipole forces are the dominant intermolecular forces present between acetonitrile molecules. With pure organic solvents ($B\% = 100$) variations in V_M are expected as pure solvent molecules interact through dipole-dipole interactions between themselves and the hydrophobic adsorption sites. At this stage, although there is a gradient of molecules preferring the surface of the resin, there is no clear distinction between the adsorption layer and the bulk liquid (Figure 11.2). If equilibration is not performed correctly between runs, some residual molecules of other solvents can stay adsorbed onto a high-energy adsorption site. When a prolonged equilibration step is performed, displacement of these molecules is achieved.

Kazakevich and McNair[126] reported that the water retention volume in a C18 column equilibrated with 100% acetonitrile was inversely proportional to the injection volume. Similar results were observed with column C8 while injecting ACN in a column equilibrated with 100% water. The infinite dilution retention behaviour of probe solvents should always be prioritised by injecting the smallest quantity possible. However, going to a very small injection may make the detection of the eluting peak impossible. The chromatographer should choose a minimum possible injection volume that produces a reasonable signal at the HPLC detector.

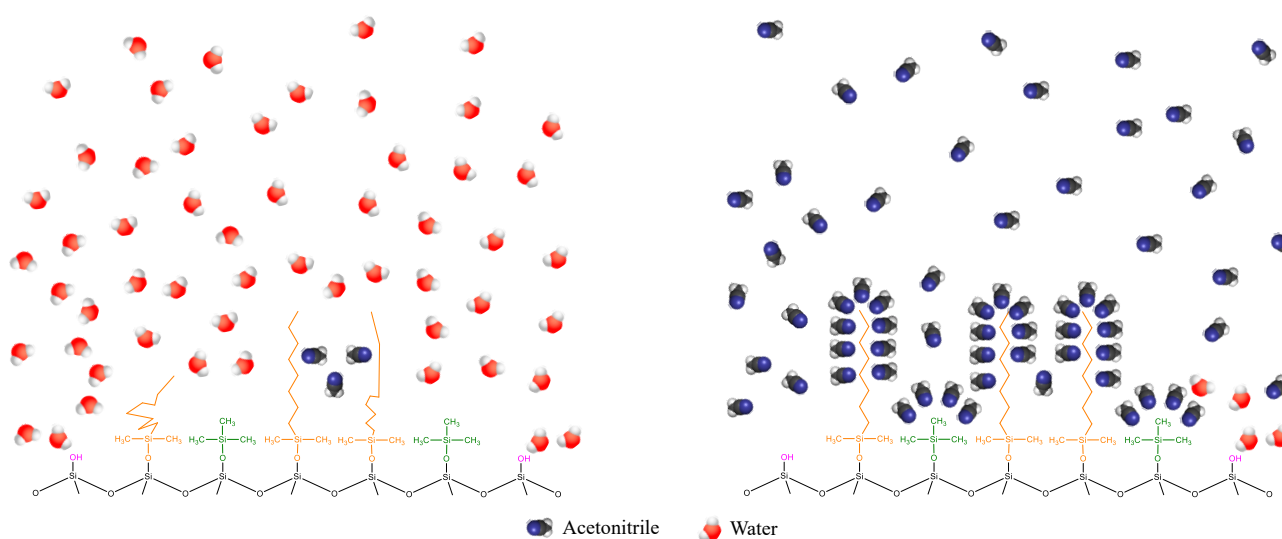


Figure 11.2 | Migration of pure acetonitrile and water molecules in an RPLC system.

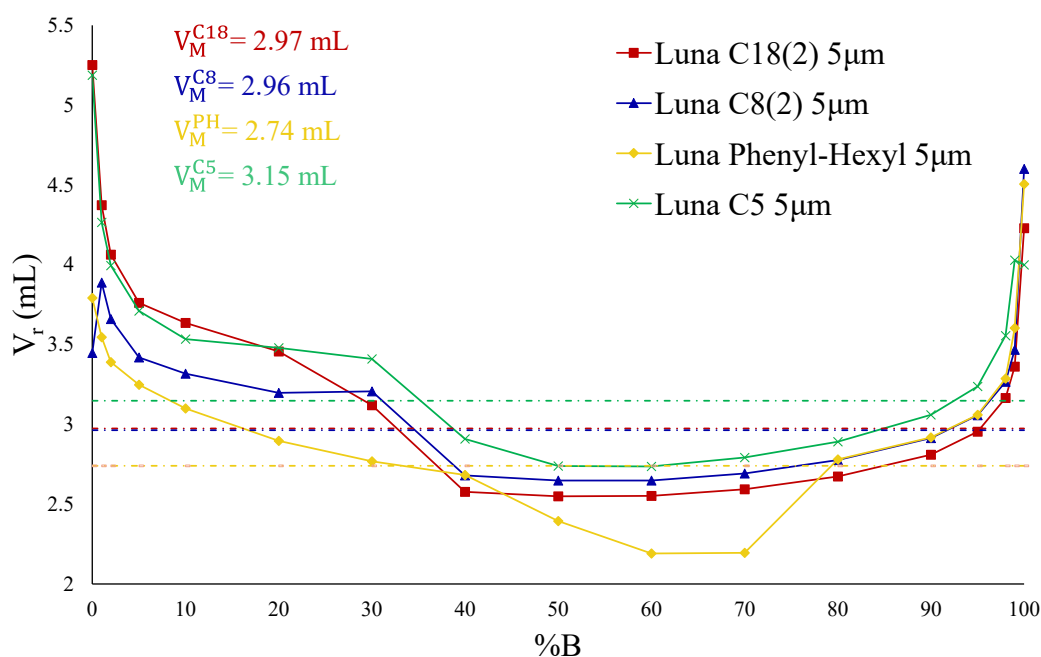


Figure 11.3 | Retention volume using the minor disturbance method in 5 μm columns using acetonitrile (ACN). The mobile phase comprised DI water (line A) and ACN (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 1 mL/min. Column details can be found in Table 6.4.

The V_r -value reduces when the column is equilibrated with a solvent mixture ($0 > B\% > 100$), with a substantial drop around the 1:1 volume ratio when the adsorbent surface is completely filled. A mixture of water and acetonitrile forms the interface during this process. This layer has a theoretically complex structure, being primarily composed of hydrogen-bonded ACN/ H₂O molecules[183, 184] with hydrophobic and hydrophilic adsorption zones. The hydrophobic zones consist of ACN-rich areas where the ACN molecules are attracted to the hydrophobic ligand, and the hydrophilic patches are enriched with water molecules drawn to the adsorbent's surface silanol groups. As the ACN content increases, H₂O molecules are displaced by ACN molecules.

The thermodynamic hold-up/retention volumes are calculated using equation 4.30 and follow the order C5>C8≅C18>Phenyl-Hexyl. Since the same manufacturer produced all columns with similar physical characteristics (Table 6.4), it is reasonable to assume that differences in retention volume are due to the chemistry of the surface, particle size and packing procedure.

From a simplistic perspective, assuming an ideal packing of the adsorbent material/resin, the only key variable is the length or linear extension of the chemical ligand. The V_M -value would be inversely proportional to the extension of the ligand. In this case, the expected order would slightly vary from the observed one, with C5 having the largest V_M , followed by C8, Phenyl-Hexyl and C18. Variations in ideality are due to the hydrophobic heterogeneity of the adsorbents in the column. The excess adsorption isotherm is a tool that can be used to analyse the surface of the packing material in the column *in-situ*.

The excess adsorption isotherms based on the MDM data from Figure 11.3 were calculated using equation 6.42 and are shown in Figure 11.4. The three distinct regions of an excess isotherm can be observed in all cases (see Section 4.6.1 and Figure 4.17).

Reaching $\Gamma_{B_{max}}^V$ is the final step of region I, after which the interface layer is filled out, and region II starts. In this region, a linear decrease of the surface excess is observed. It can be noted that this decrease causes an abrupt drop in V_r during the MDM method.

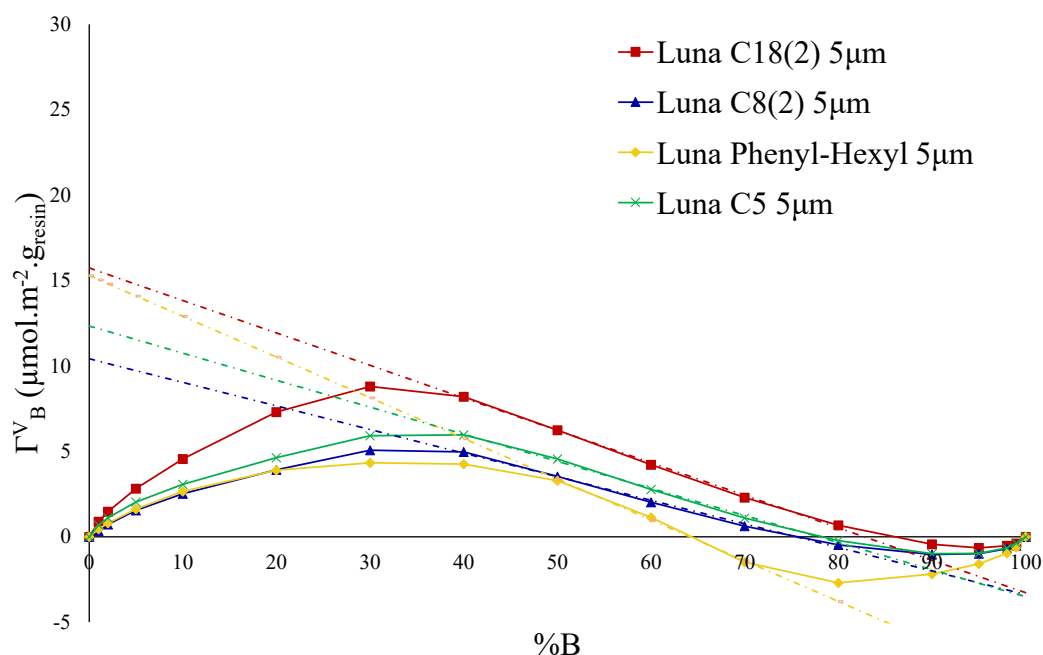


Figure 11.4 | Excess adsorption isotherms of acetonitrile (ACN) in water on 5 μm columns. The mobile phase comprised DI water (line A) and ACN (line B). Disturbance method in Figure 11.3. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

The surface of the resin/adsorbent material, which is in contact with the bulk liquid mobile phase, attracts some solution components. The excess isotherm measures the extra number of aqueous/organic molecules in the surface layer compared to the bulk liquid. Therefore, the linear decrease implies that as the adsorbent surface is filled, no extra molecules are accepted in the interface layer. Instead, to maintain thermodynamic equilibrium, organic molecules are replaced by the water molecules embedded within the interface as the organic solvent's volumetric ratio increases. In other words, the composition of the interface layer changes while increasing the concentration of organic, but the number of molecules remains the same until no or negative excess is observed where a $\Gamma_{B\min}^V$ is reached (end of region II).

When that linear decrease is extrapolated to the y-axis, Equation 4.34 becomes valid, and the value at %B=0 is the maximum amount of molecules on the surface because it is saturated with pure components. From Figure 11.4, we can see that C18 and Phenyl-Hexyl surfaces have similar surface capacity, around $15.5 \mu\text{mol.g}_{\text{resin}}/\text{m}^2$, followed by C5 and C8; 12.34 and $10.42 \mu\text{mol.g}_{\text{resin}}/\text{m}^2$, respectively.

From the same linear regression analysis, it is possible to calculate the contribution volumes of H₂O and ACN molecules at the saturation of the adsorption layer. From that data, it is possible to estimate the hydrophobic heterogeneity of the column surface based on the volumetric ratio of these solvent volumes (V_B/V_{H_2O}). Table 11.1 shows that Luna C18(2) has a higher surface hydrophobicity than the other tested columns. Despite its very high surface capacity, PH has the least hydrophobic surface based on the V_B/V_{H_2O} value. A potential explanation could be the formation of π -OH interactions between the phenyl groups at the end of the ligand and the H₂O molecules.

Although there is limited experimental information on differences in binding energy between different interaction types. It is generally believed that π -OH are thermodynamically less stable than other interactions, such as hydrophobic and π - π . However, through MD simulations, it has been estimated that π -OH interactions are stable at a distance of around 2.94 to 3.50 Å from the OH group to the centre of the phenyl group[185-187]. Although this distance is unlikely to occur in most liquid-solid systems, it could be possible in a pressurised system such as an HPLC, which could explain the high adsorption and low surface hydrophobicity. However, there is not enough evidence yet to support this hypothesis.

Table 11.1 | ACN/H₂O excess isotherm characteristics for 5 μ m RPLC columns.

Column	Particle size [μ m]	Solvent B	V_0 [mL]	Γ_B^V max [μ mol.g _{resin} /m ²]	C_B^L for $\Gamma_B^V = 0$	Γ_B^V min [μ mol.g _{resin} /m ²]	Linear range
Luna C5	5	ACN	3.15	6.08×10^{-3} (6.81 mol/L – 35.70%)	15.05 (78.90%)	-1.09×10^{-3} (15.05 mol/L – 78.90%)	7.66 – 15.32 (40–80%)
Luna C8(2)	5	ACN	2.96	5.15×10^{-3} (6.49 mol/L – 34.02%)	14.46 (75.81%)	-1.13×10^{-3} (17.68 mol/L – 92.69%)	7.66 – 15.32 (40–80%)
Luna C18(2)	5	ACN	2.97	8.72×10^{-3} (6.23 mol/L – 32.66%)	16.48 (86.40%)	-0.78×10^{-3} (18.22 mol/L – 95.52%)	7.66 – 15.3 (40–80%)
Luna Phenyl-Hexyl	5	ACN	2.74	4.36×10^{-3} (6.34 mol/L – 33.24%)	12.28 (64.38%)	-2.69×10^{-3} (15.45 mol/L – 81.00%)	9.57 – 13.40 (50–70%)

From the same linear relationship/regression analysis shown in Figure 11.4, a comparison of the layer adsorption and thickness of all columns was completed. As observed from Table 11.2, the adsorption layer is thicker for C18 and Phenyl-Hexyl, followed by C5 and C8. The number of total molecules adsorbed onto the surface is depicted in Figure 11.5. From these graphs, Phenyl-Hexyl provides the most information regarding the adsorption phenomenon occurring in the column. On first inspection, the adsorption is similar to that of C5, but as the content of ACN

builds up, the adsorption behaviour drastically changes and becomes closer to that of C18. This behaviour can be explained by the rapid replacement of highly dense structured H₂O molecules with ACN molecules within the interfacial layer.

Table II.2 | Surface properties of 5 μm RPLC columns with an ACN/H₂O adsorption interface layer.

Column	Particle size (μm)	Solvent B	m ($\times 10^3$)	C _{ads} or b ($\mu\text{mol.g}_{\text{resin}}/\text{m}^2$)	τ ($\text{\AA.g}_{\text{resin}}$)	V _{ads} (mL)	V _B (mL)	V _{H₂O} (mL)	V _B + V _{H₂O}	V _B /V _{H₂O}	V _B /(V _B + V _{H₂O})	(V _B + V _{H₂O})/V _{ads}
Luna C5	5	ACN	-0.83	12.34	6.47	0.26	0.71	0.20	0.91	3.52	0.78	3.50
Luna C8(2)	5	ACN	-0.72	10.42	5.46	0.22	0.54	0.18	0.72	3.08	0.75	3.27
Luna C18(2)	5	ACN	-0.99	15.74	8.25	0.330	0.82	0.17	0.99	4.80	0.83	3.00
Luna Phenyl-Hexyl	5	ACN	-1.25	15.29	8.02	0.321	0.80	0.45	1.25	1.79	0.64	3.89

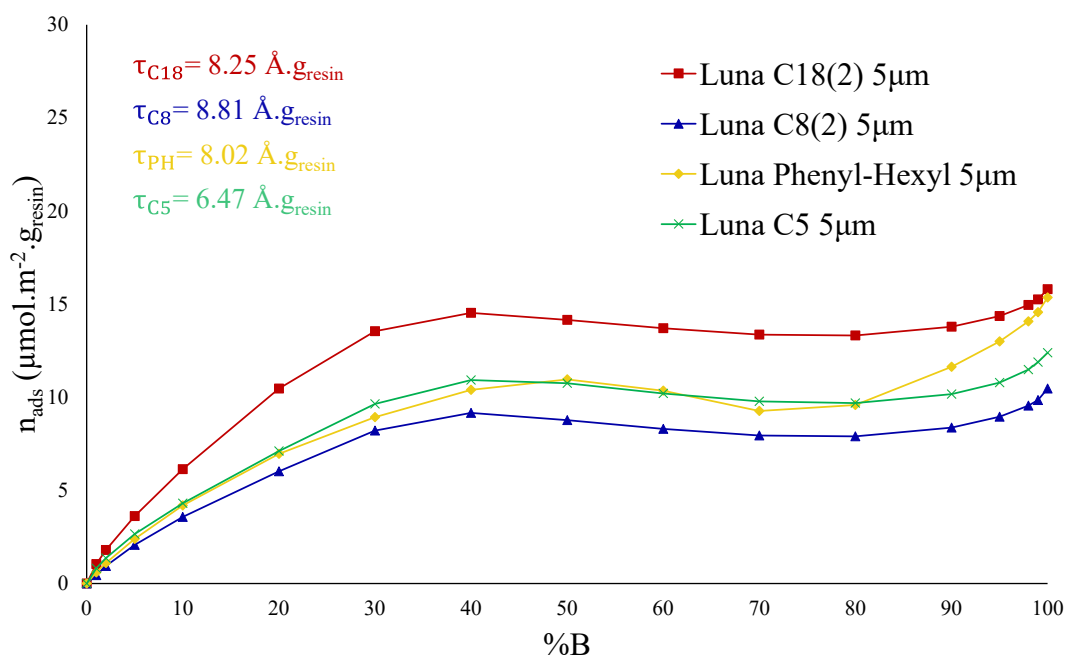


Figure II.5 | Total water/acetonitrile eluent mass adsorbed onto the 5 μm columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

Another important observation is that V_{ads} is 3 to 4 times lower than $V_{\text{B}} + V_{\text{H}_2\text{O}}$, despite having the same physical meaning. The main difference is that V_{ads} is derived from equation 4.35, which requires an estimation of surface area from N₂ adsorption experiments and comes from the

properties of the filled liquid stationary phase/adsorption layer based on the excess isotherm. I believe that since ACN/H₂O molecules are larger and the distance between atoms is greater than for N₂ molecules, then the adsorption area and volume of the packing material are higher and thicker in liquid than during dry conditions, leading to a higher apparent S_{BET} value with a conversion ratio of $(V_{\text{B}} + V_{\text{H}_2\text{O}})/V_{\text{ads}}$.

11.2 Effect of particle size

The same set of experiments were conducted in columns C8, C18 and Phenyl-Hexyl with an adsorbent particle size of 10 μm . Apart from the particle size, all other physical properties are the same as those of 5 μm , according to the manufacturer's information. Phenomenex also offers a selection of the same columns in a preparative (PREP) form. The surface area, carbon load and end-capping information are proprietary for these columns. Therefore, to investigate the effects of larger particle sizes in this Chapter, the other particle properties should be assumed to be the same as the non-PREP columns.

Figure 11.6 -Figure 11.9 show excess isotherms of all 10- μm columns based on the MDM methodology. The hold-up volume of all columns increased for C8 and Phenyl-Hexyl and reduced for C18 compared to the 5- μm columns. The hold-up volume differences for PREP and non-PREP columns were similar, with the most significant variation being for C18 (4.4%), followed by C8 (1.6%) and Phenyl-Hexyl (0.3%). A higher hold-up volume is expected for columns with larger particles as the packing becomes less efficient, and therefore, more intraparticle spaces can fill with the mobile phase. It is unclear why the opposite was observed for C18. Potential explanations could involve using highly polydispersed resin particles, denser ligand distribution or less efficient end-capping procedures; however, further experimentation involving unpacking the columns would be required to recover the adsorbent materials. This goes further than the scope of the current Thesis.

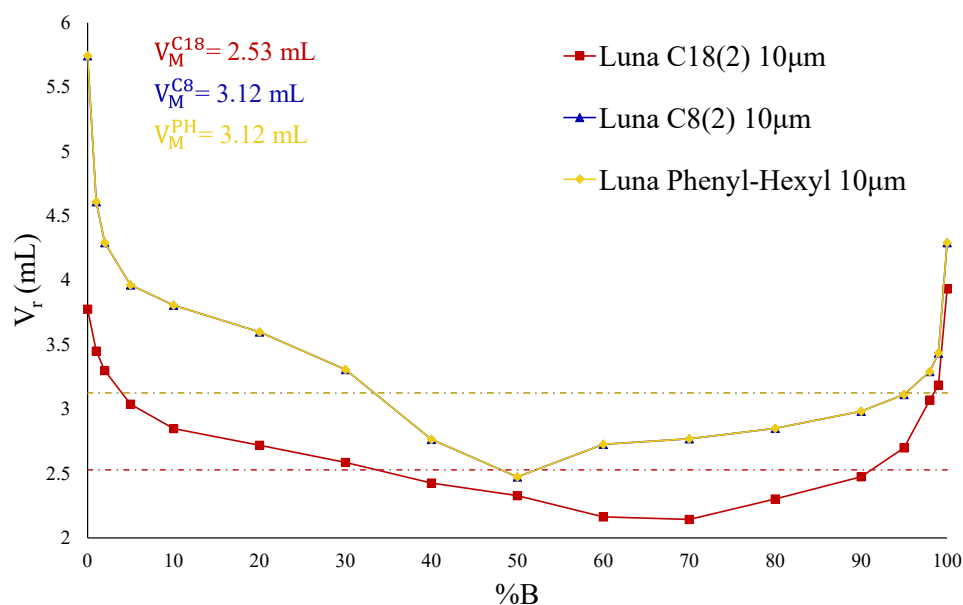


Figure 11.6 | Determination of V_M using the minor disturbance method in 10- μ m columns using acetonitrile (ACN). The mobile phase comprised DI water (line A) and ACN (line B). The disturbance was produced by injecting 1-5 μ L of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 1 mL/min. Column details can be found in Table 6.4.

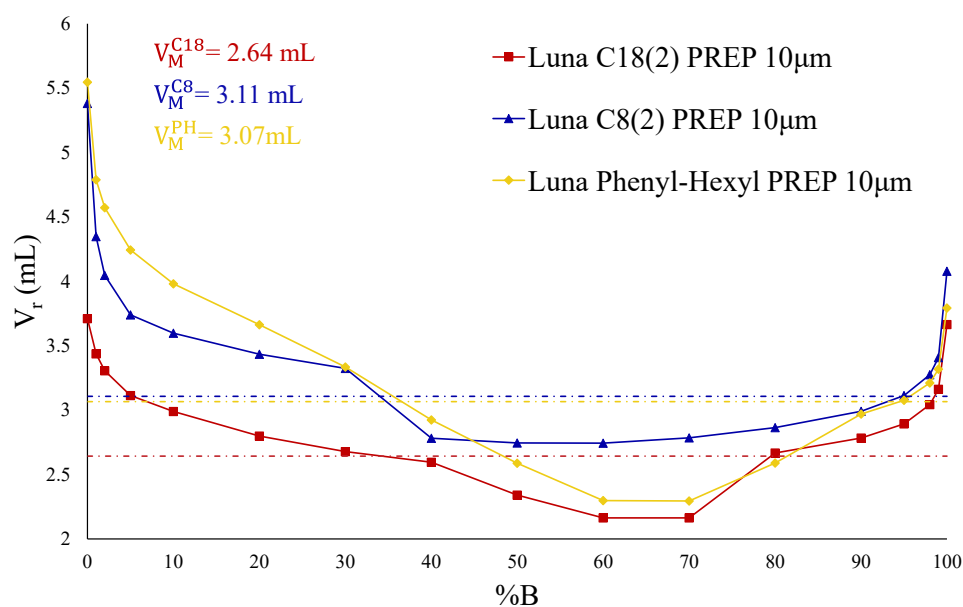


Figure 11.7 | Determination of V_M using the minor disturbance method in 10 μ m PREP columns using acetonitrile (ACN). The mobile phase comprised DI water (line A) and ACN (line B). The disturbance was produced by injecting 1-5 μ L of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 1 mL/min. Column details can be found in Table 6.4.

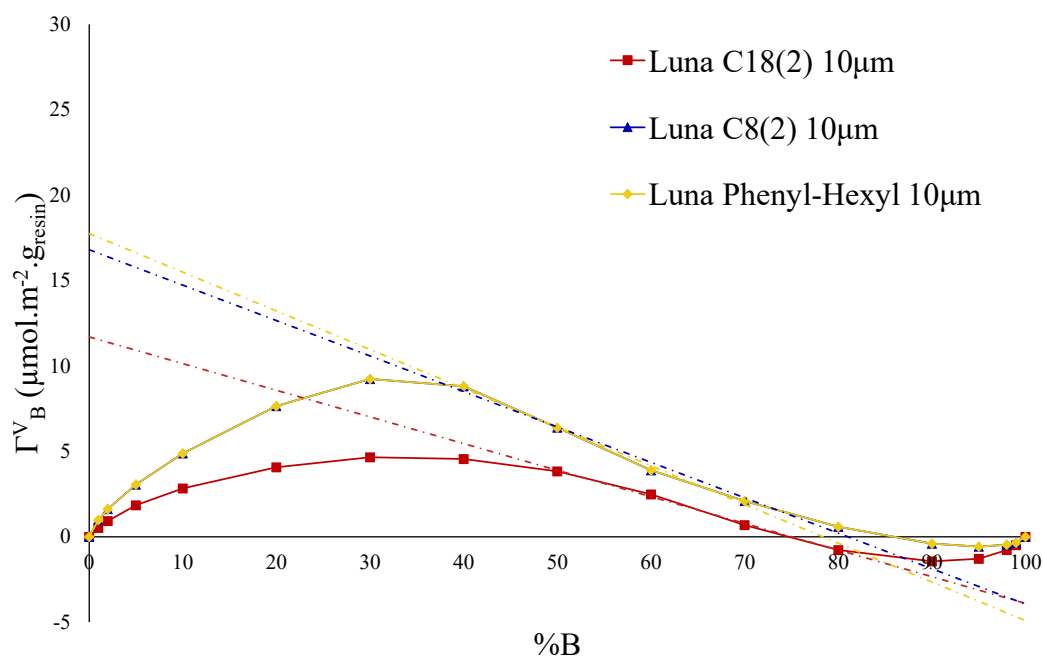


Figure 11.8 | Excess adsorption isotherms of acetonitrile (ACN) in water on 10 µm columns. The mobile phase comprised DI water (line A) and ACN (line B). Disturbance method in Figure 11.6. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

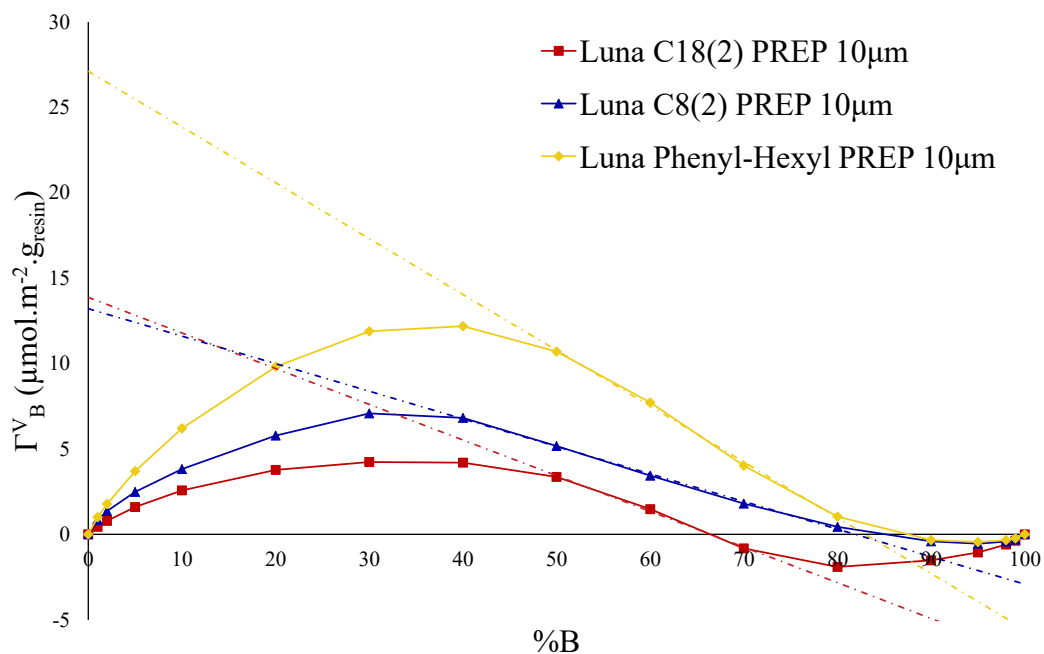


Figure 11.9 | Excess adsorption isotherms of acetonitrile (ACN) in water on 10 µm PREP columns. The mobile phase comprised DI water (line A) and ACN (line B). Disturbance method in Figure 11.7. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

In terms of surface capacity and surface hydrophobic heterogeneity (Table 11.3), similarities, such as those for V_M , are observed between PREP and non-PREP columns. It is evident that both properties of Phenyl-Hexyl and C8 are higher in the 10- μm columns and lower for C18. The PREP columns generally show more interesting results than their non-PREP counterparts. On the one hand, the surface adsorption of ACN/H₂O molecules on Phenyl-Hexyl PREP is significantly higher than in all columns. The surface hydrophobic heterogeneity is as high as the 10- μm C8 columns. On the other hand, despite the lower values of V_M , from the MDM experiments, the surface adsorption capacity and surface hydrophobic heterogeneity of C18 are very similar to those of C8.

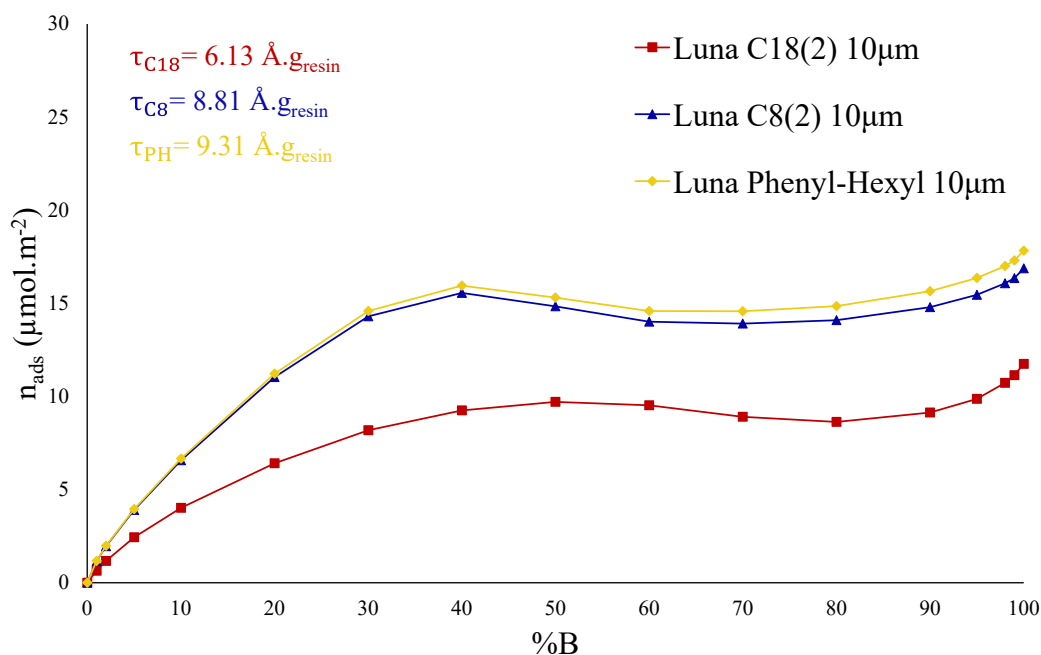
Table 11.3 | ACN/H₂O excess isotherm characteristics for 10- μm RPLC columns.

Column	Particle size [μm]	Solvent B	V_0 [mL]	$\Gamma_B^V \text{ max}$ [$\mu\text{mol.g}_{\text{resin}}/\text{m}^2$]	C_B^L for $\Gamma_B^V = 0$ [mol/L]	$\Gamma_B^V \text{ min}$ [$\mu\text{mol.g}_{\text{resin}}/\text{m}^2$]	Linear range
Luna C8(2)	10	ACN	3.12	9.26×10^{-3} (6.27 mol/L – 32.87%)	16.59 (86.98%)	-0.69×10^{-3} (18.19 mol/L – 95.36)	7.66 – 15.32 (40–80%)
Luna C8(2) PREP	10	ACN	3.11	7.17×10^{-3} (6.39 mol/L – 33.50%)	16.15 (84.67%)	-0.58×10^{-3} (17.88 mol/L – 93.74%)	7.66 – 15.32 (40–80%)
Luna C18(2)	10	ACN	2.53	4.72×10^{-3} (6.66 mol/L – 34.92%)	14.30 (74.97%)	-1.53×10^{-3} (14.31 mol/L – 75.02%)	9.57 – 15.32 (50–80%)
Luna C18(2) PREP	10	ACN	2.64	4.32×10^{-3} (6.96 mol/L – 36.49%)	12.67 (66.42%)	-1.93×10^{-3} (15.81 mol/L – 82.89%)	9.57 – 13.40 (50–70%)
Luna Phenyl-Hexyl	10	ACN	3.12	9.26×10^{-3} (6.29 mol/L – 32.98%)	16.59 (86.98%)	-0.69×10^{-3} (18.19 mol/L – 95.36%)	7.66 – 13.40 (40 – 70%)
Luna Phenyl-Hexyl PREP	10	ACN	3.07	12.26×10^{-3} (7.07 mol/L – 37.07%)	16.35 (85.72%)	-0.44×10^{-3} (17.68 mol/L – 92.69%)	9.574 – 15.32 (50–80%)

By analysing the properties of the interface layer in the 10- μm columns (Table 11.4 and Figure 11.10), it can be noticed that Phenyl-Hexyl columns show a thicker interface layer and more hydrophobic heterogeneity than the rest of the columns. If we consider complete homogeneity, the Berendsen-de Galan equation ligand density calculated for the 150-mm columns in Chapter 9 is the same for all these columns. the increased particle size increases the distance between residual silanol surface molecules and bonded ligands. This ‘minor’ change allows more ACN/H₂O molecules to fit within the surface layer, resulting in a thicker and more heterogeneous layer.

Table II.4 | Surface properties of 10- μm RPLC columns with an ACN/H₂O adsorption interface layer.

Column	Particle size (μm)	Solvent B	m ($\times 10^3$)	C_{ads} or b ($\mu\text{mol.g}_{\text{resin}}/\text{m}^2$)	τ ($\text{\AA.g}_{\text{resin}}$)	V_{ads} (mL)	V_{B} (mL)	$V_{\text{H}_2\text{O}}$ (mL)	$V_{\text{B}} + V_{\text{H}_2\text{O}}$	$V_{\text{B}}/V_{\text{H}_2\text{O}}$
Luna C8(2)	10	ACN	-1.08	16.81	8.81	0.35	0.88	0.26	1.14	4.26
Luna C8(2) PREP	10	ACN	-0.84	13.22	6.93	0.277	0.69	0.15	0.84	4.53
Luna C18(2)	10	ACN	-0.81	11.70	6.13	0.245	0.61	0.20	0.81	3.00
Luna C18(2) PREP	10	ACN	-1.09	13.88	7.28	0.291	0.72	0.37	1.09	1.98
Luna Phenyl-Hexyl	10	ACN	-1.18	17.76	9.31	0.372	0.93	0.26	1.19	3.62
Luna Phenyl-Hexyl PREP	10	ACN	-1.71	27.12	14.22	0.569	1.42	0.29	1.71	4.87

Figure II.10 | Total water/acetonitrile eluent mass adsorbed onto the 10 μm columns. Eluent adsorbed within interfaces of theoretical finite layer thickness τ .

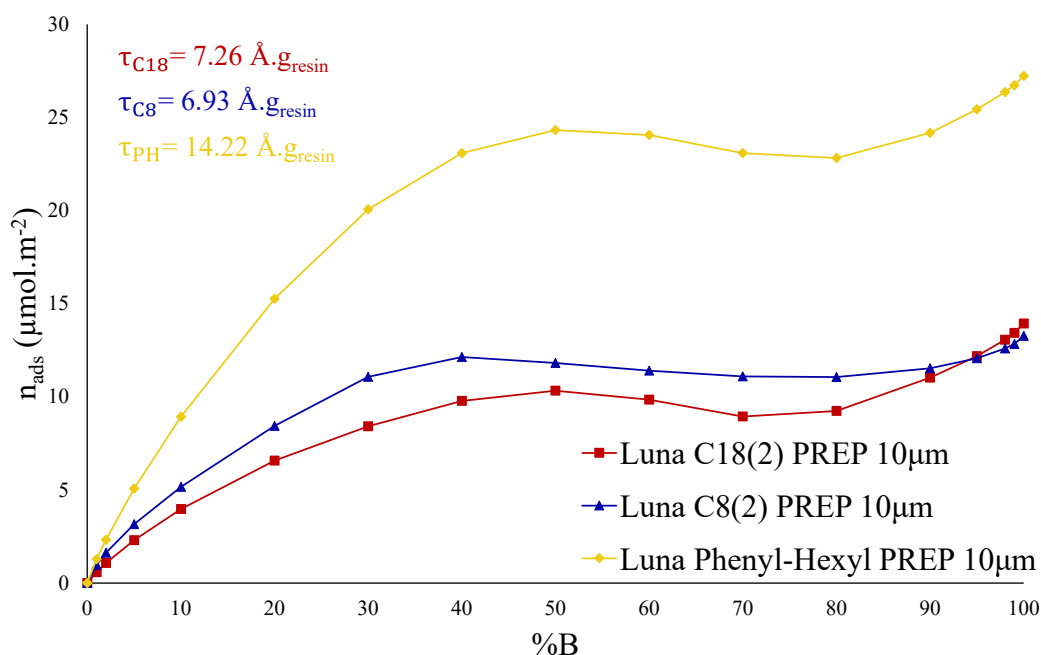


Figure 11.11 | Total water/acetonitrile eluent mass adsorbed onto the 10 μm PREP columns. Eluent adsorbed within interfaces of theoretical finite layer thickness τ .

11.3 Effect of organic solvent

Although the most common organic modifier used for peptide separations is exclusively ACN due to its higher eluotropic strength, reduced backpressure and the preference of regulatory agencies, other applications, especially for small molecules, involve using methanol (MeOH). The literature has explored empirical conversions and rules of thumb for adapting methods between these two solvents while preserving similar capacity factors and elution orders[188]. Excess isotherms of MeOH/H₂O were measured in all 10 columns to understand the fundamental effect of changing the organic solvent. Ethanol (EtOH) was also evaluated as its potential use in peptide separation is desired due to its green credentials despite its very weak eluotropic strength.

According to the classical misunderstanding of chromatography, the V_M of a column should not change regardless of the solvent used. This notion is valid exclusively for the absolute void volume of a column, which can only be estimated in dried conditions and represents no practical value to real separation phenomena[189]. During separation conditions, the eluent molecules adsorbed onto the surface, forming a (liquid) stationary phase composed of the ligands attached to the surface and the solvent molecules. This adsorbed layer reduces the absolute void volume, leaving what is known as the thermodynamic void volume or hold-up volume[190]. The

properties of this layer are dependent on the inter- and intraparticle properties of the solvents. Therefore, the V_M varies from molecule to molecule and exhibits complex behaviour with solvent mixtures such as those used as mobile phases in RPLC separations.

Tables Table 11.5 and Table 11.6 summarise the results from these experiments. Figures showing the MDM experimental results (Figure 11.12 - Figure 11.17), excess isotherms (Figure 11.18 - Figure 11.23), and layer compositions (Figure 11.24 - Figure 11.29) follow in the Chapter.

Table II.5 | Organic/H₂O excess isotherm characteristics for RPLC columns.

Column	Particle size [μm]	Solvent B	V ₀ [mL]	Γ _B ^V max [μmol.g ⁻¹ .resin/m ²]	C _B ^L for Γ _B ^V = 0 [mol/L]	Γ _B ^V min [μmol.g ⁻¹ .resin/m ²]	Linear range
Luna C5	5	MeOH	2.99	1.00 x10 ⁻³ (2.04 mol/L – 8.25%)	5.82 (23.56%)	-5.70 x10 ⁻³ (19.63 mol/L – 79.41%)	4.94 – 17.30 mol/L (20 – 70%)
		EtOH	2.76	0.50 x10 ⁻³ (1.24 mol/L – 7.24%)	3.13 (18.28%)	-2.15 x10 ⁻³ (12.38 mol/L – 72.29%)	1.71 – 6.85 mol/L (10–40%)
Luna C8(2)	5	MeOH	2.87	0.57 x10 ⁻³ (1.28 mol/L – 5.18%)	3.47 (14.04%)	-7.47 x10 ⁻³ (19.43 mol/L – 80%)	4.94 – 17.30 mol/L (20–70%)
		EtOH	2.66	6.21 x10 ⁻⁷ (0 mol/L – 0.00%)	0.00 (0.00%)	-3.49 x10 ⁻³ (4.91 mol/L – 28.67%)	0.34 – 3.42 mol/L (2 – 20%)
Luna C8(2)	10	MeOH	3.00	2.64 x10 ⁻³ (4.46 mol/L – 18.04%)	14.68 (59.39%)	-3.09 x10 ⁻³ (21.83 mol/L – 88.30%)	14.83 – 22.25 (60–90%)
		EtOH	3.04	3.43 x10 ⁻³ (4.64 mol/L – 27.09%)	11.17 mol/L (65.22%)	-2.33 x10 ⁻³ (15.18 mol/L – 88.64%)	6.85 – 13.70 (40 – 80%)
Luna C8(2) PREP	10	MeOH	2.92	1.31 x10 ⁻³ (3.31 mol/L – 13.39%)	8.96 (36.25%)	-1.90 x10 ⁻³ (22.70 mol/L – 91.83%)	1.24 – 22.25 (5–90%)
		EtOH	2.97	3.82 x10 ⁻³ (4.68 mol/L – 27.33%)	13.12 (76.61%)	-1.60 x10 ⁻³ (15.77 mol/L – 92.08%)	6.85 – 15.41 (40–90%)
Luna C18(2)	5	MeOH	2.85	2.48 x10 ⁻³ (4.50 mol/L – 18.20%)	14.37 mol/L (58.13%)	-3.46 x10 ⁻³ (22.25 mol/L – 90%)	9.89 – 22.25 (40–90%)
		EtOH	2.75	4.10 x10 ⁻³ (4.85 mol/L – 28.32%)	15.10 mol/L (88.17%)	-0.89 x10 ⁻³ (16.42 mol/L – 95.90%)	6.85 – 16.27 (40–95%)
Luna C18(2)	10	MeOH	2.80	3.01 x10 ⁻³ (6.90 mol/L – 27.91%)	19.77 mol/L (79.98%)	-0.61 x10 ⁻³ (19.77 mol/L – 79.98%)	(12.36 – 19.77) (50–80%)
		EtOH	2.81	3.35 x10 ⁻³ (4.92 mol/L – 28.73%)	12.95 (75.62%)	-2.46 x10 ⁻³ (15.76 mol/L – 92.02%)	8.56 – 13.70 (50 – 80%)
Luna C18(2) PREP	10	MeOH	2.81	2.24 x10 ⁻³ (5.68 mol/L – 22.98%)	16.50 (66.75%)	-1.10 x10 ⁻³ (20.81 mol/L – 84.19%)	12.36 – 19.77 (50–80%)
		EtOH	2.87	2.79 x10 ⁻³ (3.76 mol/L – 21.95%)	15.12 (88.29%)	-3.18 x10 ⁻³ (10.07 mol/L – 58.80%)	8.56 – 11.99 (50–70%)
Luna Phenyl- Hexyl	5	MeOH	2.96	2.24 x10 ⁻³ (5.68 mol/L – 22.98%)	16.50 (9%)	-1.11 x10 ⁻³ (20.81mol/L – 84.19%)	12.36 – 19.77 (50 – 80%)
		EtOH	3.04	3.76 x10 ⁻³ (3.86 mol/L – 22.54%)	11.41 (66.62%)	-2.10 x10 ⁻³ mol/L (15.31mol/L – 89.40%)	8.56 – 11.99 (50 – 70%)
Luna Phenyl- Hexyl	10	MeOH	2.98	2.41 x10 ⁻³ (5.01 mol/L – 20.27%)	15.10 (61.09%)	-0.34 x10 ⁻³ (16.44 mol/L – 66.51%)	8.56 – 13.70 (50 – 80 %)
		EtOH	2.92	3.83 x10 ⁻³ (4.27 mol/L – 24.93%)	12.67 (73.98%)	-1.20 x10 ⁻³ (16.03 mol/L – 93.60%)	8.56 – 11.99 (50 – 70%)
Luna Phenyl- Hexyl PREP	10	MeOH	3.05	3.28 x10 ⁻³ (6.29 mol/L – 32.98%)	19.16 (77.51%)	-1.29 x10 ⁻³ (23.21 mol/L – 93.90%)	12.36 – 19.77 (50 – 80%)
		EtOH	3.05	4.35 x10 ⁻³ (4.66 mol/L – 27.21%)	13.10 (76.49%)	-1.19 x10 ⁻³ (16.07 mol/L – 93.83%)	8.5634 – 13.70 (50 – 80%)

Table II.6 | Surface properties of RPLC columns with an organic/H₂O adsorption interface layer.

Column	Particle size (μm)	Sol-vent B	m (x10 ³)	C _{ads} or b (μmol.g _{resin} /m ²)	τ (Å.g _{resin})	V _{ads} (mL)	V _B (mL)	V _{H₂O} (mL)	V _B + V _{H₂O}	V _B /V _{H₂O}	V _B /(V _B + V _{H₂O})	(V _B + V _{H₂O})/V _{ads}
Luna C5	5	MeOH	-0.43	2.29	1.20	0.048	0.10	0.37	0.47	0.27	0.21	9.79
		EtOH	-0.48	1.37	0.80	0.032	0.08	0.40	0.48	0.20	0.17	15.00
Luna C8(2)	5	MeOH	-0.50	1.46	0.77	0.031	0.06	0.44	0.50	0.14	0.12	16.13
		EtOH	-0.94	0.20	0.12	0.005	0.01	0.92	0.93	0.01	0.01	186.00
Luna C8(2)	10	MeOH	-0.42	6.19	3.25	0.130	0.25	0.17	0.42	1.47	0.60	3.23
		EtOH	-0.62	6.92	4.04	0.162	0.40	0.22	0.62	1.85	0.65	3.83
Luna C8(2) PREP	10	MeOH	-0.16	1.52	0.80	0.032	0.06	0.09	0.15	0.65	0.40	4.69
		EtOH	-0.59	7.68	4.48	0.179	0.45	0.14	0.59	3.24	0.76	3.30
Luna C18(2)	5	MeOH	-0.43	6.05	3.17	0.127	0.24	0.18	0.42	1.35	0.57	3.31
		EtOH	-0.48	7.27	4.24	0.170	0.42	0.06	0.48	7.32	0.88	2.82
Luna C18(2)	10	MeOH	-0.31	6.20	3.25	0.130	0.25	0.06	0.31	4.17	0.81	2.38
		EtOH	-0.51	6.54	3.82	0.153	0.38	0.13	0.51	2.94	0.75	3.33
Luna C18(2) PREP	10	MeOH	-0.34	5.57	2.92	0.117	0.22	0.11	0.33	2.05	0.67	2.82
		EtOH	-0.67	6.79	3.96	0.159	0.40	0.28	0.68	1.43	0.59	4.28
Luna Phenyl-Hexyl	5	MeOH	-0.34	5.57	2.92	0.117	0.27	0.04	0.31	7.38	0.87	2.65
		EtOH	-0.69	7.85	4.59	0.183	0.50	0.23	0.73	2.00	0.68	3.99
Luna Phenyl-Hexyl	10	MeOH	-0.31	4.67	2.45	0.098	0.27	0.04	0.31	7.50	0.87	3.16
		EtOH	-0.53	6.71	3.92	0.157	0.39	0.14	0.53	2.86	0.74	3.38
Luna Phenyl-Hexyl PREP	10	MeOH	-0.36	6.81	3.57	0.143	0.28	0.08	0.36	3.44	0.78	2.52
		EtOH	-0.60	7.85	4.58	0.183	0.46	0.14	0.60	3.28	0.77	3.28

In the 5- μm column data sets, the V_M reduces in the order of $\text{ACN} > \text{MeOH} > \text{EtOH}$ for the alkyl-bonded phases (C5, C8, C18), but the opposite is observed for the phenyl columns. For all 10- μm columns, the compared V_M values follow the order $\text{ACN} > \text{MeOH} \cong \text{EtOH}$, and the difference between non-PREP and PREP columns is minimal (maximum error of 4.5%). One interesting observation is that when MeOH or EtOH are used as the organic modifier solvent, the graph of the retention volume based on the MDM changes shape. Instead of the sudden decrease around 50%B (v/v) as seen with ACN, the V_r gradually decreases until an abrupt change is observed at higher concentrations of organic solvent (80%). In some cases, a minor increase is observed at around 20 or 40%B organic solvent content in the mobile phase.

Regarding surface capacity, using MeOH and EtOH as the organic solvent drastically reduces the surface adsorption capacity by a factor of between 2 and 5, compared to ACN. Interestingly, the value of C_{ads} is relatively constant amongst all 10- μm columns (PREP or non-PREP) when EtOH was used as the organic solvent (ca. 6.5-7.9 $\mu\text{mol.g}_{\text{resin}}/\text{m}^2$). Similar reductions can be observed in the surface hydrophobic heterogeneity. However, these variations are more evident for all columns. Interestingly, in all 10- μm columns, one of the Phenyl-Hexyl columns (PREP or non-PREP) shows the highest $V_B/V_{\text{H}_2\text{O}}$ value.

To understand these observed variations, it is essential to acknowledge the chemical properties of the different organic modifiers evaluated in these experiments. For example, MeOH and EtOH are polar protic solvents and, therefore, can form strong hydrogen bonds. Mixing these solvents with water creates complex hydrogen-bonded solvent networks where water molecules surround the alcohol molecules through hydrogen bonding from the R-OH group and can encage the hydrophobic methyl group by the hydrophobic effect[191, 192]. Due to this, the molecular mesh of these solvents is similar to a simple water network (Figure 11.2). This characteristic can explain (1) the lack of sudden decrease V_M values during the MDM method around 50%B as the distance between alcohol molecules is similar to that of pure water molecules and (2) the apparent higher hydrophobic heterogeneity ($V_B/V_{\text{H}_2\text{O}}$ values) of Phenyl-Hexyl where the ratio increases because π -OH interactions, not hydrophobic interactions, drive adsorption of MeOH and H_2O molecules at similar rate. These π -OH interactions are formed between the phenyl group in the stationary phase and the water and alcohol molecules in the mobile phase and in the adsorption layer/interface.

Water on its own forms a 3-D network of hydrogen-bonded H₂O molecules, when MeOH or EtOH are added into water, these organic hydroxyl solvents form clusters with surrounding water molecules[193]. As discussed before, in the presence of a silica surface, the water molecules form hydrogen bonds with the residual silanols on the surface; as the concentration of alcohol increases, the hydrophobic alkyl ligands on the surface adsorb the alkyl groups in the alcohol molecule while the opposite alcohol group forms hydrogen bonds with the surrounding hydrogen-bonded water network, forming multiple hydrophobic shells. After all ligands are saturated, the interface formed is a complex network of alcohol/water clusters. The presence of these clusters could explain the gradual reduction of V_M values in the MDM method and reduced surface adsorption capacity.

It has been found through MD simulations that water clusters are stable up to around 80% (v/v) of alcohol content in the aqueous mixture for both alcohols[194]. This observation could explain the rapid increase in the retention volume of the MDM experiments (See for example Figure 11.12).

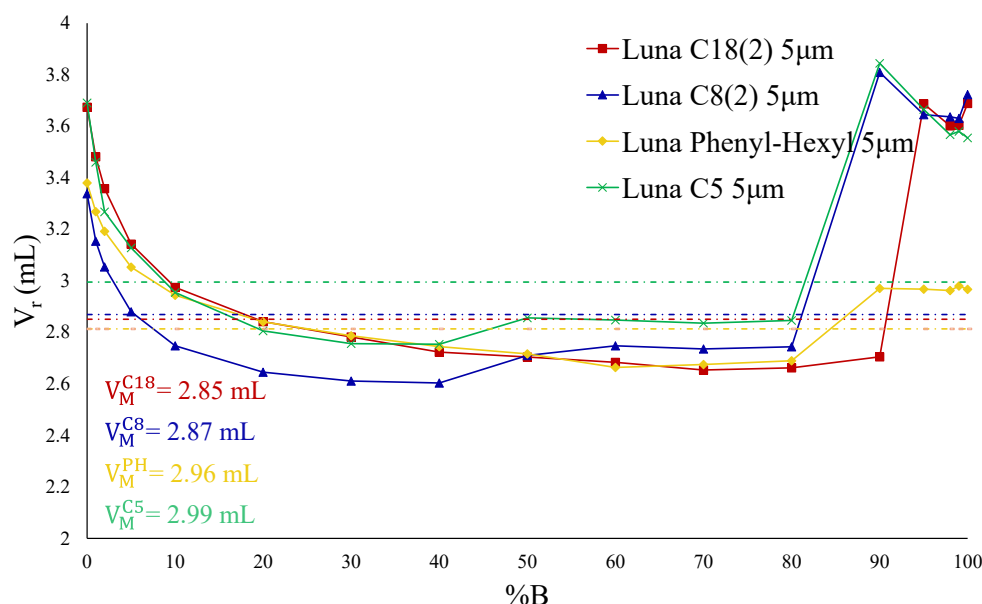


Figure 11.12 | Determination of V_M using the minor disturbance method in 5 μm columns using methanol (MeOH). The mobile phase comprised DI water (line A) and MeOH (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 1 mL/min. Column details can be found in Table 6.4.

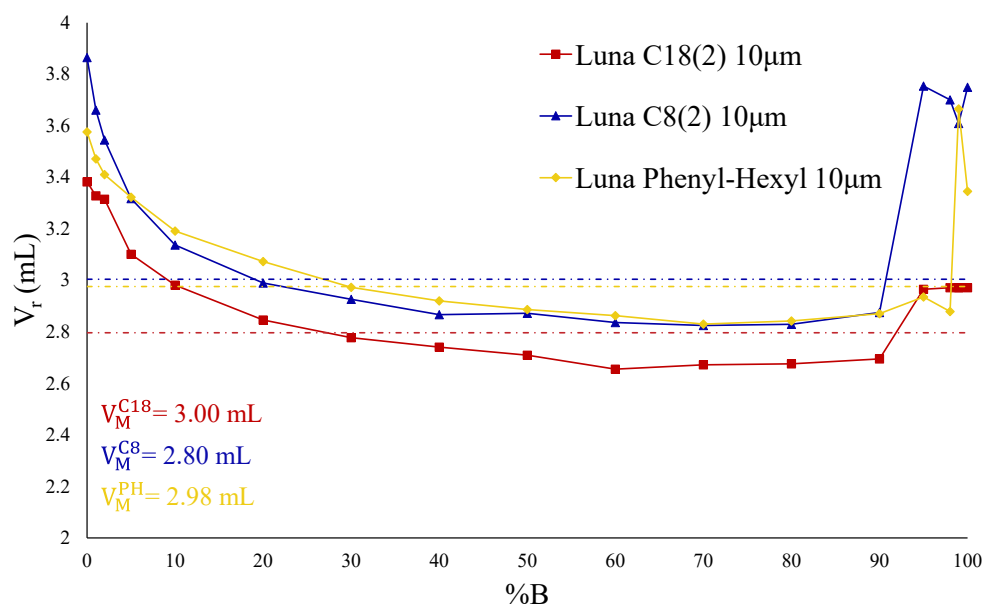


Figure 11.13 | Determination of V_M using the minor disturbance method in 10 μm columns using methanol (MeOH). The mobile phase comprised DI water (line A) and MeOH (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 0.5 mL/min. Column details can be found in Table 6.4.

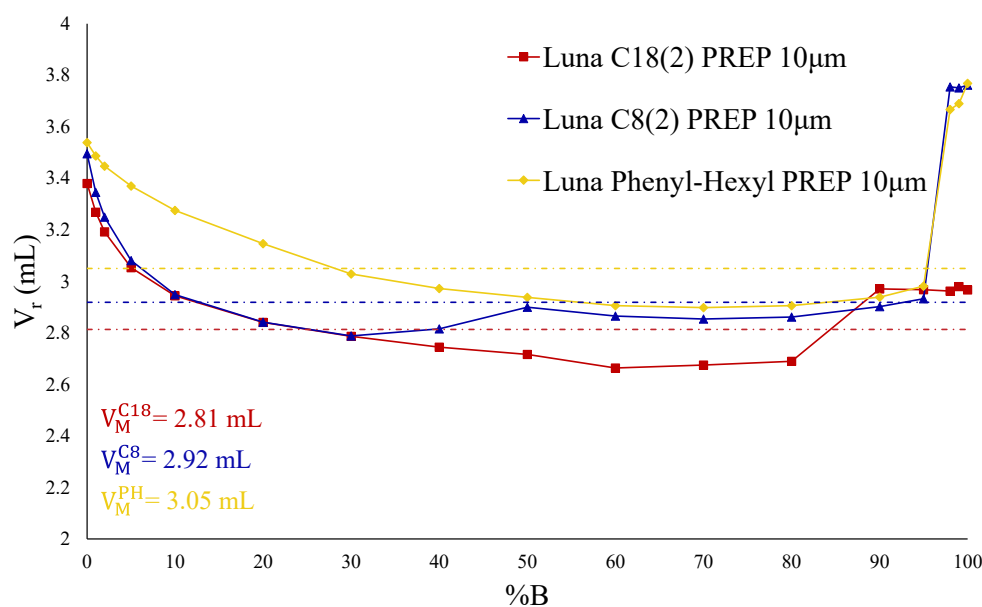


Figure 11.14 | Determination of V_M using the minor disturbance method in 10 μm PREP columns using methanol (MeOH). The mobile phase comprised DI water (line A) and MeOH (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 1 mL/min. Column details can be found in Table 6.4.

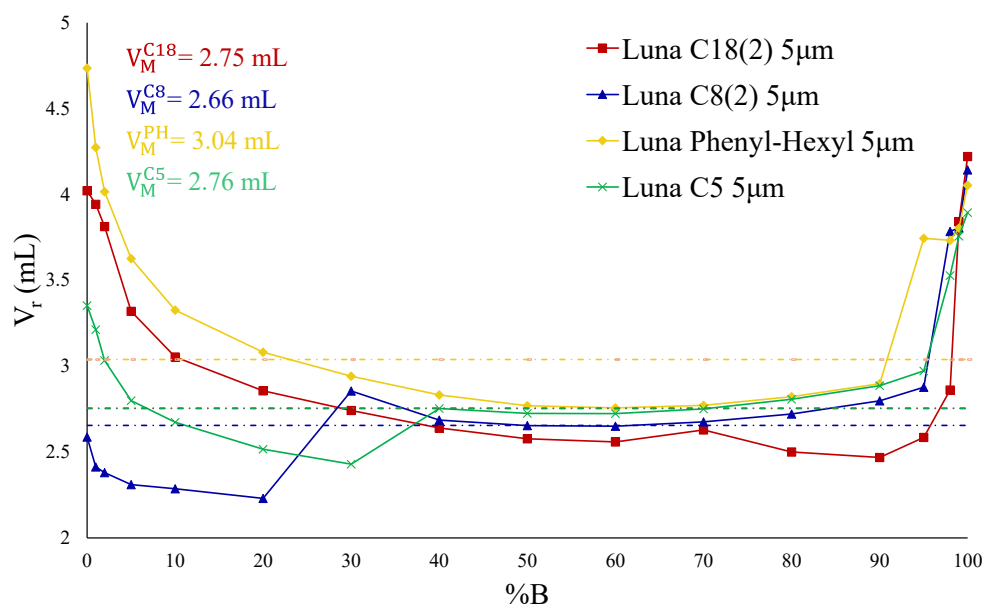


Figure 11.15 | Determination of V_M using the minor disturbance method in 5 μm columns using ethanol (EtOH). The mobile phase comprised DI water (line A) and EtOH (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 0.25 mL/min. Column details can be found in Table 6.4.

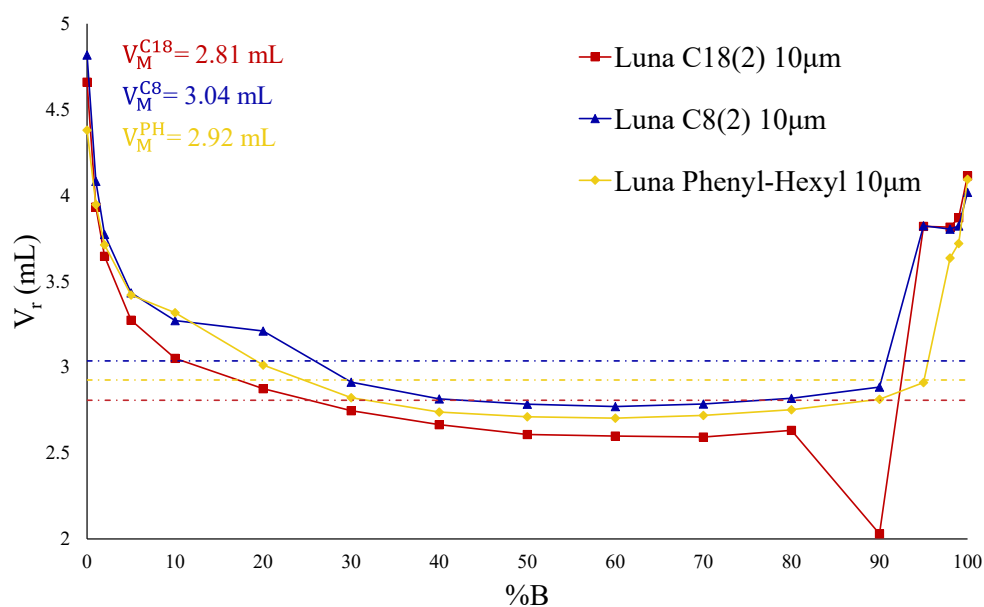


Figure 11.16 | Determination of V_M using the minor disturbance method in 10 μm columns using ethanol (EtOH). The mobile phase comprised DI water (line A) and EtOH (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 0.5 mL/min. Column details can be found in Table 6.4.

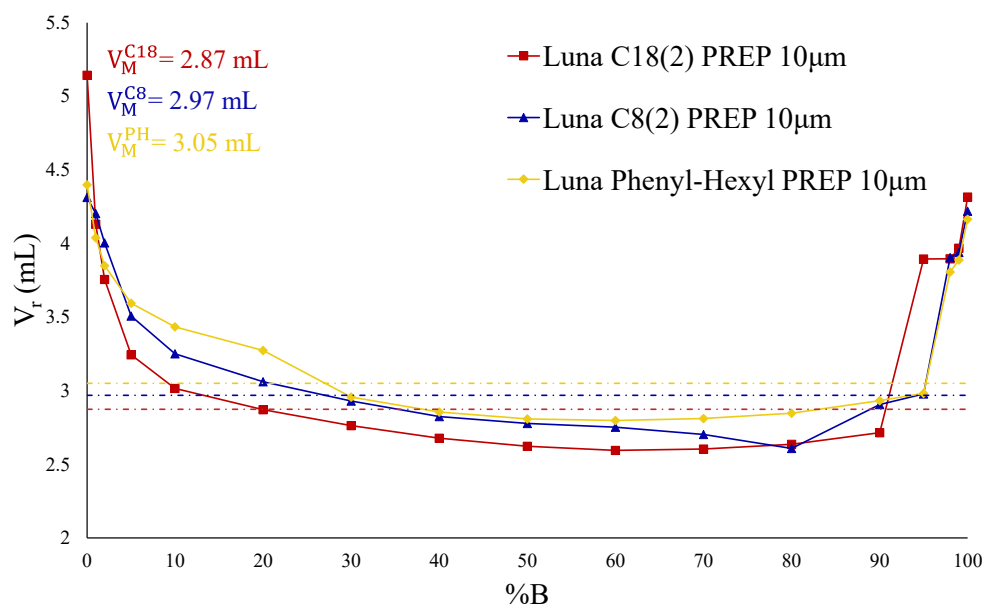


Figure 11.17 | Determination of V_M using the minor disturbance method in 10 μm PREP columns using ethanol (EtOH). The mobile phase comprised DI water (line A) and EtOH (line B). The disturbance was produced by injecting 1-5 μL of pure water or acetonitrile. All points have a coefficient of variation of less than 2%. Flowrate = 0.5 mL/min. Column details can be found in Table 6.4.

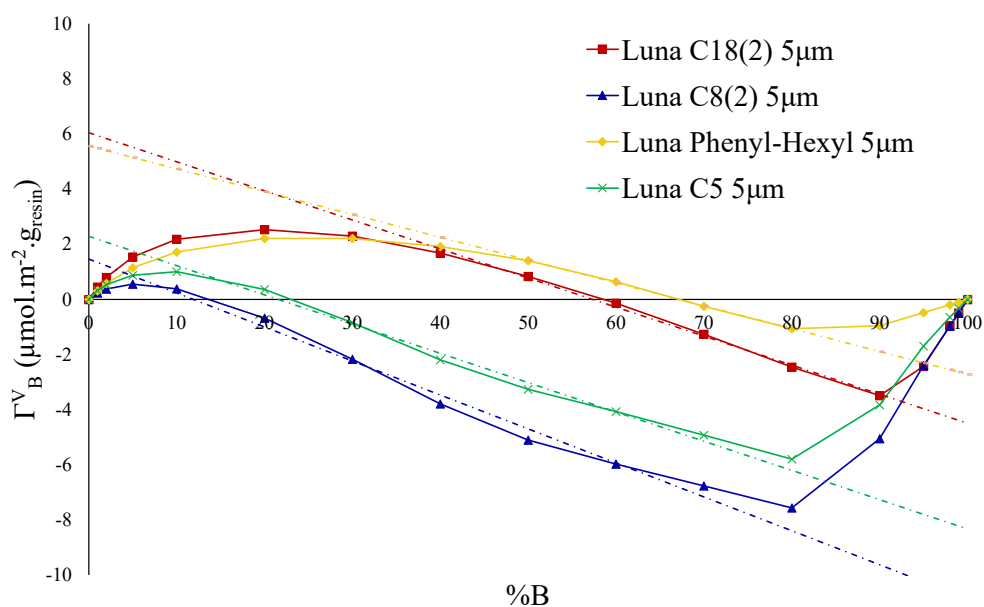


Figure 11.18 | Excess adsorption isotherms of methanol (MeOH) in water on 5 μm columns. The mobile phase comprised DI water (line A) and MeOH (line B). Disturbance method in Figure 11.12. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

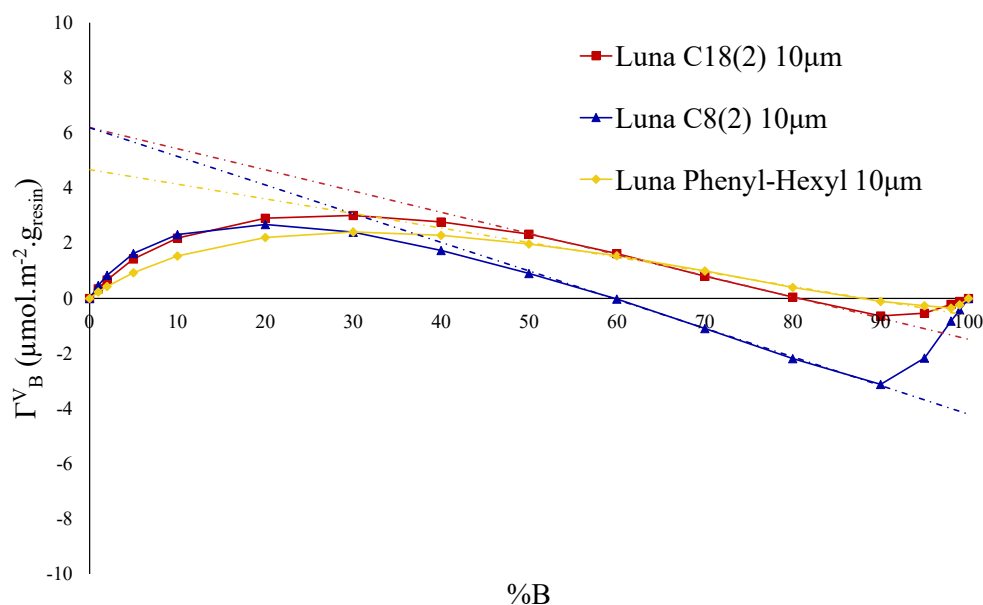


Figure 11.19 | Excess adsorption isotherms of methanol (MeOH) in water on 10 µm columns. The mobile phase comprised DI water (line A) and MeOH (line B). Disturbance method in Figure 11.3. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

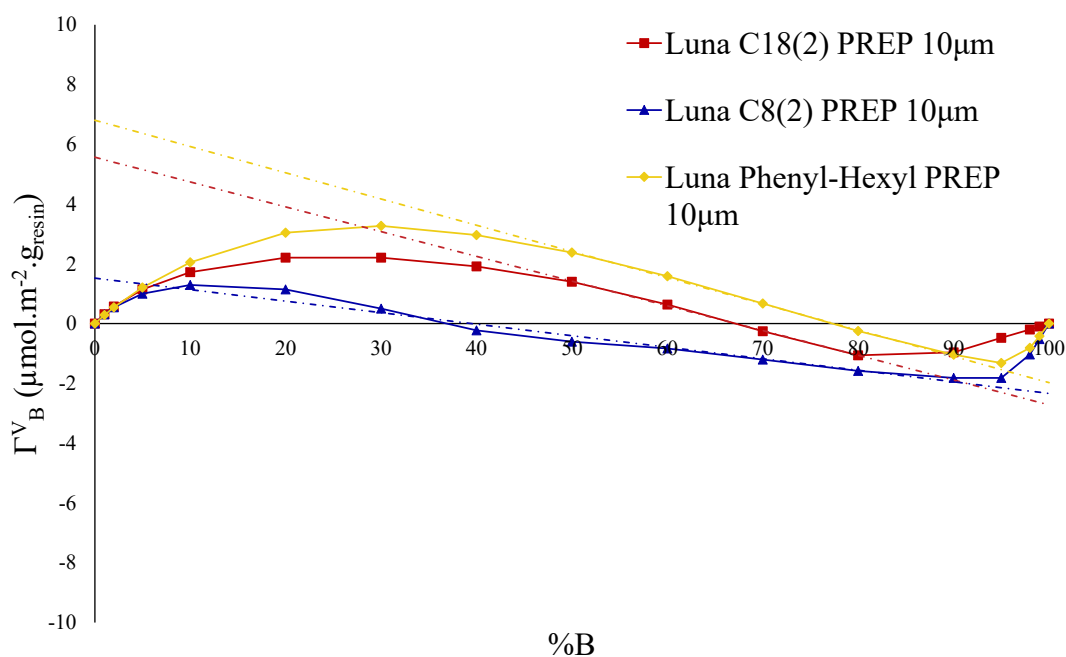


Figure 11.20 | Excess adsorption isotherms of methanol (MeOH) in water on 10 µm PREP columns. The mobile phase comprised DI water (line A) and MeOH (line B). Disturbance method in Figure 11.4. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

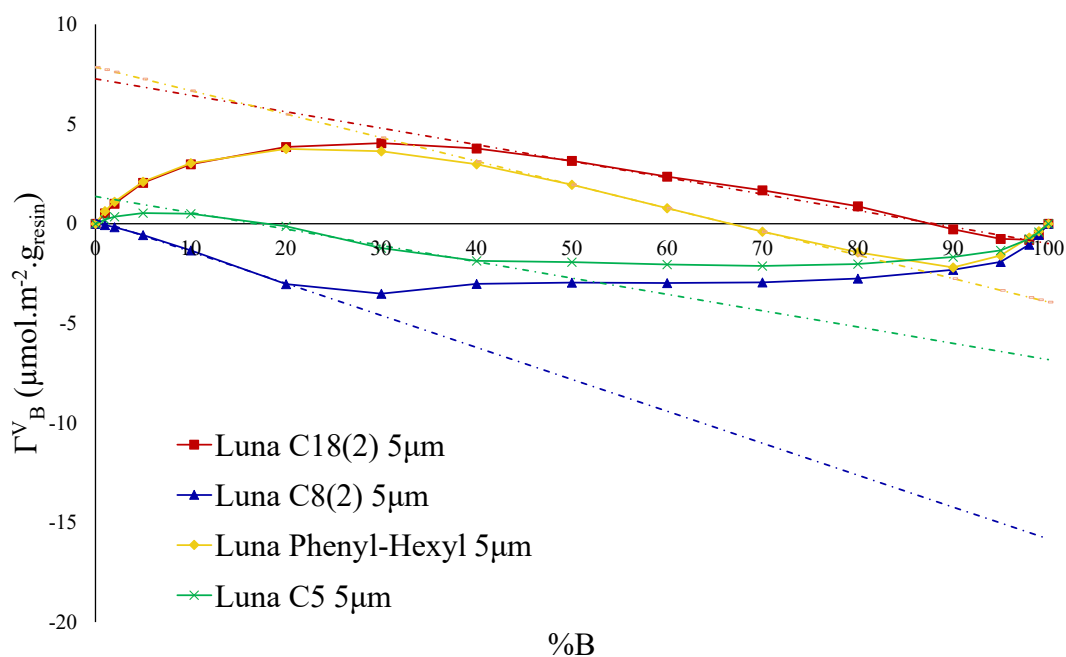


Figure 11.21 | Excess adsorption isotherms of ethanol (EtOH) in water on 5 µm columns. The mobile phase comprised DI water (line A) and EtOH (line B). Disturbance method in Figure 11.15. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

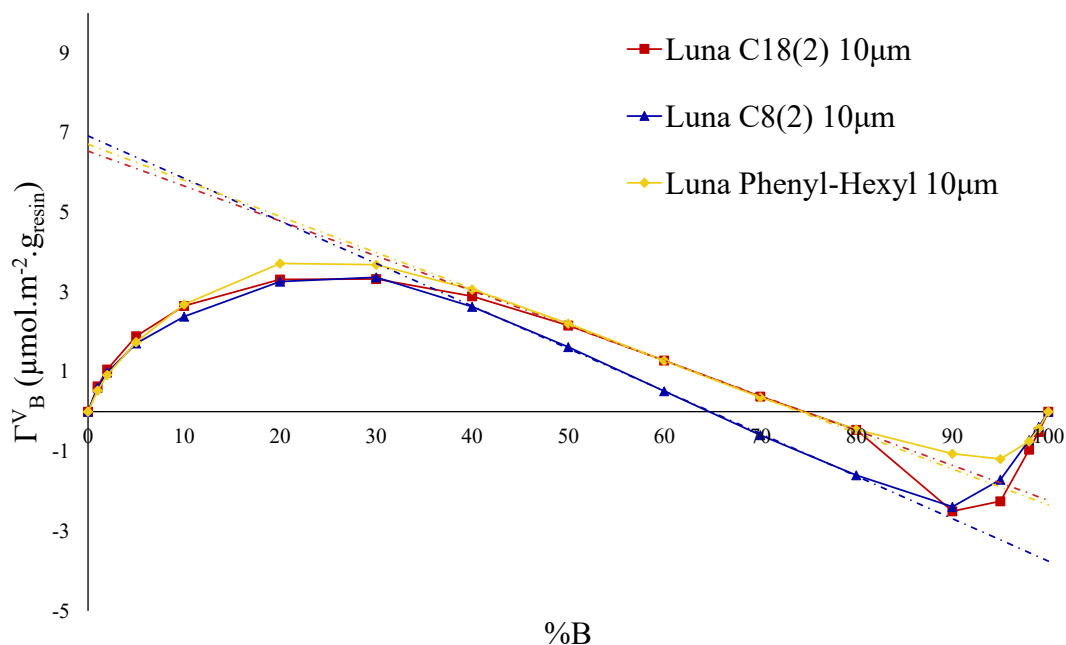


Figure 11.22 | Excess adsorption isotherms of ethanol (EtOH) in water on 10 µm columns. The mobile phase comprised DI water (line A) and EtOH (line B). Disturbance method in Figure 11.16. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

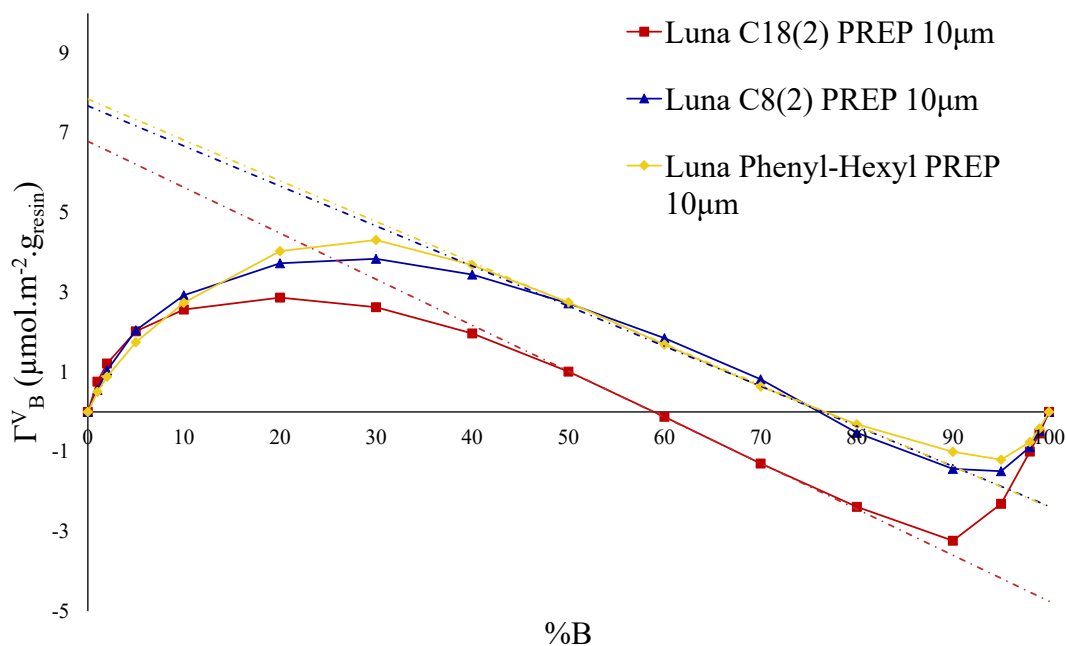


Figure 11.23 | Excess adsorption isotherms of ethanol (EtOH) in water on 10 μm columns. The mobile phase comprised DI water (line A) and EtOH (line B). Disturbance method in Figure 11.17. Column details can be found in Table 6.4. The dotted lines represent the extrapolated linear region of the isotherm.

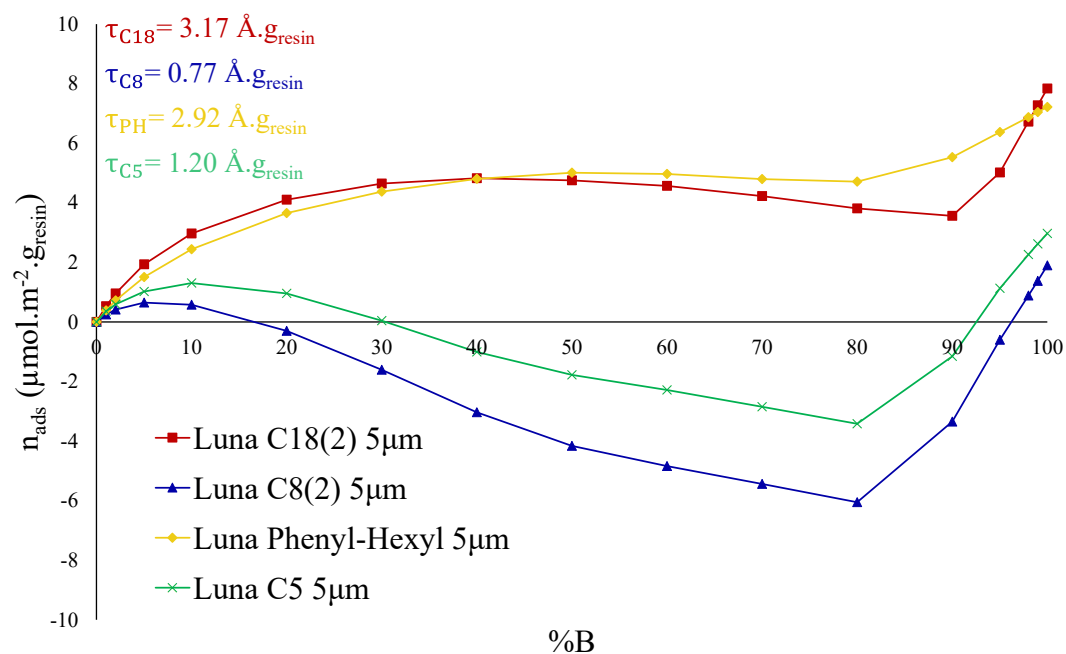


Figure 11.24 | Total water/methanol eluent mass adsorbed onto the 5 μm columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

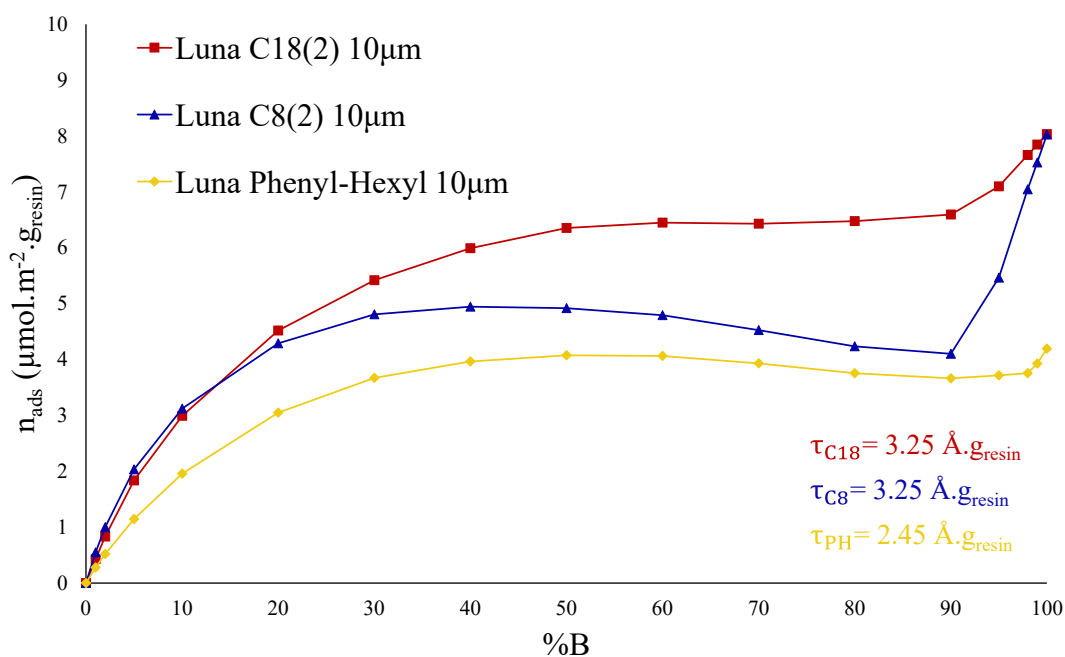


Figure 11.25 | Total water/methanol eluent mass adsorbed onto the 10 μm columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

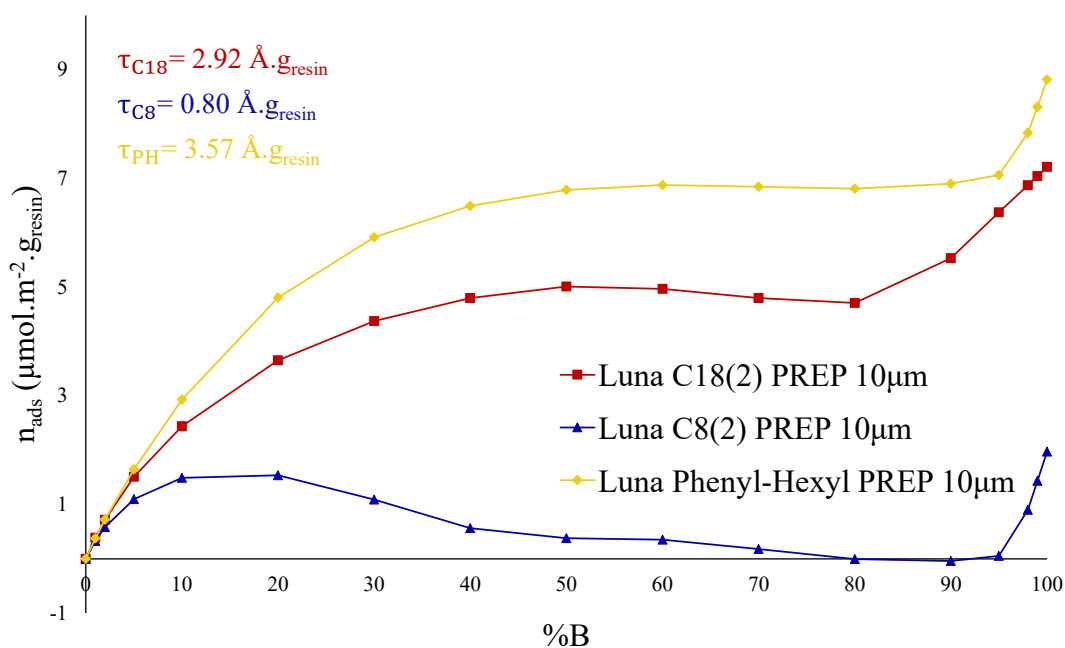


Figure 11.26 | Total water/methanol eluent mass adsorbed onto the 10 μm PREP columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

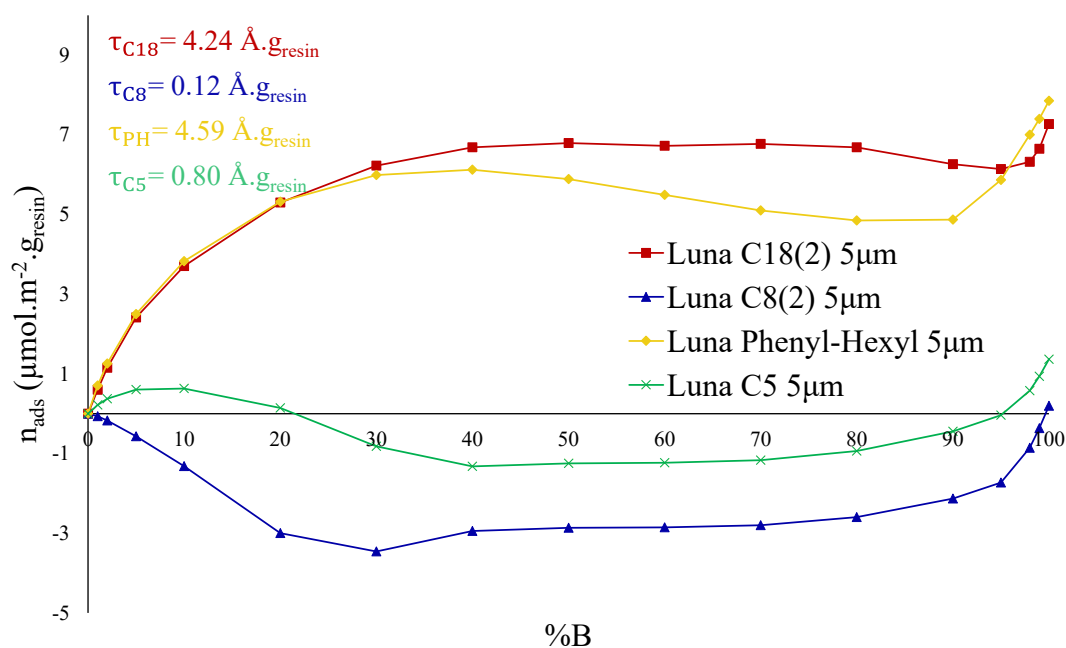


Figure 11.27 | Total water/ethanol eluent mass adsorbed onto the 5 μm columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

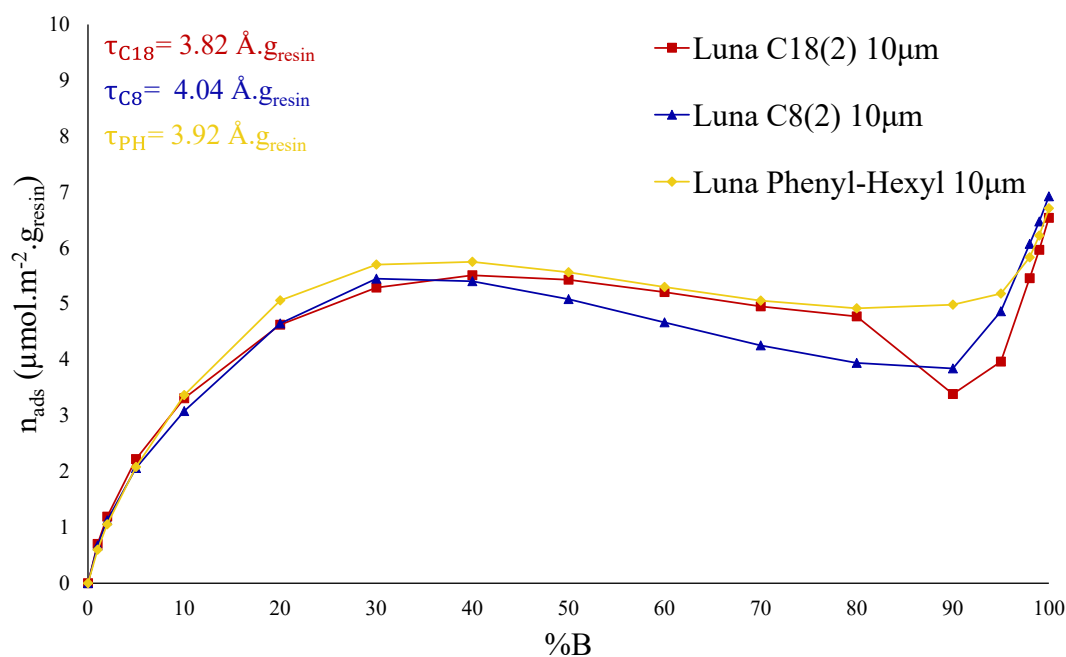


Figure 11.28 | Total water/ethanol eluent mass adsorbed onto the 10 μm columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

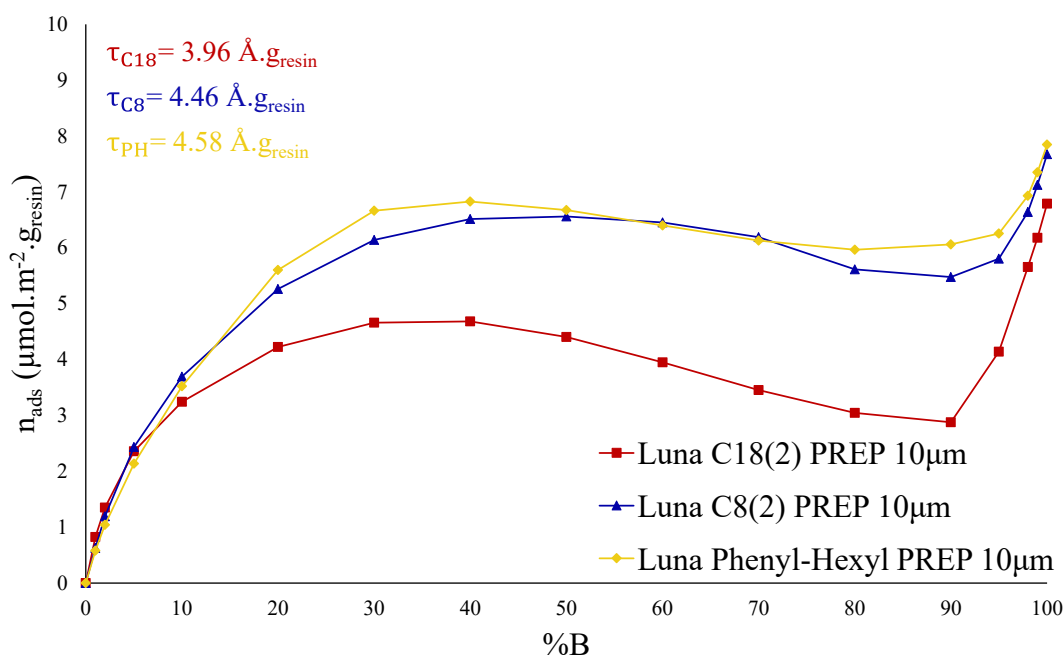


Figure 11.29 | Total water/ethanol eluent mass adsorbed onto the 10 μm PREP columns. Eluent adsorbed within interfaces of theoretical finite thickness τ .

11.4 Effect of column length

By comparing results from columns C18, C8 and Phenyl-Hexyl in Sections 9.1 and Figure 11.3, the variation in the estimation of V_M through excess isotherms can be assessed by comparing the porosity calculated by dividing the calculated V_M by the column volume of the column. The porosity for all 150-mm columns from Chapter 9 had an average porosity of 0.68, whereas the 250-mm columns used in this Chapter had a marginal error of +0.1. This observation indicates that calculating V_M values through excess isotherms is robust for different-sized columns, as the result is a cumulative calculation of all interactions between the mobile and stationary phases. This same observation can also explain the very clear overlap between analytical and semi-PREP SMART experiments in Figures 10.12-14.

11.5 Effect of ion-pairing modifier

Although ion-pairing agents/modifiers were not included due to the scope of this Chapter, some information can be obtained from the TFA-added experiments in Section 9.1 and those without TFA in Section 8.1. It is clearly noticeable that the V_M values for the same column are reduced by 26% when the TFA is present in the mobile phase. I believe that this is due to the caging of surface silanol groups by TFA groups, which creates and pushes the surrounding water/acetonitrile network, reducing the available volume.

If this hypothesis is true, it can also be used to complete the fundamental understanding of the role of ion-pairing agents like TFA in RPLC separation/adsorption. Currently, it is believed that the peak narrowing effect when ion-pairing agents are present is due to the lack of electrostatic interactions due to the caging of the silanol groups; however, I would argue that the narrowing effect is also due to a reduced available adsorption layer/volume. Thus, a lower extent of Eddy diffusion would be expected. All the experiments in this Chapter must be repeated with added ion-pairing agents in order to provide more insight and draw clearer conclusions.

11.6 Use of excess isotherm as a column characterisation system

Excess isotherms have the potential of being used as a column characterisation system as it provides:

1. Description of the RPLC column at operational (wet) conditions and without unpacking the column (in-situ measurement).
2. Characterises the adsorption layer that analytes partition during the chromatographic run.
3. Comparison of the complex interaction and formation of the available adsorption volume of a column with different organic modifiers/solvents.

Ideally, excess isotherms would have more chromatographic value if they can be used as CSS based on the fundamentals of RPLC. This would mean that by analysing the adsorption of the mobile phase onto the packing material, a rational prediction can be made regarding the interactions between the analyte and the stationary phase.

As demonstrated in this Chapter, it is possible to start understanding the complex interactions required to form the interface layer. However, further experiments with relevant probes, such

as the Krokhin standards, are required to translate this into knowledge into adsorption prediction.

11.7 Conclusions

This Chapter explores the intricate dynamics of liquid-solid surface adsorption in reversed-phase liquid chromatography (RPLC). By focusing on how surface interactions between the mobile and stationary phases drive the retention and separation processes, I have gained valuable insights into the behaviour of various chromatographic systems. While often governed by hydrophobic forces, these interactions are also influenced by more subtle physicochemical phenomena specific to the column's surface chemistry and structure.

The results presented in this Chapter clearly demonstrate the critical role of the stationary phase's surface properties, particularly in relation to its ligand structure, particle size, and overall hydrophobicity. One of the key findings is that the length and nature of surface ligands significantly affect the adsorption capacity of the column. Columns with longer alkyl chains, such as C18, exhibit stronger hydrophobic interactions, resulting in higher retention volumes compared to shorter chains like C5 or C8.

Using excess adsorption isotherms, I have quantitatively measured how organic and aqueous solvents are distributed on the surface of the RPLC stationary phase. These results show that the surface hydrophobicity, characterised by the amount of ACN/H₂O molecules adsorbed on the RPLC stationary phase, varies depending on the stationary phase chemical composition and the bulk liquid mobile phase composition. Notably, the surface adsorption capacity for organic solvents is greater in C18 and Phenyl-Hexyl columns, indicating a higher degree of hydrophobic interaction with the mobile phase. This information is critical for understanding how different solvents will behave in separation processes and provides a method for optimising the choice of stationary phase for specific separations.

Additionally, the particle size of the packing material is shown to have a measurable impact on the chromatographic process. Larger particles tend to create less efficient packing, increasing the hold-up volume and altering the surface adsorption behaviour. This behaviour was particularly evident in the comparisons made between 5 µm and 10 µm particle sizes, where significant

differences in surface adsorption capacity and solvent interactions were observed. The larger particles allowed for more pronounced solvent interactions within the stationary phase, which could be advantageous or disadvantageous depending on the separation goals.

The study also highlighted the utility of the minor disturbance method (MDM) for characterising column performance. By injecting minor disturbances of pure solvents (e.g., acetonitrile or water), I was able to observe how the surface of the column's stationary phase interacts with the mobile phase under different mobile phase compositions. The technique revealed how solvent molecules progressively fill the available adsorption sites, forming complex interface layers that influence the overall solute retention mechanism. This approach provides an invaluable tool for fine-tuning chromatographic conditions in both analytical and preparative RPLC systems.

Another important observation relates to forming hydrophobic and hydrophilic "patches" within the interface layer, particularly in mixed solvent systems. These patches are composed of clusters of organic and water molecules, which arrange themselves depending on the hydrophobicity of the ligand surface. The dynamic exchange between these patches as the mobile phase composition changes is a key factor in achieving precise separations, especially for complex mixtures where multiple analytes exhibit different affinities for the stationary phase.

In summary, this Chapter has underscored the importance of understanding the physicochemical properties of the stationary phase and their role in liquid-solid adsorption in RPLC. Excess adsorption isotherms and methods like MDM have allowed for a detailed characterisation of these interactions, offering practical guidance for improving chromatographic performance. These findings suggest that continued exploration of surface interactions, mainly through advanced modelling and experimental techniques, will be crucial for developing more efficient and selective chromatographic methods.

The conclusions drawn from this study have significant implications for both analytical and industrial chromatography. The ability to accurately predict and control adsorption behaviour on various chromatography column types opens new avenues for optimising solute separation efficiency, reducing solvent consumption, and enhancing the scalability of RPLC processes. As the industry moves toward increasingly complex biomolecular separations, the insights gained from this work will provide a foundation for further innovations in chromatographic separation and purification science.

Future research should focus on expanding this methodology to a broader range of stationary phases and mobile phase compositions, particularly in the context of large biomolecules and pharmaceuticals. Moreover, continued efforts to refine the characterisation of surface interactions using both experimental and theoretical approaches will be essential for fully realising the potential of these findings in practical applications. This Chapter serves as a stepping stone toward a deeper understanding of liquid-solid interactions in chromatography and sets the stage for future advancements in the field.

Chapter 12

Final remarks and future work

12.1 Final remarks

The research presented in this Thesis provides a comprehensive exploration of peptide adsorption mechanisms in RPLC, contributing significantly to our theoretical understanding and practical industrial applications. The investigation starts by focusing on how critical chromatographic parameters influence peptide retention and separation efficiency, including the effects of solvent composition, column ligand properties, peptide hydrophobicity, and mobile phase interactions. Comprehensive and detailed experimental studies of the fundamentals of RPLC have enabled the development of optimised methods for peptide purification that address several current challenges in the field, thereby demonstrating the immediate impact of the research on practical applications.

One of the most impactful contributions of this work is the deepened understanding of peptide adsorptive interactions with chromatographic stationary phases, especially the role of hydrophobicity, peptide structure and mobile phase composition in determining retention behaviour. The detailed characterisation of various RPLC columns and their associated adsorbate adsorption mechanisms has allowed the identification of key factors that influence peptide separation, such as ligand density, column pore size, and surface chemistry. This level of understanding not only improves peptide separation but also has implications for other fields, demonstrating the interdisciplinary nature and broad applicability of the research contained within this Thesis.

The creation and application of the SMART algorithm for mobile phase composition represents a significant advancement in chromatographic methodologies. By systematically optimising elution profiles and column selection, this algorithm has demonstrated its ability to improve the resolution of peptide mixtures, minimise solvent consumption, and reduce overall process time. These achievements are particularly important for industrial applications, where efficiency, scalability, and cost-effectiveness are key concerns. The successful implementation of the SMART algorithm in separating glucagon and other peptides not only demonstrates its potential

for broader application in the purification of more complex biopharmaceutical products but also opens up exciting possibilities for the future of liquid chromatography, suggesting potential future applications in other fields and industries where gradient elution chromatographic approaches are commonly used.

Environmental sustainability has also been an important consideration in this research. Despite being a highly effective method for peptide purification, RPLC is traditionally associated with high solvent usage, contributing to both environmental impact and operational costs. This Thesis has critically examined the environmental footprint of RPLC processes and proposed solutions for reducing solvent consumption and waste production. The research underscores the importance of adopting green chemistry principles in chromatography, such as using environmentally friendly solvents, improving solvent recovery systems, and optimising chromatographic methods to require less organic solvent per run. These considerations are vital as the pharmaceutical industry seeks to balance high purification standards with increasing sustainability and resource efficiency demands.

Moreover, this Thesis has shed light on some of the limitations of current RPLC techniques, such as the challenges of separating closely related peptide impurities. These impurities, often generated during peptide synthesis, can be difficult to resolve using traditional chromatographic methods due to their similar hydrophobic properties. By exploring alternative chromatographic strategies, including column selection based on specific peptide characteristics and multi-isocratic elution profiles, this research has provided practical solutions to enhance the resolution of these difficult-to-separate impurities within reaction mixtures.

In conclusion, this Thesis has significantly contributed to real-world peptide purification by advancing the theoretical understanding of and practical methodologies for RPLC. The insights gained from this research could have immediate applications in the pharmaceutical industry, where peptide-based drugs are becoming increasingly important. Understanding the fundamentals of RPLC adsorption provides a more efficient approach to addressing critical challenges in peptide separation, developing more efficient purification methods, and proposing solutions to reduce environmental impact; this work lays the foundation for future innovations in chromatography. Looking forward, the techniques and findings presented in this Thesis have the potential to be further refined and expanded, not only for peptide purification but for a wide range of biomolecules. The work here opens new possibilities for improving chromatographic processes'

efficiency, sustainability, and scalability, ensuring they meet the evolving needs of academic research and industrial applications in the biopharmaceutical sector.

12.2 Future Work

Building on the work presented here, several avenues for future research can be explored:

- **Extension of SMART Algorithm to Complex Peptides:** While the SMART algorithm has demonstrated its effectiveness with relatively simple peptide mixtures and glucagon, its application to more complex and larger peptides remains to be tested. Further studies could expand the algorithm's adaptability to a broader range of peptide structures, particularly those with intricate folding patterns, larger hydrophobic domains, and more complex close-eluting impurities.
- **Improving Stationary Phase Design:** The role of stationary phase ligand density and ligand type in adsorption efficiency opens another research direction. Future work could explore the synthesis of novel stationary phases tailored specifically for peptides, possibly incorporating more selective ligands to enhance peptide-specific interactions and improve separation performance.
- **Exploration of Alternative RPLC Packing Materials:** One of the exciting areas for future research lies in exploring novel materials for RPLC packing. Current stationary phases, typically based on silica with hydrophobic ligands (such as C18), have limitations in separating closely related impurities without using excessive amounts of resources. Future work could explore materials with different surface chemistries, porosities, and mechanical properties. Alternative materials like polymer-based stationary phases, metal-organic frameworks (MOFs), or mixed-mode resins may offer improved selectivity and stability. Research into these materials could lead to enhanced separation of complex peptides and greater adaptability for industrial applications. I did initial work on exploring using polyethylene fibres as a reference hydrophobic material for RPLC separations (See Appendix V)
- **Green Chemistry and Sustainable Practices:** As highlighted in this Thesis, solvent usage in peptide purification is an area ripe for innovation. Future research should focus on the development of greener solvents and solvent-recycling techniques. Reducing the environmental impact of solvent disposal could align peptide purification methods with global sustainability goals while still maintaining or improving purification efficacy.
- **Industrial Process Scale-Up:** While this Thesis has demonstrated the scalability of peptide purification methods, further work is necessary to test these methods under real industrial conditions.

Overall, the groundwork laid in this research opens numerous possibilities for future advancements in peptide purification, particularly in enhancing separation efficiency, reducing environmental impact, and increasing scalability. These future directions aim to push the boundaries of what is possible in peptide chemistry, ensuring that the methods developed here remain relevant and beneficial for the evolving needs of the biotechnology and pharmaceutical industries.

Chapter 13

References

1. Kalidas, C. and M.V. Sangaranarayanan, *Peptides*, in *Biophysical Chemistry*, C. Kalidas and M.V. Sangaranarayanan, Editors. 2023, Springer Nature Switzerland: Cham. p. 129-141.
2. Sachdeva, S., *Peptides as 'Drugs': The Journey so Far*. International Journal of Peptide Research and Therapeutics, 2016. **23**: p. 49-60.
3. Rossino, G., et al., *Peptides as Therapeutic Agents: Challenges and Opportunities in the Green Transition Era*. Molecules, 2023. **28**: p. 7165.
4. Bottens, R.A. and T. Yamada, *Cell-Penetrating Peptides (CPPs) as Therapeutic and Diagnostic Agents for Cancer*. Cancers, 2022. **14**: p. 5546.
5. Xia, X., et al., *Tissue-Specific Delivery of Oligonucleotides*. 2019, Springer New York. p. 17-50.
6. Vall, H., P. Patel, and M. Parmar, *Teriparatide*, in *StatPearls*. 2024, StatPearls Publishing: Treasure Island (FL).
7. Liang, J., et al., *Cortistatin, a novel cardiovascular protective peptide*. Cardiovascular Diagnosis and Therapy, 2018. **9**: p. 394-399.
8. Precedence Research. *Peptide Synthesis Market Size, Share, and Trends 2024 to 2033*. July 2024; Available from: <https://www.precedenceresearch.com/peptide-synthesis-market>.
9. Shane Madden, T.K. *The power of peptide therapeutics: Advancements and methods in peptide drug development*. 2024 [cited 2024 18/09/2024]; Available from: <https://www.sterlingpharmasolutions.com/articles/the-power-of-peptide-therapeutics/>.
10. Sharma, K., et al., *Peptide-based drug discovery: Current status and recent advances*. Drug Discovery Today, 2023. **28**: p. 103464.
11. Pereira, A.J., et al., *Peptide-based therapeutics: challenges and solutions*. Medicinal Chemistry Research, 2024. **33**: p. 1275-1280.
12. Blankenship, K. *The top 20 drugs by global sales in 2019*. 2020; Available from: <https://www.fiercepharma.com/special-report/top-20-drugs-by-global-sales-2019>.
13. Rampell, C. and D.L. Estes, *The many ways new weight loss drugs could transform the economy*. 2024.
14. Jillian, F.-S. and D. Ochoa-Bravo, *Generic Victoza (liraglutide) GLP-1 now available in the U.S.* 2024.
15. Al Musaimi, O. and D.s.M.M. Jaradat, *Advances in Therapeutic Peptides Separation and Purification*. Separations, 2024. **11**: p. 233.
16. Mant, C.T., et al., *HPLC Analysis and Purification of Peptides*, in *Peptide Characterization and Application Protocols*, G.B. Fields, Editor. 2007, Humana Press: Totowa, NJ. p. 3-55.
17. Fosgerau, K. and T. Hoffmann, *Peptide therapeutics: current status and future directions*. Drug Discovery Today, 2015. **20**: p. 122-8.
18. Mallela, K., *Pharmaceutical biotechnology - concepts and applications*. Human Genomics, 2010. **4**: p. 218.

19. EMA, *Guideline on the Development and Manufacture of Synthetic Peptides*, C.f.M.P.f.H. Use, Editor. 2023.
20. USP, *USP Chromatographic Column Application*, United States Pharmacopeia, Editor. 2024.
21. FDA, *ANDAs for Certain Highly Purified Synthetic Peptide Drug Products That Refer to Listed Drugs of rDNA Origin (Guidance for Industry)*. 2021.
22. FDA, *Definition of the Term "Biological Product"*, U.S.F.a.D. Administration, Editor. 2018. p. 63817-63824.
23. Williams, L., *Macromolecular Structure*. Gatech.edu, 2021.
24. Choudhuri, S., *Fundamentals of Genes and Genomes*, in *Bioinformatics for Beginners*, S. Choudhuri, Editor. 2014, Academic Press: Oxford. p. 1-25.
25. Görbitz, C.H., *Peptide structures*. Current Opinion in Solid State and Materials Science, 2002. **6**: p. 109-116.
26. Nowick, J.S., *Exploring beta-sheet structure and interactions with chemical model systems*. Accounts of Chemical Research, 2008. **41**: p. 1319-30.
27. A.J., d.J. and H. Yin, *5.14 - Supramolecular Membrane Chemistry*. 2017: p. 311-328.
28. Eisenberg, D., et al., *Analysis of membrane and surface protein sequences with the hydrophobic moment plot*. J Mol Biol, 1984. **179**: p. 125-42.
29. Al Musaimi, O., O.M.M. Valenzo, and D.R. Williams, *Prediction of peptides retention behavior in reversed-phase liquid chromatography based on their hydrophobicity*. J Sep Sci, 2023. **46**: p. e2200743.
30. Al Musaimi, O., O.M. Mercado-Valenzo, and D.R. Williams, *Factors Influencing the Prediction Accuracy of Model Peptides in Reversed-Phase Liquid Chromatography*. Separations, 2023. **10**.
31. Akbarian, M., et al., *Bioactive Peptides: Synthesis, Sources, Applications, and Proposed Mechanisms of Action*. International Journal of Molecular Sciences, 2022. **23**.
32. Craik, D.J., et al., *The future of peptide-based drugs*. Chem Biol Drug Des, 2013. **81**: p. 136-47.
33. Komoda, T. and T. Matsunaga, *Chapter 4 - Metabolic Pathways in the Human Body*, in *Biochemistry for Medical Professionals*, T. Komoda and T. Matsunaga, Editors. 2015, Academic Press: Boston. p. 25-63.
34. Chambers, R.S., *High-throughput antibody production*. Curr Opin Chem Biol, 2005. **9**: p. 46-50.
35. Friedberg, F., T. Winnick, and D.M. Greenberg, *Peptide synthesis in vivo*. J Biol Chem, 1947. **169**: p. 763.
36. Gruenewald, S., et al., *In vivo production of artificial nonribosomal peptide products in the heterologous host Escherichia coli*. Applied and Environmental Microbiology, 2004. **70**: p. 3282-91.
37. Sharma, A., et al., *Liquid-Phase Peptide Synthesis (LPPS): A Third Wave for the Preparation of Peptides*. Chem Rev, 2022. **122**: p. 13516-13546.
38. Iwaoka, M. and S. Shimodaira, *Synthesis and catalytic functions of selenopeptides*, in *Organochalcogen Compounds*, E.J. Lenardão, et al., Editors. 2022, Elsevier. p. 195-218.
39. Petrou, C. and Y. Sarigiannis, *Peptide synthesis*, in *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*, S. Koutsopoulos, Editor. 2018, Woodhead Publishing. p. 1-21.
40. NIH, *MAINTENANCE AND OPERATION OF SYNTHETIC PEPTIDE FACILITY*, NIH, Editor. 1994, NIH Guide.

41. D'Hondt, M., et al., *Related impurities in peptide medicines*. J Pharm Biomed Anal, 2014. **101**: p. 2-30.
42. Zhuravlev, L.T., *Surface characterization of amorphous silica—a review of work from the former USSR*. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 1993. **74**: p. 71-90.
43. Borges, E.M., *Silica, hybrid silica, hydride silica and non-silica stationary phases for liquid chromatography*. J Chromatogr Sci, 2015. **53**: p. 580-97.
44. Sandoval, J.E. and J.J. Pesek, *Synthesis and characterization of a hydride-modified porous silica material as an intermediate in the preparation of chemically bonded chromatographic stationary phases*. Analytical Chemistry, 2002. **61**: p. 2067-2075.
45. Thommes, M., et al., *Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC Technical Report)*. Pure and Applied Chemistry, 2015. **87**: p. 1051-1069.
46. Nawrocki, J. and B. Buszewski, *Influence of silica surface chemistry and structure on the properties, structure and coverage of alkyl-bonded phases for high-performance liquid chromatography*. Journal of Chromatography A, 1988. **449**: p. 1-24.
47. Nawrocki, J., *Silica surface controversies, strong adsorption sites, their blockage and removal. Part II*. Chromatographia, 1991. **31**: p. 193-205.
48. Inkson, B.J., *2 - Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) for materials characterization*. 2016: p. 17-43.
49. Jaćkowska, M., et al., *Influence of chemical modification on the porous structure of polymeric adsorbents*. Materials Chemistry and Physics, 2011. **130**: p. 644-650.
50. Gritti, F., et al., *Physical properties and structure of fine core-shell particles used as packing materials for chromatography Relationships between particle characteristics and column performance*. J Chromatogr A, 2010. **1217**: p. 3819-43.
51. Ravikovitch, P.I. and A.V. Neimark, *Density functional theory model of adsorption on amorphous and microporous silica materials*. Langmuir, 2006. **22**: p. 11171-9.
52. Lastoskie, C., K.E. Gubbins, and N. Quirke, *Pore size distribution analysis of microporous carbons: a density functional theory approach*. The Journal of Physical Chemistry, 2002. **97**: p. 4786-4796.
53. Berendsen, G.E. and L.d. Galan, *A Geometrical Model for Chemically Bonded TMS and PDS Phases*. Journal of Liquid Chromatography, 2006. **1**: p. 403-426.
54. Berendsen, G.E. and L.d. Galan, *Preparation and Chromatographic Properties of Some Chemically Bonded Phases for Reversed-Phase Liquid Chromatography*. Journal of Liquid Chromatography, 1978. **1**: p. 561-586.
55. Broek, A.P., et al., *Characterization of hemodialysis membranes by inverse size exclusion chromatography*. Journal of Membrane Science, 1995. **99**: p. 217-228.
56. Athalye, A.M., S.J. Gibbs, and E.N. Lightfoot, *Predictability of chromatographic protein separations*. Journal of Chromatography A, 1992. **589**: p. 71-85.
57. Billen, J., et al., *Influence of the packing heterogeneity on the performance of liquid chromatography supports*. J Chromatogr A, 2005. **1073**: p. 53-61.
58. Farkas, T., M.J. Sepaniak, and G. Guiochon, *Column radial homogeneity in high-performance liquid chromatography*. Journal of Chromatography A, 1996. **740**: p. 169-181.
59. Baur, J.E. and R.M. Wightman, *Microcylinder electrodes as sensitive detectors for high-efficiency, high-speed liquid chromatography*. J Chromatogr, 1989. **482**: p. 65-73.

60. Abia, J.A., K.S. Mriziq, and G.A. Guiochon, *Radial heterogeneity of some analytical columns used in high-performance liquid chromatography*. J Chromatogr A, 2009. **1216**: p. 3185-91.
61. Gritti, F. and G. Guiochon, *A chromatographic estimate of the degree of heterogeneity of RPLC packing materials. 1. Non-endcapped polymeric C30-bonded stationary phase*. J Chromatogr A, 2006. **1103**: p. 43-56.
62. Gritti, F. and G. Guiochon, *A chromatographic estimate of the degree of surface heterogeneity of reversed-phase liquid chromatography packing materials. II-Endcapped monomeric C18-bonded stationary phase*. J Chromatogr A, 2006. **1103**: p. 57-68.
63. Gritti, F. and G. Guiochon, *A chromatographic estimate of the degree of surface heterogeneity of RPLC packing materials. III. Endcapped amido-embedded reversed phase*. J Chromatogr A, 2006. **1103**: p. 69-82.
64. Nawrocki, J., *The silanol group and its role in liquid chromatography*. Journal of Chromatography A, 1997. **779**: p. 29-71.
65. Scott, R.P.W. and P. Kucera, *Some Aspects of the Chromatographic Properties of Thermally Modified Silica Gel*. Journal of Chromatographic Science, 1975. **13**: p. 337-346.
66. Grumbach, E.S., D.M. Diehl, and U.D. Neue, *The application of novel 1.7 microm ethylene bridged hybrid particles for hydrophilic interaction chromatography*. J Sep Sci, 2008. **31**: p. 1511-8.
67. Pursch, M., L.C. Sander, and K. Albert, *Chain order and mobility of high-density c(18) phases by solid-state NMR spectroscopy and liquid chromatography*. Anal Chem, 1996. **68**: p. 4107-13.
68. Albert, K. and E. Bayer, *Characterization of bonded phases by solid-state NMR spectroscopy*. Journal of Chromatography A, 1991. **544**: p. 345-370.
69. Perkins, T.W., et al., *Protein retention in hydrophobic interaction chromatography: modeling variation with buffer ionic strength and column hydrophobicity*. Journal of Chromatography A, 1997. **766**: p. 1-14.
70. Kimata, K., et al., *Chromatographic Characterization of Silica C18 Packing Materials. Correlation between a Preparation Method and Retention Behavior of Stationary Phase*. Journal of Chromatographic Science, 1989. **27**: p. 721-728.
71. Walters, M.J., *Classification of octadecyl-bonded liquid chromatography columns*. Journal of the Association of Official Analytical Chemists, 1987. **70**: p. 465-9.
72. Engelhardt, H. and M. Jungheim, *Comparison and characterization of reversed phases*. Chromatographia, 1990. **29**: p. 59-68.
73. Cruz, E., et al., *Chromatographic classification of commercially available reverse-phase HPLC columns*. Chromatographia, 1997. **44**: p. 151-161.
74. Petersson, P. and M.R. Euerby, *An evaluation of the robustness of the Tanaka characterization protocol for reversed-phase liquid chromatography columns*. J Sep Sci, 2005. **28**: p. 2120-9.
75. ACDLabs. *ACD Column Selection Database*. 2021 [cited 2021 February 20]; Available from: <https://www.acdlabs.com/resources/freeware/colsel/>.
76. Keith J, L., *CHAPTER 1 - Reactions in Solution: General Principles*. 1963: p. 1-83.
77. Abraham, M.H., et al., *Hydrogen bonding. Part 13. A new method for the characterisation of GLC stationary phases—the laffort data set*. Journal of the Chemical Society, Perkin Transactions 2, 1990: p. 1451-1460.
78. Abraham, M.H., et al., *Hydrogen bonding*. Journal of Chromatography A, 1990. **518**: p. 329-348.

-
79. Baczek, T. and R. Kaliszan, *Combination of linear solvent strength model and quantitative structure-retention relationships as a comprehensive procedure of approximate prediction of retention in gradient liquid chromatography*. J Chromatogr A, 2002. **962**: p. 41-55.
 80. Michel, M., et al., *Comparative evaluation of high-performance liquid chromatography stationary phases used for the separation of peptides in terms of quantitative structure-retention relationships*. J Chromatogr A, 2007. **1175**: p. 49-54.
 81. Plass, M., K. Valko, and M.H. Abraham, *Determination of solute descriptors of tripeptide derivatives based on high-throughput gradient high-performance liquid chromatography retention data*. Journal of Chromatography A, 1998. **803**: p. 51-60.
 82. Mant, C.T., D. Cepenine, and R.S. Hodges, *Reversed-phase HPLC of peptides: Assessing column and solvent selectivity on standard, polar-embedded and polar endcapped columns*. J Sep Sci, 2010. **33**: p. 3005-21.
 83. Mant, C.T. and R.S. Hodges, *Design of peptide standards with the same composition and minimal sequence variation to monitor performance/selectivity of reversed-phase matrices*. J Chromatogr A, 2012. **1230**: p. 30-40.
 84. Field, J.K., et al., *Investigation into reversed phase chromatography peptide separation systems part I: Development of a protocol for column characterisation*. J Chromatogr A, 2019. **1603**: p. 113-129.
 85. Field, J.K., M.R. Euerby, and P. Petersson, *Investigation into reversed phase chromatography peptide separation systems part II: An evaluation of the robustness of a protocol for column characterisation*. J Chromatogr A, 2019. **1603**: p. 102-112.
 86. Field, J.K., M.R. Euerby, and P. Petersson, *Investigation into reversed phase chromatography peptide separation systems part III: Establishing a column characterisation database*. J Chromatogr A, 2020. **1622**: p. 461093.
 87. Field, J.K., et al., *Investigation into reversed-phase chromatography peptide separation systems Part IV: Characterisation of mobile phase selectivity differences*. J Chromatogr A, 2021. **1641**: p. 461986.
 88. Fekete, S., et al., *Importance of instrumentation for fast liquid chromatography in pharmaceutical analysis*. J Pharm Biomed Anal, 2014. **87**: p. 105-19.
 89. Guillarme, D., et al., *Method transfer for fast liquid chromatography in pharmaceutical analysis: application to short columns packed with small particle. Part II: gradient experiments*. European Journal of Pharmaceutics and Biopharmaceutics, 2008. **68**: p. 430-40.
 90. Fountain, K.J., et al., *Effects of extra-column band spreading, liquid chromatography system operating pressure, and column temperature on the performance of sub-2-microm porous particles*. J Chromatogr A, 2009. **1216**: p. 5979-88.
 91. Dasgupta, P.K., et al., *Flow-Cell-Induced Dispersion in Flow-through Absorbance Detection Systems: True Column Effluent Peak Variance*. Anal Chem, 2018. **90**: p. 2063-2069.
 92. Ettre, L.S., *Nomenclature for chromatography (IUPAC Recommendations 1993)*. Pure and Applied Chemistry, 1993. **65**: p. 819-872.
 93. International Union of, P. and C. Applied, *dead-volume*, in *The IUPAC Compendium of Chemical Terminology*. 2014.
 94. Knox, J.H. and R. Kaliszan, *Theory of solvent disturbance peaks and experimental determination of thermodynamic dead-volume in column liquid chromatography*. Journal of Chromatography A, 1985. **349**: p. 211-234.
 95. Alhedai, A., D.E. Martire, and R.P.W. Scott, *Column "dead volume" in liquid chromatography*. The Analyst, 1989. **114**: p. 869-875.

-
96. Gritti, F., Y. Kazakevich, and G. Guiochon, *Measurement of hold-up volumes in reverse-phase liquid chromatography Definition and comparison between static and dynamic methods*. J Chromatogr A, 2007. **1161**: p. 157-69.
 97. McCormick, R.M. and B.L. Karger, *Distribution phenomena of mobile-phase components and determination of dead volume in reversed-phase liquid chromatography*. Analytical Chemistry, 2002. **52**: p. 2249-2257.
 98. Rimmer, C.A., C.R. Simmons, and J.G. Dorsey, *The measurement and meaning of void volumes in reversed-phase liquid chromatography*. J Chromatogr A, 2002. **965**: p. 219-32.
 99. Krstulovic, A.M., H. Colin, and G. Guiochon, *Comparison of methods used for the determination of void volume in reversed-phase liquid chromatography*. Analytical Chemistry, 2002. **54**: p. 2438-2443.
 100. Slaats, E.H., et al., *Study of the influence of competition and solvent interaction on retention in liquid-solid chromatography by measurement of activity coefficients in the mobile phase*. Journal of Chromatography A, 1978. **149**: p. 255-270.
 101. Tsopelas, F., M. Ochsenkuhn-Petropoulou, and A. Tsantili-Kakoulidou, *Void volume markers in reversed-phase and biomimetic liquid chromatography*. J Chromatogr A, 2010. **1217**: p. 2847-54.
 102. Luo, H. and Y.K. Cheng, *A comparative study of void volume markers in immobilized-artificial-membrane and reversed-phase liquid chromatography*. J Chromatogr A, 2006. **1103**: p. 356-61.
 103. Gritti, F. and K. Smith, *Harmonization of Experimental Methods Used to Measure the True Hold-Up Volume of Liquid Chromatography Columns*. LCGC North America, 2023: p. 28-33.
 104. Bidlingmeyer, B.A., et al., *Some Practical Considerations when Determining the Void Volume in High-Performance Liquid Chromatography*. Journal of Chromatographic Science, 1991. **29**: p. 275-279.
 105. Engelhardt, H., H. Müller, and B. Dreyer, *Is there a "true" dead volume for HPLC columns?* Chromatographia, 1984. **19**: p. 240-245.
 106. Möckel, H.J., *Influence of temperature on dead volume of ODS columns and on n-alkane retention in high-performance liquid chromatography*. Journal of Chromatography A, 1994. **675**: p. 13-28.
 107. Pous-Torres, S., J.R. Torres-Lapasió, and M.C. García-Álvarez-Coque, *Performance of Markers and the Homologous Series Method for Dead Time Estimation in Reversed-Phase Liquid Chromatography*. Journal of Liquid Chromatography & Related Technologies, 2009. **32**: p. 1065-1083.
 108. Schellinger, A.P. and P.W. Carr, *Isocratic and gradient elution chromatography: a comparison in terms of speed, retention reproducibility and quantitation*. J Chromatogr A, 2006. **1109**: p. 253-66.
 109. Hage, D.S., *Chromatography*, in *Principles and Applications of Clinical Mass Spectrometry*, N. Rifai, A.R. Horvath, and C.T. Wittwer, Editors. 2018, Elsevier. p. 1-32.
 110. Vailaya, A. and C. Horváth, *Retention in reversed-phase chromatography: partition or adsorption?* Journal of Chromatography A, 1998. **829**: p. 1-27.
 111. Gritti, F. and G. Guiochon, *Adsorption mechanism in RPLC. Effect of the nature of the organic modifier*. Anal Chem, 2005. **77**: p. 4257-72.
 112. Guo, Y. and D. Baran, *Hydrophilic Partitioning or Surface Adsorption? A Quantitative Assessment of Retention Mechanisms for Hydrophilic Interaction Chromatography (HILIC)*. Molecules, 2023. **28**.

-
113. Dose, E.V., S. Jacobson, and G. Guiochon, *Determination of isotherms from chromatographic peak shapes*. Analytical Chemistry, 2002. **63**: p. 833-839.
 114. Wang, J. and X. Guo, *Adsorption isotherm models: Classification, physical meaning, application and solving method*. Chemosphere, 2020. **258**: p. 127279.
 115. Taher, A. and H.J.H. Brouwers, *Sorption isotherm measurements for porous materials: A new hygroscopic method*. Construction and Building Materials, 2023. **379**: p. 131166.
 116. Lin, W., et al., *Determining Granular Activated Carbon Adsorption Isotherms of Benzene as a Volatile Organic Compound*. Water Quality Research Journal, 2000. **35**: p. 219-230.
 117. Hlady, V., J. Buijs, and H.P. Jennissen, *[26] Methods for studying protein adsorption*, in *Methods in Enzymology*. 1999, Academic Press. p. 402-429.
 118. Reddy, T.V., S. Chauhan, and S. Chakraborty, *Adsorption isotherm and kinetics analysis of hexavalent chromium and mercury on mustard oil cake*. Environmental Engineering Research, 2016. **22**: p. 95-107.
 119. Przybyciel, M. and R. Majors, *Phase collapse in reversed-phase liquid chromatography*. LC-GC North America, 2002. **20**: p. 516+520-523.
 120. Gorner, T., et al., *Inverse Liquid Chromatography Investigation of Adsorption on Heterogeneous Solid Surfaces: Phenylalanine on Activated Carbon*. Langmuir, 2002. **18**: p. 8546-8552.
 121. Sabharwal, A.P. and H.A. Chase, *Preparative Reversed-Phase High Performance Liquid Chromatography of Insulin Variants*. Food and Bioproducts Processing, 1999. **77**: p. 18-26.
 122. Arnell, R., P. Forssen, and T. Fornstedt, *Accurate and rapid estimation of adsorption isotherms in liquid chromatography using the inverse method on plateaus*. J Chromatogr A, 2005. **1099**: p. 167-74.
 123. Horvath, S., et al., *Unbiased Determination of Adsorption Isotherms by Inverse Method in Liquid Chromatography*. Molecules, 2023. **28**: p. 1031.
 124. Norde, W., *Adsorption of proteins from solution at the solid-liquid interface*. Adv Colloid Interface Sci, 1986. **25**: p. 267-340.
 125. Vajda, P., A. Felinger, and G. Guiochon, *Evaluation of surface excess isotherms in liquid chromatography*. Journal of Chromatography A, 2013. **1291**: p. 41-47.
 126. Kazakevich, Y.V. and H.M. McNair, *Study of the Excess Adsorption of the Eluent Components on Different Reversed-Phase Adsorbents*. Journal of Chromatographic Science, 1995. **33**: p. 321-327.
 127. Riedo, F. and E.s. Kováts, *Adsorption from liquid mixtures and liquid chromatography*. Journal of Chromatography A, 1982. **239**: p. 1-28.
 128. Rodrigues, A.E., M.M. Dias, and J.C.B. Lopes, *Theory of Linear and Nonlinear Chromatography*, in *Chromatographic and Membrane Processes in Biotechnology*, C.A. Costa and J.S. Cabral, Editors. 1991, Springer Netherlands: Dordrecht. p. 25-52.
 129. Guiochon, G., A. Felinger, and D.G. Shirazi, *Fundamentals of preparative and nonlinear chromatography*. 2006: Elsevier.
 130. Dolan, J.W., *Count the Cost, Part I: Increasing Resolution by Increasing Column Efficiency*. LCGC North America, 2017. **35**.
 131. Hayes, R., et al., *Core-shell particles: preparation, fundamentals and applications in high performance liquid chromatography*. J Chromatogr A, 2014. **1357**: p. 36-52.
 132. Hayes, R., et al., *Monodisperse sphere-on-sphere silica particles for fast HPLC separation of peptides and proteins*. Analyst, 2014. **139**: p. 5674-7.

-
133. Krokhin, O.V. and V. Spicer, *Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides*. Anal Chem, 2009. **81**: p. 9522-30.
 134. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein*. J Mol Biol, 1982. **157**: p. 105-32.
 135. Tanford, C., *Contribution of Hydrophobic Interactions to the Stability of the Globular Conformation of Proteins*. Journal of the American Chemical Society, 2002. **84**: p. 4240-4247.
 136. Ponnuswamy, P.K., *Hydrophobic characteristics of folded proteins*. Prog Biophys Mol Biol, 1993. **59**: p. 57-103.
 137. Euerby, M.R. and P. Petersson, *Chromatographic classification and comparison of commercially available reversed-phase liquid chromatographic columns containing polar embedded groups/amino endcappings using principal component analysis*. J Chromatogr A, 2005. **1088**: p. 1-15.
 138. Thorsteinsdottir, U.A. and M. Thorsteinsdottir, *Design of experiments for development and optimization of a liquid chromatography coupled to tandem mass spectrometry bioanalytical assay*. Journal of Mass Spectrometry, 2021. **56**: p. e4727.
 139. Everett, D.H., *Reporting data on adsorption from solution at the solid/solution interface (Recommendations 1986)*. Pure and Applied Chemistry, 1986. **58**: p. 967-984.
 140. Bertie, J.E. and Z. Lan, *Liquid Water–Acetonitrile Mixtures at 25 °C: The Hydrogen-Bonded Structure Studied through Infrared Absolute Integrated Absorption Intensities*. The Journal of Physical Chemistry B, 1997. **101**: p. 4111-4119.
 141. Yang, Y., *A model for temperature effect on column efficiency in high-temperature liquid chromatography*. Analytica Chimica Acta, 2006. **558**: p. 7-10.
 142. Yang, Y., et al., *Temperature effect on peak width and column efficiency in subcritical water chromatography*. J Chromatogr Sci, 2002. **40**: p. 107-12.
 143. Edge, A.M., et al., *Temperature as a variable in liquid chromatography: development and application of a model for the separation of model drugs using water as the eluent*. J Chromatogr A, 2006. **1132**: p. 206-10.
 144. Chen, Y., C.T. Mant, and R.S. Hodges, *Temperature selectivity effects in reversed-phase liquid chromatography due to conformation differences between helical and non-helical peptides*. J Chromatogr A, 2003. **1010**: p. 45-61.
 145. Blackler, A.R., A.E. Speers, and C.C. Wu, *Chromatographic benefits of elevated temperature for the proteomic analysis of membrane proteins*. Proteomics, 2008. **8**: p. 3956-64.
 146. Bobaly, B., et al., *Impact of organic modifier and temperature on protein denaturation in hydrophobic interaction chromatography*. J Pharm Biomed Anal, 2016. **131**: p. 124-132.
 147. Lenco, J., et al., *Sense and Nonsense of Elevated Column Temperature in Proteomic Bottom-up LC-MS Analyses*. J Proteome Res, 2021. **20**: p. 420-432.
 148. Karamizadeh, S., et al., *An Overview of Principal Component Analysis*. Journal of Signal and Information Processing, 2013. **04**: p. 173-175.
 149. Stasiak, J., et al., *Principal component analysis of HPLC retention data and molecular modeling structural parameters of cardiovascular system drugs in view of their pharmacological activity*. International Journal of Molecular Sciences, 2010. **11**: p. 2681-98.
 150. Lee, S.C., et al., *Effects of textural properties on the response of a SnO₂-based gas sensor for the detection of chemical warfare agents*. Sensors (Basel), 2011. **11**: p. 6893-904.

151. Marchetti, N., et al., *Determination of adsorption isotherms by means of HPLC: adsorption mechanism elucidation and separation optimization*. J Sep Sci, 2009. **32**: p. 727-41.
152. Miyabe, K. and G. Guiochon, *Characterization of monolithic columns for HPLC*. J Sep Sci, 2004. **27**: p. 853-73.
153. Bracho, D., et al., *Functionalization of Silica Nanoparticles for Polypropylene Nanocomposite Applications*. Journal of Nanomaterials, 2012. **2012**: p. 1-8.
154. Hong, P., S. Koza, and E.S. Bouvier, *Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates*. Journal of Liquid Chromatography & Related Technologies, 2012. **35**: p. 2923-2950.
155. Irvine, G.B., *High-performance size-exclusion chromatography of peptides*. Journal of Biochemical and Biophysical Methods, 2003. **56**: p. 233-42.
156. Zucchetto, N. and D. Brühwiler, *Strategies for Localizing Multiple Functional Groups in Mesoporous Silica Particles through a One-Pot Synthesis*. Chemistry of Materials, 2018. **30**: p. 7280-7286.
157. Lecot, S., et al., *Arrangement of Monofunctional Silane Molecules on Silica Surfaces: Influence of Alkyl Chain Length, Head-Group Charge, and Surface Coverage, from Molecular Dynamics Simulations, X-ray Photoelectron Spectroscopy, and Fourier Transform Infrared Spectroscopy*. The Journal of Physical Chemistry C, 2020. **124**: p. 20125-20134.
158. Rybka, J., et al., *Stationary-Phase Contributions to Surface Diffusion in Reversed-Phase Liquid Chromatography: Chain Length versus Ligand Density*. The Journal of Physical Chemistry C, 2019. **123**: p. 21617-21628.
159. Yamada, S., *Cation- π Interactions in Organic Synthesis*. Chem Rev, 2018. **118**: p. 11353-11432.
160. Tatko, C.D. and M.L. Waters, *Comparison of C-H... π and hydrophobic interactions in a beta-hairpin peptide: impact on stability and specificity*. J Am Chem Soc, 2004. **126**: p. 2028-34.
161. Ainaravapu, S.R., et al., *Contour length and refolding rate of a small protein controlled by engineered disulfide bonds*. Biophysical Journal, 2007. **92**: p. 225-33.
162. Carr, P.W., et al., *Contributions to reversed-phase column selectivity. I. Steric interaction*. J Chromatogr A, 2011. **1218**: p. 1724-42.
163. Aguilar-Armenta, G. and L. Díaz-Jiménez, *Characterization of the porous structure of two naturally occurring materials through N₂-adsorption (77 K) and gas chromatographic methods*. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 2001. **176**: p. 245-252.
164. Rustamov, I., et al., *Geometry of chemically modified silica*. J Chromatogr A, 2001. **913**: p. 49-63.
165. Gritti, F. and G. Guiochon, *Mass transfer equation for proteins in very high-pressure liquid chromatography*. Anal Chem, 2009. **81**: p. 2723-36.
166. Lenco, J., et al., *Reversed-Phase Liquid Chromatography of Peptides for Bottom-Up Proteomics: A Tutorial*. J Proteome Res, 2022. **21**: p. 2846-2892.
167. Al Musaimi, O., V. Gavva, and D.R. Williams, *Greener Cleavage of Protected Peptide Fragments from Sieber Amide Resin*. ChemistryOpen, 2022. **11**: p. e202200236.
168. Alhassan, M., et al., *Cleaving protected peptides from 2-chlorotriyl chloride resin. Moving away from dichloromethane*. Green Chemistry, 2020. **22**: p. 2840-2845.
169. Tyteca, E., et al., *Applicability of linear and nonlinear retention-time models for reversed-phase liquid chromatography separations of small molecules, peptides, and intact proteins*. J Sep Sci, 2016. **39**: p. 1249-57.

170. Poole, C.F. and S.N. Atapattu, *Analysis of the solvent strength parameter (linear solvent strength model) for isocratic separations in reversed-phase liquid chromatography*. J Chromatogr A, 2022. **1675**: p. 463153.
171. Gonzalez-Lopez, N.M., et al., *Gradient Retention Factor Concept Applied to Method Development for Peptide Analysis by Means of RP-HPLC*. ACS Omega, 2022. **7**: p. 44817-44824.
172. Hruby, V.J., *Glucagon*, in *Molecular and Cellular Endocrinology*, E.E. Bittar and N. Bittar, Editors. 1997, Elsevier. p. 387-401.
173. Spezani, R. and C.A. Mandarim-de-Lacerda, *The current significance and prospects for the use of dual receptor agonism GLP-1/Glucagon*. Life Sciences, 2022. **288**: p. 120188.
174. Gastaldelli, A., M. Gaggini, and R. DeFronzo, *Glucose kinetics: an update and novel insights into its regulation by glucagon and GLP-1*. Current Opinion in Clinical Nutrition and Metabolic Care, 2017. **20**: p. 300-309.
175. Bakhtiani, P.A., et al., *A novel, stable, aqueous glucagon formulation using ferulic acid as an excipient*. Journal of Diabetes Science and Technology, 2015. **9**: p. 17-23.
176. Joshi, A.B., E. Rus, and L.E. Kirsch, *The degradation pathways of glucagon in acidic solutions*. Int J Pharm, 2000. **203**: p. 115-25.
177. Mountain, R.D., *Molecular Dynamics Simulation of Water–Acetonitrile Mixtures in a Silica Slit*. The Journal of Physical Chemistry C, 2013. **117**: p. 3923-3929.
178. Rybka, J., et al., *Molecular Dynamics Study of the Relation between Analyte Retention and Surface Diffusion in Reversed-Phase Liquid Chromatography*. The Journal of Physical Chemistry C, 2019. **123**: p. 3672-3681.
179. Trebel, N., et al., *Consequences of Cylindrical Pore Geometry for Interfacial Phenomena in Reversed-Phase Liquid Chromatography*. J Phys Chem B, 2021. **125**: p. 11320-11336.
180. Gritti, F. and G. Guiochon, *Thermodynamics of adsorption of binary aqueous organic liquid mixtures on a RPLC adsorbent*. J Chromatogr A, 2007. **1155**: p. 85-99.
181. Gritti, F., J. Sehajpal, and J. Fairchild, *Using the fundamentals of adsorption to understand peak distortion due to strong solvent effect in hydrophilic interaction chromatography*. J Chromatogr A, 2017. **1489**: p. 95-106.
182. Yeung, L.-S., *Assessment of Energetic Heterogeneity of Reversed-Phase Surfaces Using Excess Adsorption Isotherms for HPLC Column Characterization*, in *Department of Chemistry and Biochemistry*. 2018, Seton Hall University: South Orange, New Jersey.
183. Mountain, R.D., *Microstructure and hydrogen bonding in water-acetonitrile mixtures*. J Phys Chem B, 2010. **114**: p. 16460-4.
184. Rao, Y., N.J. Turro, and K.B. Eisenthal, *Water Structure at Air/Acetonitrile Aqueous Solution Interfaces*. The Journal of Physical Chemistry C, 2009. **113**: p. 14384-14389.
185. Bartalucci, E., et al., *Probing a Hydrogen- π Interaction Involving a Trapped Water Molecule in the Solid State*. Angew Chem Int Ed Engl, 2023. **62**: p. e202217725.
186. Chawla, M., et al., *Occurrence and stability of lone pair- π stacking interactions between ribose and nucleobases in functional RNAs*. Nucleic Acids Research, 2017. **45**: p. 11019-11032.
187. Paik, D., et al., *Thermodynamics of π - π Interactions of Benzene and Phenol in Water*. International Journal of Molecular Sciences, 2022. **23**: p. 9811.
188. Yabre, M., et al., *Greening Reversed-Phase Liquid Chromatography Methods Using Alternative Solvents for Pharmaceutical Analysis*. Molecules, 2018. **23**.

189. Do, D.D. and H.D. Do, *Appropriate volumes for adsorption isotherm studies: the absolute void volume, accessible pore volume and enclosing particle volume*. J Colloid Interface Sci, 2007. **316**: p. 317-30.
190. Kazakevich, Y.V. and H.M. McNair, *Thermodynamic Definition of HPLC Dead Volume*. Journal of Chromatographic Science, 1993. **31**: p. 317-322.
191. Petong, P., R. Pottel, and U. Kaatze, *Water–Ethanol Mixtures at Different Compositions and Temperatures. A Dielectric Relaxation Study*. The Journal of Physical Chemistry A, 2000. **104**: p. 7420-7428.
192. Zhong, Y., G.L. Warren, and S. Patel, *Thermodynamic and structural properties of methanol-water solutions using nonadditive interaction models*. Journal of Computational Chemistry, 2008. **29**: p. 1142-52.
193. Pal, J., A. Patla, and R. Subramanian, *Thermodynamic properties of forming methanol-water and ethanol-water clusters at various temperatures and pressures and implications for atmospheric chemistry: A DFT study*. Chemosphere, 2021. **272**: p. 129846.
194. Bako, I., et al., *Water-methanol mixtures: topology of hydrogen bonded network*. Physical Chemistry Chemical Physics, 2008. **10**: p. 5004-11.

Appendix I

Peptide hydrophobicity estimation: methods and scales

In 1961, Tanford and Lovrien [1] proved the relevance of the hydrophobicity effect and electrostatic interactions in the structure and stability of proteins by experimenting with sedimentation patterns for crystalline catalase under different pH levels. They observed a constant reduction in the enzymatic activity of crystalline protein compared to lyophilised samples. Such reduction is caused by the drying phase of the lyophilisation process, where some stable water molecules are forced out of the inner protein structure, creating a more compacted state due to an increase in hydrophobic interactions [2].

A decade later, Nozaki and Tanford [3] tried to experimentally estimate the solubility of standard amino acids in aqueous ethanol and dioxane solutions. During this work, they calculated the free energy of transfer in terms of solubility measurements of 11 amino acids, thus proposing the first hydrophobic scale of individual amino acids. After their publication, multiple hydrophobicity scales have been proposed to predict peptide hydrophobicity and understand the folding process of proteins. However, due to the disparity between scales, the estimation accuracy depends on the phenomenon being studied and how the hydrophobicity contribution of individual residues is considered.

The following segments of the Appendix describe experimental hydrophobicity scales; some computational work is also considered due to its relevance in the field. However, a chapter of the final Thesis will be dedicated only to computationally derived values, as they represent the great majority of the methods to calculate peptide/protein hydrophobicity.

1.1 Computational scales

1.1.1 Considering the contribution of all residues in the sequence equally

As peptide-based systems are formed of concatenated amino acids, it makes sense to determine their hydrophobicity by quantifying individual contributions from all residues in the structure. This type of calculation is called overall hydrophobicity. It can be expressed as the summation of the hydrophobicity values of each amino acid present in the sequence from N-terminal to C-terminal.

Inspired by the work of Nozaki and Tanford, Bull and Breese [4] developed what can be considered the first experimental approximation of a complete amino acid hydrophobicity scale. They calculated the free energy of transfer of individual amino acids from solution to surface by measuring the surface tension of amino acid solutions in 0.1 M NaCl following the Jones-Ray method. This technique was designed as a differential capillary rise method, where the surface tension is measured from the difference in height between the meniscus in a capillary and a reference line using a microscope camera and corrected by comparing with the calculations of a blank reference solution. Although the experiment could not achieve the expected accuracy, an apparent division of those amino acids with hydrophobic side chains from the polar group with the only exception of Ala, which might be due to its size.

Wolfenden and Cullis [5] estimated the affinity of peptides for watery surroundings by calculating the free energy transfer from a dilute vapour phase to an overdiluted aqueous solution. This vapour-liquid dynamic technique expresses these free-energy changes as the hydration potential with a negligible contribution of intermolecular forces. Since the intermolecular forces are virtually absent in the vapour phase and the buffers in aqueous solution prevent charged states, the changes in free energy could be quantified with negligible effect from solute-solute interactions. Although the hydration potential was proposed as a hydrophobicity scale, the underpinning analysis is too simplistic for the complexity of the system; thereby, no distinguishable differences between residue hydrophobicities are observable within the data.

A second attempt at calculating transfer free energy was performed by Heijne and Blomberg [6] from 9 prepiece sequences assumed to form a membrane-spanning α -helix. The hydrophobic contribution was measured from the free energy difference of the transfer of a residue from a random coil conformation in water to its alpha-helical conformation in a lipophilic phase. This approximation of a hydrophobicity scale based on the transfer of amino acids between liquid

phases of different polarities was an improved version of the analysis performed by Wolfenden and Cullis. However, since sequences used for the analysis in their work form alpha-helix structure, the scale is only valid for predicting trans-membrane translocating motifs of proteins.

Computational approaches have been able to expand the motif predicting power of these scales to other structures such as the hydrophobic moment for the detection of helical membrane segments, the alternating hydrophobicity for the detection of membrane-embedded β -sheets or the exact β -strand score considering the frequency of amino acids placed inward and outward.

1.1.2 Considering structural information

In the early years of protein hydrophobicity research, considering every residue that conforms to a peptide-based molecule had worked well at a small scale as a motif predictor. However, this approach does not offer realistic estimations on more complex applications where the contribution of the microenvironment surrounding the biomolecule is capable of altering the mechanistic of the system. Some phenomena include polarisation and nucleation during crystallisation, retention time in chromatographic columns or solubility in complex solutions.

This limitation is addressed by considering the bulk structure of a protein or peptide and its solvent-accessible surface area defined by the area in square Årmstrongs over which a solvent molecule can travel along the protein surface while making van der Waals contact and not penetrating any residue atom.

Janin [7] attempted to approach this statistically by calculating the transfer free energy of individual amino acids from an organic solvent to plain water. The data analysed was obtained from structural information of 22 globular proteins. They discovered that the free energy also represents the transfer of residues from the inside of a protein to the surface related to the ratio of buried/accessible molar fractions. They also noticed that the thermodynamics are governed by the presence or absence of oxygen and nitrogen atoms in the side chain and the spatial position of the residue, an effect that could be due to the ability to form hydrogen bonds.

Mavalan and Ponnuswamy [8] proposed the bulk hydrophobic character which accounts the hydrophobicity contribution of a residue i present in the protein surface as the sum of the hydrophobicity indices in the Tanford and Jones scale of those residues j that appear within a radius

of 8 Å. They determined this number as the optimum distance up to which a residue can exert influence on its surroundings members and vice versa by analysing 14 globular protein templates with significant and frequent structural characteristic [9].

The same authors later expanded the previous experiment, they introduced information of crystal structures from different membrane proteins from *Rhodopseudomonas viridis*, porins from *Rhodobacter capsulatus* and a set of 64 globular proteins [10]. They confirmed that even for more complex structures such as homo- and hetero-oligomers, the limit of 8 Å is sufficient to characterise the accommodation of local and non-local interactions into the hydrophobic behaviour of a residue. Therefore, this scale has been used to understand the folding stability and transition state structures of proteins.

Later, Ponnuswamy [11] tested their method with crystallographic data of 21 protein molecules whose structure represent examples of the four known structural classes (alpha, beta, alpha + beta, alpha/beta) proposed by Levit and Chothia [12]. After more than a decade, Cid et al. [13] expanded the Ponnuswamy method using a database of 60 proteins with their respective structural information. They found that by using the hydrophobicity coefficients calculated by this method, it is possible to predict the structural classification of the sequence according to the four structural classes. These results placed this hydrophobicity scale as one of the most common scales used to predict secondary and tertiary structures from amino acid sequences.

Berggren [14] introduced a more complex parameter called the average surface hydrophobicity (ASH, also symbolised as ϕ_{surface}) that can be explain mathematically through the following equation:

$$\phi_{\text{surface}} = \sum (\phi_{aax} r_{aax}) \quad (\text{I.1})$$

where ϕ_{aax} is the calculated value of the surface hydrophobicity for each amino acid in the surface and r_{aax} is the relative surface area exposed for each amino acid x. The original application of this parameter was to estimate the partitioning in aqueous two-phase system (ATPS), but the method was adapted by Lienqueo et al. [15] to chromatography to obtain a new hydrophobicity scale. Up to date, this a standard chromatographic method to measure hydrophobicity, thus it will be discussed in the following section.

1.2 Experimental

1.2.1 Via RPLC

Despite the evident advances in the field of hydrophobicity prediction as a result of computational methods, there are still some questions remaining as to what extent these statistical descriptors are valid. The main problem was that there were not many experimental designs that could confirm values to the level of detail as the computer could calculate.

A few years after the invention of RPLC, Meek [16] suggested that the variability in retention times for peptides could be predicted as the summing of the contribution to retention of each residue on its sequence just as those scales in section I.a. He injected 25 different and arbitrary peptides at pH 7.4 and 2.1 into a C₁₈ column and calculated retention coefficients for each amino acid according to the retention times of the biomolecules. In a later experiment [17], he expanded his work to 100 peptides with 20 residues or less to test the method. He iterated the values of the retention coefficients by adding sequences to the calculation progressively until a clear correlation between the predicted and the actual retentions was evident and within an acceptable level of variability. This is considered the first fully experimental peptide hydrophobicity scale using chromatography.

Just like Ponnuswamy became one of the most important scientists in the field of peptide/protein hydrophobicity prediction through the bulk hydrophobic character, Hodges' group could be considered the most influential research group in the application of the RPLC technique to calculate hydrophobicity scales. Their first approach [18] was to design novel model synthetic peptides where their retention time would be mostly affected by only one amino acid species; thus, giving new retention coefficients exclusive to the individual effect of each aa. They used the sequence Ac-Gly-X-X-(Leu)₃-(Lys)₂-amide, where X is substituted by the 20 amino acids found in proteins.

The distribution of amino acids along the hydrophobicity scale was similar in both methods described above, which showed the potential of using RPLC as a method to predict peptide and protein hydrophobicity. However, the perspective on how this data had to be interpreted had to change because during a chromatographic run only interactions between the surface of the biomolecule and the resin inside the column occur. Therefore, it was necessary to combine computational algorithms that could correlate the retention data with hydrophobicity surface contribution just like the methods described in section I.a.2.

Wilce et al. [19] combined structural and retention data to derive amino acid hydrophobicity coefficients by multiple linear regression using a matrix approach for solving sets of simultaneous structure-retention equations. They tested the model with a collection of retention data in n-alkyl silicas from 1738 peptides ranging in size from 2-50 amino acid residues with free amino and carboxyl-terminal groups. As expected, their results show mathematical consistency seem within the different hydrophobicity scales tested. However, the relevance of this calculations is still questionable because calculating hydrophobicity parameters of individual amino acids do not account peptide-ligand binding nor accommodate the influence of changes in secondary structure of solvation disorders.

Hodge's group recovered first approach and correlated the retention times to structural information of some proteins [18]. They created a more complex system composed of three parameters to predict hydrophobicity of amino acids depending on the average accessibility of each residue. The distribution of amino acids along the scale deviated from their original work, but it was consistent with some purely computational methods which suggests the relevance of using a chromatographic system to evaluate peptide hydrophobicity and the power of these structural based algorithms.

The main of using chromatography is that even when excellent separation of synthetic peptides is obtained, there are multiple parameters to control including (but not exclusive to) gradient steepness, flowrate, temperature, column packing material. If the chromatographic method is badly designed or if there are variations in any of the parameters, the retention behaviour of the peptides will change and thus the any scale developed from chromatographic techniques would not be repeatable. Cowan and Whittaker [20] proved this by using a shorter peptide sequence Z-X_{aa}-OMe and different pH levels (3.0 and 7.5). They normalised their scale using the peptide containing Gly as reference and found large variations to those of Hodges' group and Meek. Undoubtedly, this variability is provoked by the reduction in the length of the sequence, but also by the changes in to the chromatographic conditions.

In a more recent effort, Hodges' group tried to address this through a 10-residue sequence, Ac-X-G-A-K-G-A-G-V-G-L-amide, to determine the intrinsic hydrophobicity of amino acid side chains implying the maximum possible hydrophobicity of side chains in the absence of nearest-neighbour or conformation effects that would decrease the full expression of the hydrophobicity of each residue [21, 22]. Here they found that only potentially charged side chains (Lys, His, Arg,

Asp and Glu) are affected by pH, mobile phase composition, or functional groups. Also, they found that α -helical peptide models give high correlation with coefficients generated from structural information [23]. This result is expected because the basis of most computational algorithms are proteins with α -helices. For the same reason, variations with scales derived from retention data of random mixtures of peptides are not surprising because there is a lot of changes in conformation and frequency of amino acids.

Other techniques depend on the adaptation of chromatographic models to already existent hydrophobicity scales. Linqueo et al. [15] proposed a methodology to correlate dimensionless retention time estimated from hydrophobic interaction chromatography (HIC), a variant of RPLC, with ASH. This method has been proven valid in several cases; however, it should only be used to estimate hydrophobicity of proteins that exhibit highly homogenous distributions of the hydrophobic patches on their surfaces. This homogeneity can be estimated by calculating the hydrophobic contact area (HCA) using the three-dimensional structure of a protein [24]. This variable has been found to give an insight of the surface hydrophobicity distribution in the surface of proteins. Other algorithms more detailed than HCA are local hydrophobicity [25], hydrophobic imbalance [26] and statistical description of the protein surface [27]. These are not reviewed in this document because they are fully computational.

1.3 References

1. Tanford, C. and R. Lovrien, *Dissociation of Catalase into Subunits*. Journal of the American Chemical Society, 2002. **84**(10): p. 1892-1896.
2. Lale, S.V., M. Goyal, and A.K. Bansal, *Development of lyophilization cycle and effect of excipients on the stability of catalase during lyophilization*. Int J Pharm Investig, 2011. **1**(4): p. 214-21.
3. Nozaki, Y. and C. Tanford, *The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale*. J Biol Chem, 1971. **246**(7): p. 2211-7.
4. Bull, H.B. and K. Breese, *Surface tension of amino acid solutions: a hydrophobicity scale of the amino acid residues*. Arch Biochem Biophys, 1974. **161**(2): p. 665-70.
5. Wolfenden, R.V., P.M. Cullis, and C.C. Southgate, *Water, protein folding, and the genetic code*. Science, 1979. **206**(4418): p. 575-7.
6. von Heijne, G. and C. Blomberg, *Trans-membrane translocation of proteins. The direct transfer model*. Eur J Biochem, 1979. **97**(1): p. 175-81.
7. Janin, J., *Surface and inside volumes in globular proteins*. Nature, 1979. **277**(5696): p. 491-2.

8. Manavalan, P. and P.K. Ponnuswamy, *Hydrophobic character of amino acid residues in globular proteins*. Nature, 1978. **275**(5681): p. 673-4.
9. Ponnuswamy, P.K., *Hydrophobic characteristics of folded proteins*. Prog Biophys Mol Biol, 1993. **59**(1): p. 57-103.
10. Ponnuswamy, P.K. and M.M. Gromiha, *Prediction of transmembrane helices from hydrophobic characteristics of proteins*. Int J Pept Protein Res, 1993. **42**(4): p. 326-41.
11. Ponnuswamy, P.K., M. Prabhakaran, and P. Manavalan, *Hydrophobic packing and spatial arrangement of amino acid residues in globular proteins*. Biochim Biophys Acta, 1980. **623**(2): p. 301-16.
12. Levitt, M. and C. Chothia, *Structural patterns in globular proteins*. Nature, 1976. **261**(5561): p. 552-8.
13. Cid, H., et al., *Hydrophobicity and structural classes in proteins*. Protein Eng, 1992. **5**(5): p. 373-5.
14. Berggren, K., et al., *The surface exposed amino acid residues of monomeric proteins determine the partitioning in aqueous two-phase systems*. Biochim Biophys Acta, 2002. **1596**(2): p. 253-68.
15. Lienqueo, M.E., A. Mahn, and J.A. Asenjo, *Mathematical correlations for predicting protein retention times in hydrophobic interaction chromatography*. J Chromatogr A, 2002. **978**(1-2): p. 71-9.
16. Meek, J.L., *Prediction of peptide retention times in high-pressure liquid chromatography on the basis of amino acid composition*. Proc Natl Acad Sci U S A, 1980. **77**(3): p. 1632-6.
17. Meek, J.L. and Z.L. Rossetti, *Factors affecting retention and resolution of peptides in high-performance liquid chromatography*. Journal of Chromatography A, 1981. **211**(1): p. 15-28.
18. Parker, J.M., D. Guo, and R.S. Hodges, *New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites*. Biochemistry, 1986. **25**(19): p. 5425-32.
19. Wilce, M.C.J., M.-I. Aguilar, and M.T.W. Hearn, *Physicochemical Basis of Amino Acid Hydrophobicity Scales: Evaluation of Four New Scales of Amino Acid Hydrophobicity Coefficients Derived from RP-HPLC of Peptides*. Analytical Chemistry, 2002. **67**(7): p. 1210-1219.
20. Cowan, R. and R.G. Whittaker, *Hydrophobicity indices for amino acid residues as determined by high-performance liquid chromatography*. Pept Res, 1990. **3**(2): p. 75-80.
21. Kovacs, J.M., C.T. Mant, and R.S. Hodges, *Determination of intrinsic hydrophilicity/hydrophobicity of amino acid side chains in peptides in the absence of nearest-neighbor or conformational effects*. Biopolymers, 2006. **84**(3): p. 283-97.
22. Kovacs, J.M., et al., *Quantitation of the nearest-neighbour effects of amino acid side-chains that restrict conformational freedom of the polypeptide chain using reversed-phase liquid chromatography of synthetic model peptides with L- and D-amino acid substitutions*. J Chromatogr A, 2006. **1123**(2): p. 212-24.
23. Mant, C.T., et al., *Intrinsic amino acid side-chain hydrophilicity/hydrophobicity coefficients determined by reversed-phase high-performance liquid chromatography of model peptides: comparison with other hydrophilicity/hydrophobicity scales*. Biopolymers, 2009. **92**(6): p. 573-95.

24. Mahn, A., M.E. Lienqueo, and J.A. Asenjo, *Effect of surface hydrophobicity distribution on retention of ribonucleases in hydrophobic interaction chromatography*. J Chromatogr A, 2004. **1043**(1): p. 47-55.
25. Mahn, A., G. Zapata-Torres, and J.A. Asenjo, *A theory of protein-resin interaction in hydrophobic interaction chromatography*. J Chromatogr A, 2005. **1066**(1-2): p. 81-8.
26. Salgado, J.C., I. Rapaport, and J.A. Asenjo, *Predicting the behaviour of proteins in hydrophobic interaction chromatography. 1: Using the hydrophobic imbalance (HI) to describe their surface amino acid distribution*. J Chromatogr A, 2006. **1107**(1-2): p. 110-9.
27. Salgado, J.C., I. Rapaport, and J.A. Asenjo, *Predicting the behaviour of proteins in hydrophobic interaction chromatography. 2. Using a statistical description of their surface amino acid distribution*. J Chromatogr A, 2006. **1107**(1-2): p. 120-9.

Appendix II

Derivation of the Peak Maxima (PM) method

From the infinitesimal representation of the chromatographic peak migrating through the RPLC column (Figure 4.10). The following equation describes the total mass balance of the concentration of analyte (n) at that infinitesimal space:

$$\frac{\delta n}{\delta t} = cF - (c + \delta c)F = -\delta cF = -\left(\frac{\delta c_L}{\delta x}\right)_t dx F$$

where F is the bulk liquid flow rate, c is the concentration of the analyte in the bulk liquid, δc is the analyte concentration change at the infinitesimal space. This concentration change can also be expressed in terms of distance travelled x by adjusting the partial derivatives.

The total mass balance n can also be expressed in terms of where the analytes are located within the infinitesimal space. These locations can be either adsorbed onto the stationary phase ($\frac{\delta n_{MP}}{\delta t}$) or dissolved into the bulk liquid ($\frac{\delta n_{SP}}{\delta t}$). By changing the calculation base to distance travelled, the equation changes into:

$$\frac{\delta n}{\delta t} = \left(\frac{\delta n_{MP}}{\delta t}\right)_x + \left(\frac{\delta n_{SP}}{\delta t}\right)_x = \left(\frac{\delta c_L}{\delta t}\right)_x \varepsilon A_c dx + \left(\frac{\delta q}{\delta t}\right)_x (1 - \varepsilon) A_c dx$$

where $\varepsilon A_c dx$ is the partial space/volume available as mobile phase (hold-up volume) within the infinitesimal space, determined by the total porosity ε and surface area A_c of the resin. For simplicity, the analyte content in the bulk liquid and adsorbed onto the stationary phase are expressed as c_L and q , respectively.

The two equations are equal, and further simplification can be done.

$$-F \left(\frac{\delta c_L}{\delta x} \right)_t \left(\frac{\delta t}{\delta c_L} \right)_x = \varepsilon A_c + (1 - \varepsilon) A_c \left(\frac{\delta q}{\delta c_L} \right)_x$$

$$\frac{F}{\left(\frac{\delta x}{\delta t} \right)_{c_L}} = \varepsilon A_c + (1 - \varepsilon) A_c \left(\frac{\delta q}{\delta c} \right)_x$$

$$F t_r = \varepsilon A_c x + (1 - \varepsilon) A_c x \frac{dq}{dc_L}$$

From this equation, the variables for the volumes of the mobile (V_M) and stationary phases (V_{SP}) can be replaced, and equations 4.24 and 4.26.

$$V_r = V_M + V_{SP} \frac{dq}{dc_L}$$

$$\frac{dq}{dc_L} = \frac{V_r - V_M}{V_{SP}} = \frac{V_n}{V_{SP}}$$

$$q = \int_0^{c_L} \frac{V_n}{V_{SP}} dc_L$$

Appendix III

Derivation of excess isotherm equations

The total amount of solvent B (organic solvent in the case of RPLC) can be defined as $\frac{\eta_B}{A} = \Gamma_B^V + \frac{V_{ads}}{A} C_B^l = \Gamma_B^V + \tau C_B^l$, where V_{ads} represents the volume adsorbed layer per unit surface. However, this equation is only valid in the linear region of the excess adsorption isotherm where the volume of the adsorbed liquid layer V_{ads} onto the adsorbent is at its maximum.

If the volume of a binary solvent system with no adsorbent remains constant or an equilibrium is reached between the molar volume of both solvents on the adsorbed surface and in the bulk liquid phase, the following can be assumed

$$\eta_B = V^l(C_B^o - C_B^l)$$
$$\Gamma_B^V = \frac{\eta_B}{A} = \frac{V^l(C_B^o - C_B^l)}{A}$$

where, η_B is the number of moles of solvent B adsorbed onto the surface in excess of the bulk liquid above the adsorbent surface, V^l volume of the binary solvent system, C_B^o is the concentration of solvent B in the bulk liquid before adding adsorbent into the vessel and equilibrating with the adsorption surface, C_B^l is the concentration of solvent B in the bulk liquid after adding adsorbent into the vessel, Γ_B^V is the excess adsorption of solvent B per unit of surface area, A is the surface area.

However, these previous equations are tedious to apply to HPLC column characterisation because they assume that a static adsorption process is achieved.

Thus, minor disturbance peak determination method from Y.V. Kazakevich, H.M. McNair, J. Chromatogr. Sci. 33 (1995) 321 can be applied to connect HPLC behaviour and excess adsorption. The advantage of this approach is the amount of excess adsorption can be determined by

experimental measurement without a priori assumption of an adsorption model. **One model is only needed for the interpretation of the excess adsorption isotherm.**

$$v_r = v_m + A \frac{d\Gamma_B^v}{dC_B^l} \Rightarrow d\Gamma_B^v = \frac{v_r - v_m}{A} dC_B^l \Rightarrow \int_0^{C_B^l} d\Gamma_B^v = \int_0^{C_B^l} \frac{v_r - v_m}{A} dC_B^l$$

$$v_r = v_m + A \frac{d\Gamma_B^v}{dC_B^l} \Rightarrow v_r dC_B^l = v_m dC_B^l + A d\Gamma_B^v \Rightarrow$$

$$\int_{C_B^l(0\%)}^{C_B^l(100\%)} v_r dC_B^l = v_m \int_{C_B^l(0\%)}^{C_B^l(100\%)} dC_B^l + A \int_{C_B^l(0\%)}^{C_B^l(100\%)} d\Gamma_B^v$$

$$\Gamma_B^v(0\%) = \Gamma_B^v(100\%) = 0 ; \text{ excess adsorption in pure solvent}$$

$$\therefore v_m = \frac{\int_{C_B^l(0\%)}^{C_B^l(100\%)} v_r dC_B^l}{\int_{C_B^l(0\%)}^{C_B^l(100\%)} dC_B^l} = \frac{\int_{C_B^l(0\%)}^{C_B^l(100\%)} v_r dC_B^l}{C_B^l(100\%)}$$

In heterogeneous surfaces like silica-based resins, different adsorption regions will be observed due to the presence of hydrophobic alkyl chains and hydrophilic residual surface silanol groups. In mathematical representation, the following can be assumed:

Solvent B will be attracted to the hydrophobic parts of the surface

$$\frac{\eta_B}{A} = \Gamma_B^{\text{hpho}} + V^{\text{hpho}} C_B^l \Rightarrow \Gamma_B^v = \frac{\eta_B}{A} - V^{\text{hpho}} C_B^l$$

where, V^{hpho} is the volume of liquids adsorbed on the hydrophobic surfaces per unit of surface area.

Since the organic solvent B molecules are mainly accumulated on the hydrophobic part of the surface, thus:

$$\frac{\eta_B}{A} = V^{\text{hpho}} C_B^{\text{hpho}}$$

$$\therefore \Gamma_B^{\text{hpho}} = V^{\text{hpho}} (C_B^{\text{hpho}} - C_B^l)$$

Polar solvent (water in RPLC) will be attracted to the hydrophilic parts of the surface:

$$\frac{\eta_w}{A} = \Gamma_w^{\text{hphi}} + V^{\text{hphi}} C_w^l \Rightarrow \Gamma_w^{\text{hphi}} = \frac{\eta_w}{A} - V^{\text{hphi}} C_w^l$$

where, V^{hphi} is the volume of liquids adsorbed on the hydrophilic surfaces per unit of surface area.

Since water molecules are mainly accumulated on the hydrophobic part of the surface, thus

$$\frac{\eta_w}{A} = V^{\text{hphi}} C_w^{\text{hphi}}$$

$$\therefore \Gamma_w^{\text{hphi}} = V^{\text{hphi}} (C_w^{\text{hphi}} - C_w^{\text{l}})$$

The sum of the excess adsorption may express the total excess adsorption of organic solvent B and water due to hydrophobic surface interaction and hydrophilic surface interaction.

$$\Gamma_B^{\text{T}} = \Gamma_B^{\text{hpho}} + \Gamma_B^{\text{hphi}}$$

$$\Gamma_w^{\text{T}} = \Gamma_w^{\text{hphi}} + \Gamma_w^{\text{hpho}}$$

Since this is a binary system, increasing one component implies a decrease of its complementary component in the same volume.

$$\frac{\Gamma_B^{\text{hphi}}}{D_B} = - \frac{\Gamma_w^{\text{hphi}}}{D_w}$$

and

$$x_w = 1 - x_B$$

$$\frac{C_w^{\text{l}}}{D_w} = 1 - \frac{C_B^{\text{l}}}{D_B}$$

$$C_w^{\text{l}} = D_w - D_w \frac{C_B^{\text{l}}}{D_B}$$

where x_B and x_w represents the mole fraction of solvent B and water, respectively. Similarly, D_B and D_w are the molar density of solvent B and water, respectively.

$$\begin{aligned}
 \therefore \Gamma_B^T &= V^{\text{hpho}}(C_B^{\text{hpho}} - C_B^l) - V^{\text{hphi}}(C_w^{\text{hphi}} - C_w^l) \frac{D_B}{D_w} \\
 &= V^{\text{hpho}}(C_B^{\text{hpho}} - C_B^l) - V^{\text{hphi}} \left(C_w^{\text{hphi}} - \left[D_w - D_w \frac{C_B^l}{D_B} \right] \right) \frac{D_B}{D_w} \\
 &= V^{\text{hpho}}(C_B^{\text{hpho}} - C_B^l) - V^{\text{hphi}} \left(C_w^{\text{hphi}} \frac{D_B}{D_w} - D_B + C_B^l \right)
 \end{aligned}$$

During the linear region of the superposed excess adsorption isotherm, the composition and the volume of the adsorbed layer on the unit surface area of the composite surface are constant.

Thus, $V^{\text{hpho}} = V_B$, $C_B^{\text{hpho}} = D_B$, $V^{\text{hphi}} = V_w$, $C_w^{\text{hphi}} = D_w$

$$\therefore \Gamma_B^T = V_B D_B - (V_B + V_w) C_B^l$$

Appendix IV

Peptides under elution conditions

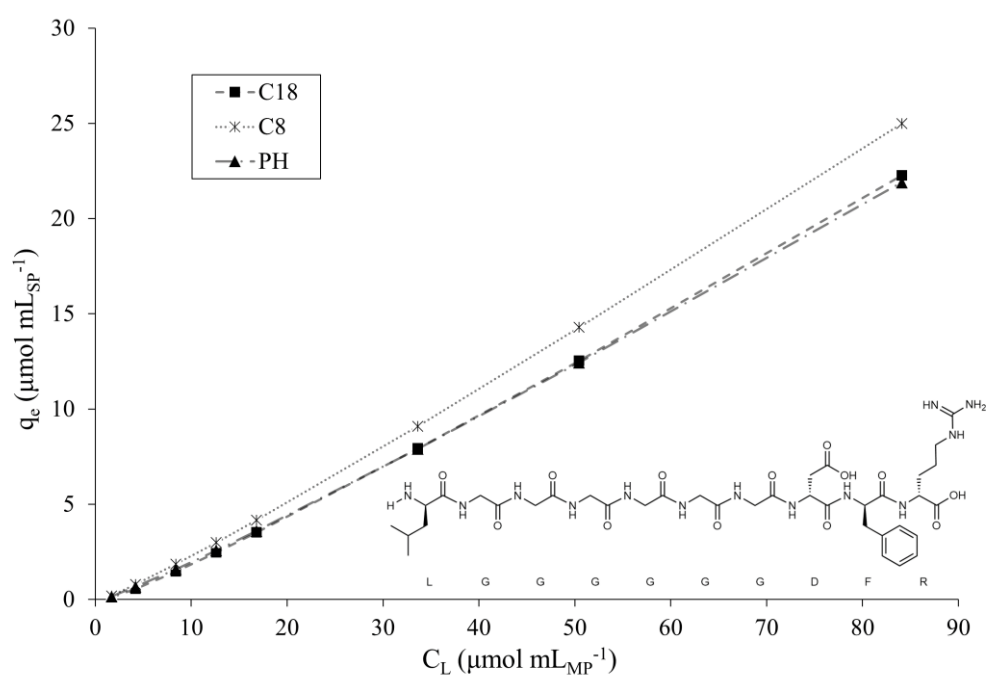


Figure IV.1 | Adsorption isotherm of peptide P2 at eluting conditions for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 25% ACN/H₂O (v/v) + 0.1% TFA (v/v).

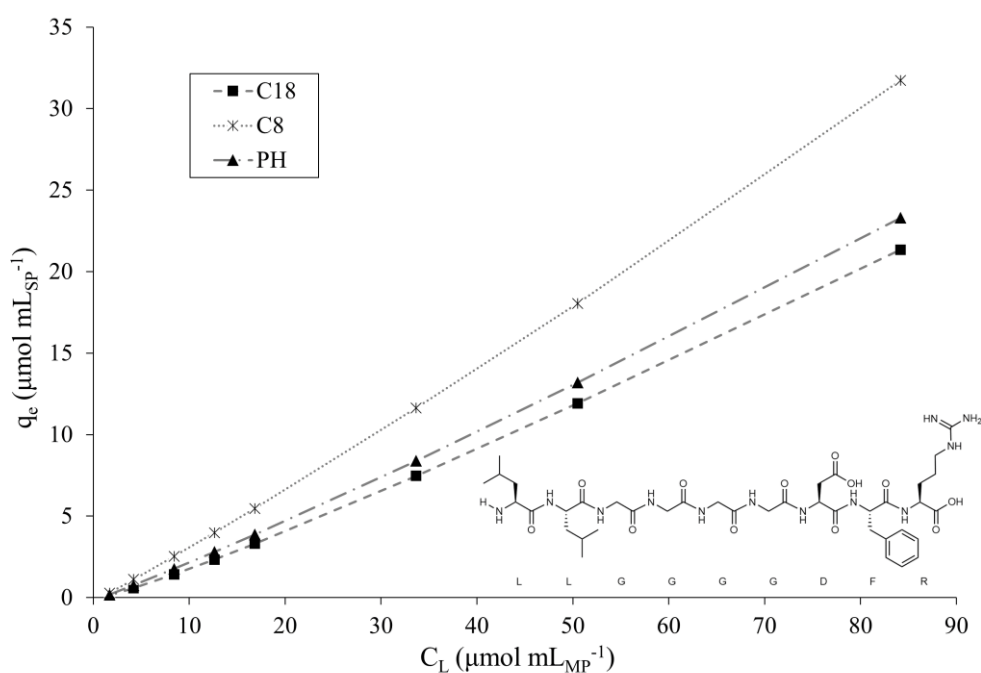


Figure IV.2 | Adsorption isotherm of peptide P3 at eluting conditions for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 30% ACN/H₂O (v/v) + 0.1% TFA (v/v).

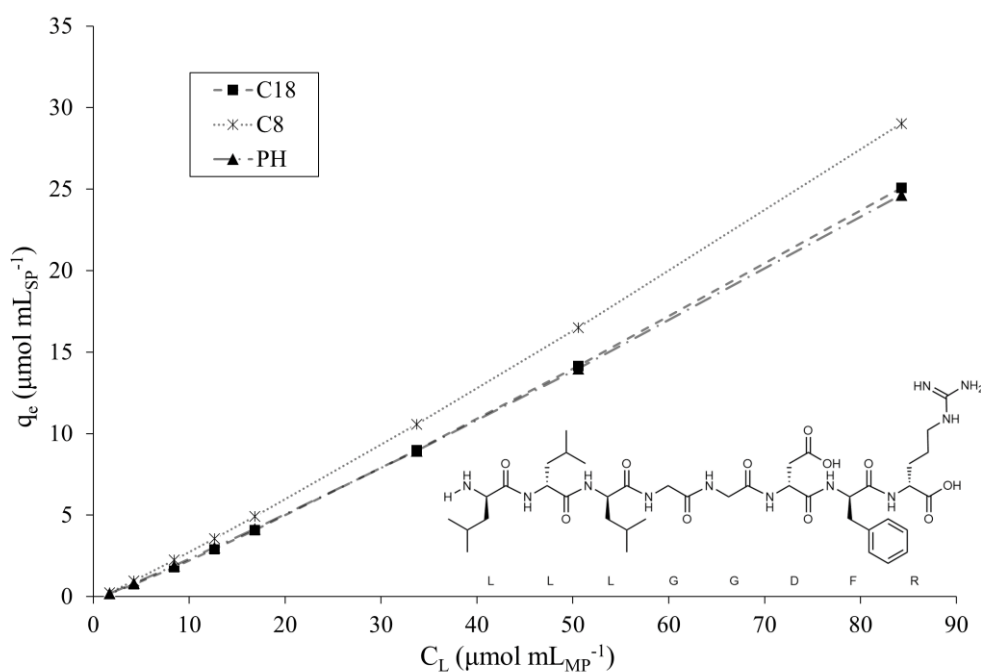


Figure IV.3 | Adsorption isotherm of peptide P4 at eluting conditions for RPLC columns. Isotherms were measured at 25 °C using the peak maxima method for three RPLC columns: Luna C18, C8 and phenyl-hexyl (PH). The mobile phase was 35% ACN/H₂O (v/v) + 0.1% TFA (v/v).

Appendix V

Characterisation of Polyethylene Fibres as a Reference Hydrophobic Material for Chromatographic Separation of Peptides

I extend my appreciation to Shivani Thangaraj and Irmantas Racaitis, who supported this Appendix by executing the experiments and contributing to research conclusions during their Chemical Engineering third-year research project at Imperial College London.

5.1 Materials and Methods

5.1.1 Column packing

All liquid chromatography (LC) experiments were performed using a stainless-steel 300 x 4.6 mm ID ($V_c \cong 4.99$ mL) column packed with 2.6 g of Spectra 75 PE fibres previously cleaned in a glass beaker with five washes of water and ethanol and dried overnight at 70 °C. This packing density resulted in a packing volume of 50–60%. The threads were spun to produce a 30 cm diameter loop of aligned cleaned PE fibres. Then, an additional thread length was tied to the end of the loop to pull the fibres through the column. An *in-situ* washing step was performed using an LC-20AD pump (Shimadzu, Japan). Water and ethanol were pumped individually at 1 mL/min for approximately $10V_c$.

For inverse gas chromatography (IGC) experiments, a silanised glass 300 mm x 4 mm ID column was packed with 1.68 g of Spectra 75 PE fibres. The same packing procedure described above was followed, excluding the final *in-situ* washing step.

5.1.2 Coating of Polyethylene and Peptides on Glass Substrates

Soda-lime glass slides (75 x 25 mm, 1.0-1.2 mm thickness, VWR, United Kingdom) and substrates with 20 nm ultra-flat quartz coating (15 x 20 mm, Ossila, United Kingdom) were cleaned by sonication in soap solution for 10 min, double rinsed with warm DI water before and after a final immersion in 1% (v/v) Hellmanex solution for 5 min. The drying procedure consisted of sonication in acetone and DI water for 10 and 20 minutes, respectively, followed by oven drying overnight at 70 °C.

Cleaned Spectra 75 fibres were cut into smaller fibre sections and fixed into a clean glass slide using double-sided adhesive tape. An additional slide with only adhesive tape was used as a control sample.

Individual peptides were spin-coated onto cleaned glass substrates using an Ossila spin coater using the static dispense method at 1000 rpm for 30 seconds. A 50 µL aliquot of 1 mg/mL peptide sample dissolved in methanol was pipetted onto the centre of the substrates, followed by immediate spinning. Control samples were produced by repeating the same procedure with pure methanol.

5.1.3 Inverse Liquid Chromatography (ILC)

A modular HPLC Prominence LC-20 (Shimadzu) was utilised for all LC experiments mentioned in this work. This instrument was equipped with a DGU-20A5R in-line degasser, an LC-20AD pump, a SIL-20A autosampler injector, two SPD-20A UV/VIS detectors (before and after column detection) and a CTO-20AC column oven. The estimated hold-up and dwell volume to the after-column detector are 0.36 and 1.46 mL, respectively. The system was controlled using a CBM-20A communication bus module and LabSolutions (ver. 5.92, Shimadzu). Data analysis was performed using the post-run analysis tool built in the same software and Microsoft Excel for Microsoft 365 MSO 64-bit version 2205 (Microsoft, United States).

A modified and expanded Tanaka method [1] was employed to characterise the stainless-steel PE fibre column and a Luna C18(2) (150 × 4.6 mm ID, 5 µm) column (Phenomenex, United States). The expansion consisted of an increase in the number of arene species used to determine parameter α_{CH_2} which describes the column hydrophobicity selectivity based on a single methylene (-CH₂-) group of difference between molecular probes. Compared to the original method, which varied between 0.6 and 1 mL/min, the flow rate was standardised to 0.8 mL/min for all runs. Samples were dissolved in methanol or water and filtered through 0.20-µm hydrophobic polytetrafluoroethylene (PTFE) or 0.22-µm hydrophilic polyvinylidene fluoride (PVDF) filter units (Merck, Germany), respectively.

Additional variations in flow rate and injection volumes were performed on the PE fibre column.

The Tanaka method was modified after initial testing of the Spectra 75 column to improve the chromatograms' quality. The methanol concentration of the mobile phase in the Spectra 75 column was varied between 0–60% to determine the percentage at which the samples were sufficiently retained to calculate all the required parameters. The flow rate was then increased to produce narrower peaks. The conditions decided upon for the mobile phase were a methanol concentration of 45% and a flowrate of 5 mL/min.

The behaviour of the Spectra 75 column was analysed by injecting Tanaka test samples at 50 and 100 µL.

5.1.4 Material characterisation

The morphology of the PE fibres was examined and photographed through scanning electron microscopy (SEM) imaging using a JSM-6010 SEM (Jeol, Japan). A small sample of cleaned fibres was mounted onto a metal stub with double-sided carbon adhesive tape and sputter-coated with gold.

Specific surface area was measured through low-temperature nitrogen adsorption (LTNA) using a TriStar 3000 gas adsorption analyser (Micromeritics, United States). A mass of 0.20 g of cleaned PE fibres was transferred into Micromeritics sample tubes and degassed under

a nitrogen atmosphere for 24 h at 240 °C by triplicate. LTNA measurement was conducted at liquid nitrogen temperature (77 K) with 30 adsorption and 20 desorption points to produce the isotherm. Results were analysed using the Micromeritics TriStar 3000 software by modelling the measurements to the Brunauer–Emmett–Teller (BET) equation.

The contact angle (CA) of DI water and diiodomethane was employed to measure the surface energy at ambient temperature (22-25 °C). A separate dispensing syringe was dedicated for each test liquid to avoid cross-contamination. The advancing contact angle for sessile droplets was measured by initially forming a 1- μ L droplet at the end of the syringe and carefully placed onto the sample surface. Pictures of the droplet were taken every second for 20 seconds or until the CA was stable. Then, the droplet was inflated with 1 μ L of test liquid dispensing syringe and remeasured. This advancing procedure was repeated until the CA drastically dropped due to the expansion of the droplet at the liquid-solid interface (increase of droplet diameter). This procedure was repeated on the spin-coated peptides, Spectra 75-coated glass slides, and their respective control samples by triplicates. Results were processed using the Fowkes surface energy model [2].

PE fibres ' specific surface area and energy were compared to inverse gas chromatography (IGC) data using an iGC-SEA system (Surface Measurement Systems, United Kingdom). Samples were run at a series of decreasing volume octane injections to obtain an adsorption isotherm at 30°C, allowing for calculation of the BET-specific surface area of the fibres using the SEA Analysis software. Similarly, several polar (ethanol, ethyl acetate, acetonitrile, dichloromethane) and non-polar (hexane, heptane, octane, nonane) probes were injected into the samples at 30°C to achieve fractional surface coverages over a range of 0.005 – 0.200. The dispersive, specific, and total surface energy profiles and distributions were calculated directly through the SEA Analysis software. The Doris and Gray method was employed for the dispersive component.

5.1.5 HPLC of Peptides

Solutions of model peptides (P1–P5) were prepared for HPLC analysis with a concentration of 1 mg/mL in water with 25% acetonitrile (ACN). All methods used 1 μ L injections of each sample.

The chromatographic runs for individual peptides were carried out with a linear gradient from 0% to 100% of ACN in H₂O (gradient slope of 0.5% ACN/min). The Spectra 75 column was run at a 2 mL/min flow rate, whereas the Luna C18(2) column was run at 1.6 mL/min due to maximum pressure safety restrictions in the HPLC system. In addition, isocratic elution was carried out for the Spectra 75 column at 5, 8 and 10% ACN in H₂O. The dead volumes of both columns were determined with acetone injections of 10 µL.

5.2 Results and Discussion

5.2.1 Inverse Liquid Chromatography

Tanaka Test

The Tanaka test was carried out using the Luna C18(2) column to compare obtained parameters with database values from the ACD/Column Selector tool developed by Advanced Chemistry Development, Inc[3]. After confirming the reproducibility of the results, the test was carried out on the Spectra 75 column.

However, it was found that the Tanaka method cannot be normalised for characterising this column due to the large difference in hydrophobicity and retention capacity of the stationary phase compared to the Luna C18(2) column. Under the standard conditions for the method, there was little to no solute retention in the Spectra 75 column. A reduction of the methanol concentration in the mobile phase from 80% v/v to at least 45% v/v was required to achieve sufficient retention in the column, indicating that the hydrophobic interactions between Spectra 75 fibres and the analytes are weaker than those in the Luna C18(2) column, and likely also when compared with other commercial columns. The changes in chromatographic conditions make a direct comparison with database values for commercial columns invalid [4]. Despite this, the nature of the samples injected using the Tanaka method allows for a better understanding of the column characteristics.

For the purpose of this study, the hydrophobic behaviour of the Spectra 75 column was particularly significant. Therefore, the hydrophobic retention capacity was of key interest. This

parameter was calculated from the expansion of Run 1 using the series of alkylbenzene injections. Uracil was injected as an indicator of the column void volume, which was determined to be 3.15 mL. The retention time of uracil (t_M) was then used to calculate the retention factors of all samples with the equation:

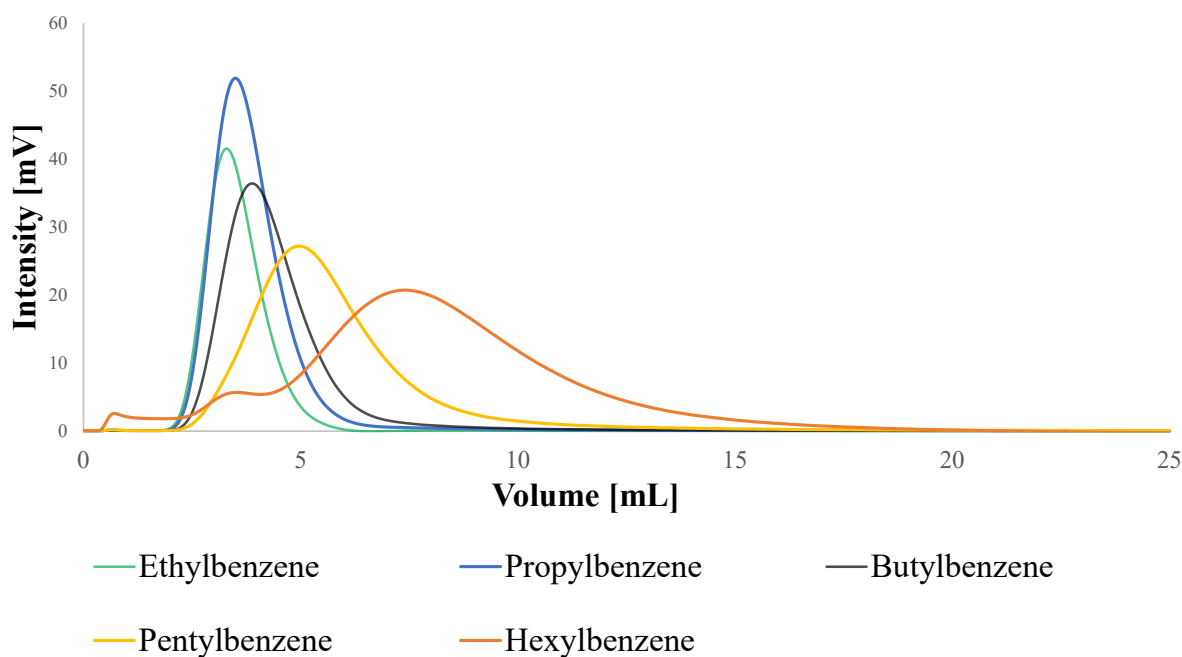


Figure V.1 | Elution profiles of Tanaka test expansion compounds in Spectra 75 column under isocratic conditions.

As shown in Figure V.1, the column demonstrates a degree of hydrophobic retention as the retention volume increases with each subsequent alkylbenzene. However, retention factors are significantly lower than those obtained with the commercial column, and the peak shapes deviate further from ideal Gaussian peaks with increasing molecule size. There is observably greater peak broadening and tailing, which can be attributed to the low surface area of the stationary phase.

The chemical structure of polyethylene ensures that the column does not have the capacity for hydrogen bonding or electrostatic interactions with solutes, which is reflected in the results of runs 2, 3 and 4. All samples in these runs eluted at the column void volume as the interactions between the solutes and the stationary phase were much weaker than interactions with the more polar mobile phase. In a typical column, the parameters calculated from these runs ($\alpha_{C/P}$, $\alpha_{B/P, pH\ 7.6}$, $\alpha_{B/P, pH\ 2.7}$) are indicators of silanol activity in the stationary phase.

The results from the Tanaka test prove that the Spectra 75 column retains hydrophobic solutes without any charged interactions. They also highlight the significant limitation of the column's ability to separate components due to the very low packing surface area, which is demonstrated in the poor resolution of the chromatogram peaks.

5.2.2 Morphological & Chemical Properties of the Fibres

Surface Area

Initially, the surface area was estimated using the physical data of the Spectra 75 fibres (Honeywell). Since the column was packed with fibres extending throughout the whole column, this calculation was carried out while assuming a cylindrical filament shape and neglecting the base area of the cylinder:

$$a = \frac{1}{2.9725 \cdot \sqrt{dpf \cdot \rho}} [m^2 g^{-1}] \quad (1)$$

Using this equation, the surface area of the fibre was estimated to be $0.248\ m^2 g^{-1}$. This theoretical value was compared with the BET surface area acquired by two techniques: nitrogen adsorption and inverse gas chromatography (IGC). The acquired values were $0.3544\ m^2 g^{-1}$ and $0.6897\ m^2 g^{-1}$, respectively. The BET surface area acquired by IGC was considerably larger than the theoretical or nitrogen adsorption values.

Even though there is a discrepancy of estimated surface area among the different techniques, it still reveals that the surface area of the Spectra 75 HPLC column is about 3 orders of magnitude lower than that of conventional reversed-phase column resins, such as Luna C18(2) with $400\ m^2$

g⁻¹ (Phenomenex). As a result, the separation power of the Spectra 75 column is greatly reduced as the analytes interact less with the stationary phase due to its limited available surface. Since the Tanaka tests for the Spectra 75 column displayed low retention and resolution of analytes, this technical limitation necessitates an alternatively manufactured polyethylene fibre to obtain a higher surface area for improved retention and resolution.

Filament Diameter

The effective diameter of the fibres was also calculated from the surface areas obtained from the physical specifications, nitrogen adsorption experiments and IGC (The calculated values are shown in Table V.1):

$$d = \frac{4}{a \cdot \rho} \tag{2}$$

Table V.1 | Spectra 75 filament diameter.

	Theoretical	N₂ Adsorption	IGC
Diameter [µm]	16.6	11.6	5.2

With the help of SEM images (see Figure V.2), the diameter of the Spectra 75 filaments was estimated to be 17.7 ± 1.9 µm. The SEM images revealed that the Spectra 75 filaments' cross-section is rectangular. This suggests a flat filament as opposed to the previously assumed cylindrical shape. This leads to the fibres being characterised by an effective diameter instead of a geometrical diameter.

Compared to other estimations, the SEM data suggest the largest diameter for the Spectra 75 filaments. This discrepancy is most likely linked to the aforementioned flat shape of the fibre, where the wider side of the fibres was seen more frequently in the SEM images. As a result, the SEM diameter estimations were skewed towards the longer dimension of the cross-section and

were not so useful. In contrast, other utilised methods (theoretical, N₂ adsorption, IGC) provided a more reasonable value (effective diameter).

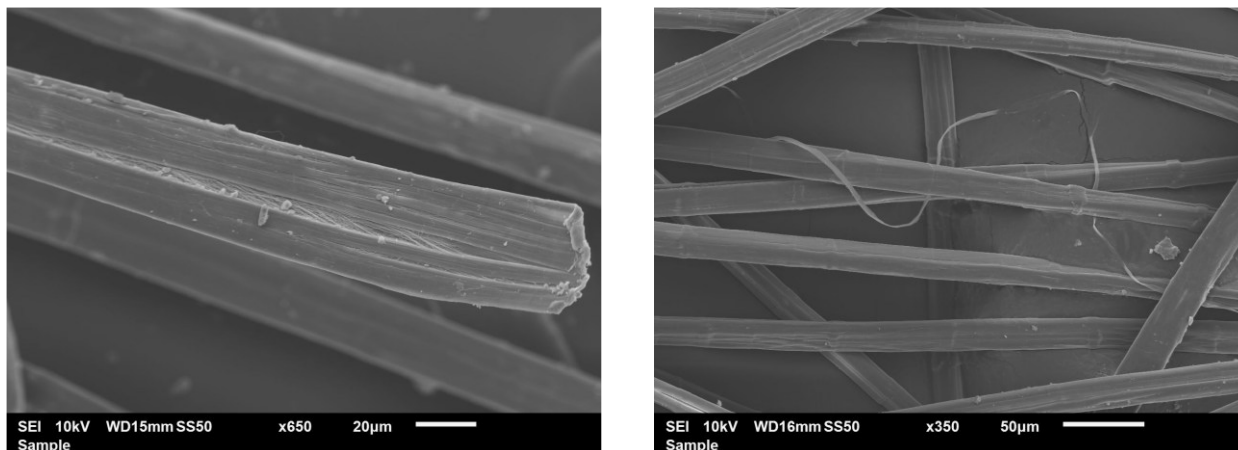


Figure V.2 | SEM imaging of Spectra 75 fibres.

Surface Energy

Contact angle measurements and inverse gas chromatography (IGC) determined the surface energy of the Spectra 75 fibres.

The advancing contact angles of water and diiodomethane on the Spectra 75 surface were measured at 131.0° and 9.7°, respectively. During experiments with diiodomethane, the liquid droplets would wet the surface completely. However, a very flat droplet would form just briefly before collapsing, so the whole measurement was recorded, and a value was taken at the frame where the droplet still held its shape. As a result, this measurement was less precise than the others.

The dispersive and polar surface energy components were then estimated from contact angle data with the Fowkes method:

$$\sqrt{\sigma_l^D \cdot \sigma_s^D} + \sqrt{\sigma_l^P \cdot \sigma_s^P} = \frac{\sigma_l(1 + \cos \theta)}{2} \quad (3)$$

The calculated values were then compared to surface energies acquired by inverse gas chromatography (see Table V.2).

Table V.2 | Surface energy of Spectra 75 fibres determined by contact angle and IGC.

	Dispersive surface energy [mJ/m ²]	Polar surface energy [mJ/m ²]	Total surface energy [mJ/m ²]
Contact angle	50.1	8.2	58.3
IGC	42.8	3.6	46.4

It was observed that the surface energy components determined by the contact angle measurements are higher than those determined by IGC. Generally, for a rough surface, contact angles below 90° tend to appear lower than they are. This phenomenon might have occurred with diiodomethane on the Spectra 75 surface, leading to a higher surface energy estimation than with IGC. In contrast, measurements of water would be affected, too, since contact angles above 90° tend to appear higher than they are. This could've contributed to the overestimation of the polar component of surface energy. The issue of surface roughness could be alleviated by modifying the fibre coating procedure, for example, using a hydraulic press to press down the fibres and/or evaluating the surface roughness by other techniques, such as white light interferometry or confocal microscopy.

All in all, contact angle measurements aren't as reliable as IGC, and the latter technique provides more insight into the energetic heterogeneity of the Spectra 75 fibres. Hence, the remaining part of this section will focus on this technique. The IGC results are shown in Figure V.7.

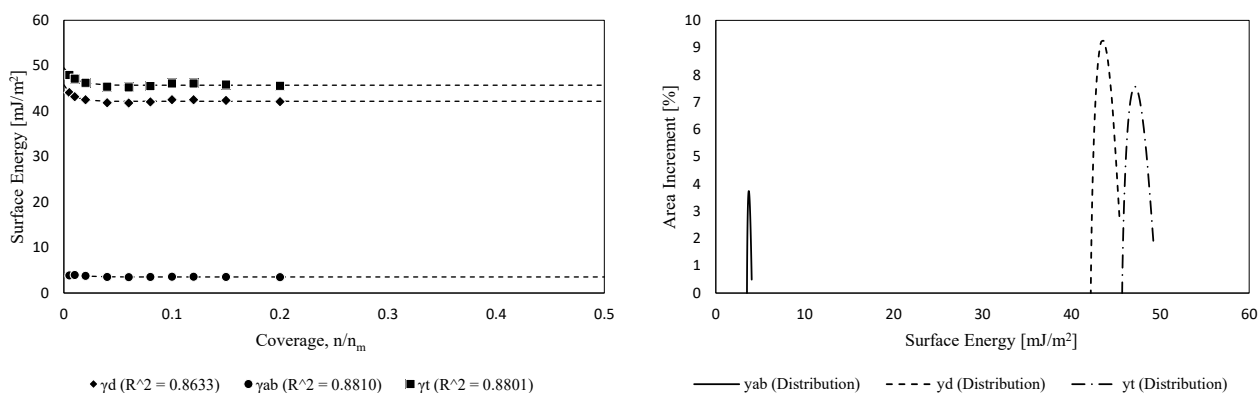


Figure V.7 | Surface energy and distributions of Spectra 75 fibres obtained by IGC.

The surface energy data was fitted with an exponential decay regression model in the form $y=y_0+a\cdot e^{-b\cdot x}$. As expected, the surface energy of the fibre sample decreases with coverage. While this is a common feature of IGC analysis for all materials, it is the range of these surface energies (Table V.2) that is paramount in evaluating energetic heterogeneity.

Figure V.7(right) depicts a relatively narrow surface energy distribution of the Spectra 75 fibres, suggesting a quite energetically homogeneous surface, which is consistent with the rigid crystalline structure of the fibre. In addition, it can be seen that the dispersive surface energy is the dominant component, accounting for over 90% of the total surface energy. This property of polyethylene fibres confirms that this resin would interact with analytes via mostly dispersive interactions and can be used as a hydrophobic reference. Further investigation by analysing an alkyl-bound silica resin would provide a better idea of how the surface properties of Spectra 75 compared to conventional resins.

5.2.3 Chromatographic Separation of Peptides

Chromatographic tests were conducted to evaluate the Spectra 75 column for the development of peptide separation and a hydrophobicity scale. The main Thesis shows that the Krokhin standards are uniformly eluted from an RPLC column (see the main body of the Thesis).

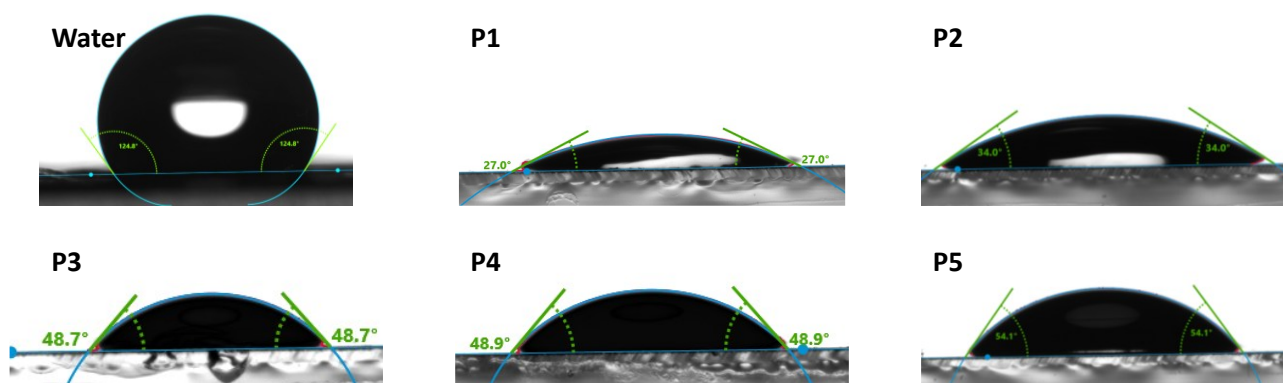


Figure V.4 | Contact angle of peptide films.

Clearly, the peptides could be ranked based on their retention times, and thus, a hydrophobicity scale was developed. A hydrophobicity index for each peptide was expressed as the concentration at which the peptide elutes. A similar approach can be found in the study by Krokhin and Spicer, where isocratic elution was utilised to convey the hydrophobicity index as the mobile phase composition at which the retention factor of the peptide is 10. This method, however, couldn't have been applied to the Spectra 75 column since there was insufficient retention of the analytes at higher acetonitrile concentrations. On the other hand, lowering the concentration resulted in extremely shallow peaks, which were hard to analyse. Due to these limitations, this project relies on gradient elution instead. Figure V.5 shows the elution profile of the 5 standard peptides through the Spectra 75 fibre column.

The developed ranking was compared with modelled GRAVY (Grand average of hydropathicity index) values and water contact angles of the peptides (Figure V.4). The GRAVY values were manually calculated by using the amino acid hydropathicity scale developed by Kyte & Doolittle [5].

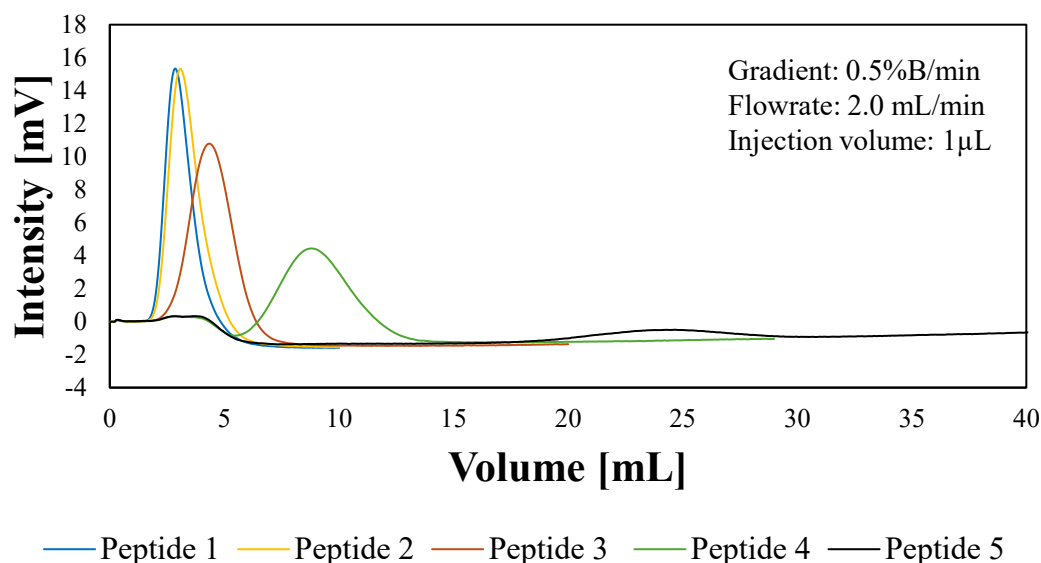


Figure V.5 | Peptide elution through Spectra 75 c

Table V.3 compares the three developed scales. The peptides' hydrophobicity rankings match. However, these scales yield very different hydrophobicity values for the model peptides after normalisation.

Table V.3 | Overview of three peptide hydrophobicity scales. GRAVY values were based on the Kyte-Doolittle amino acid scale. Contact angles were determined experimentally on spin-coated peptide films. %ACN values were determined for Spectra 75 column with gradient elution from 0 to 100% acetonitrile in water at 0.5 % /min, flow rate: 2 mL/min.

Peptide	Sequence	Hydrophobicity Measure					
		GRAVY	Normalised GRAVY	H ₂ O Contact Angle	Normalised Contact Angle	HPLC (%ACN)	Normalised %ACN
P1	LGGGGGGDGSR	-0.71	0.000	26.5	0.000	0.71	0.000
P2	LGGGGGGDFR	-0.38	0.154	33.9	0.254	0.77	0.011
P3	LLGGGGDFR	0.09	0.374	46.3	0.680	1.11	0.073
P4	LLLGGDFR	0.67	0.645	49.5	0.790	2.20	0.277

P5	LLLLDFR	1.43	1.000	55.6	1.000	6.09	1.000
----	---------	------	-------	------	-------	------	-------

As a result, a correlation was developed between the GRAVY values and the peptide's elution concentration and contact angle (Figure V.6). These correlations relate the experimental hydrophobicity indices with theoretical GRAVY values. It was found that the correlation between the GRAVY values and the elution through the Spectra 75 fibre columns followed a linear relationship with a higher coefficient of determination than that with contact angle.

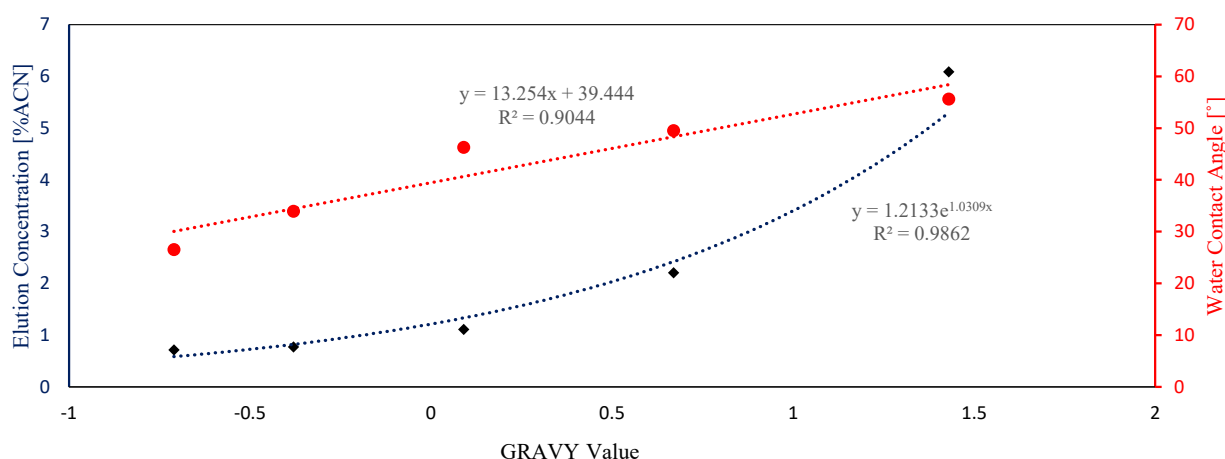


Figure V.6 | Hydrophobicity correlation plots for peptide scales.

Quality of Separation

Gradient elution was found to produce the greatest difference in retention of all five peptides in the Spectra 75 column, so the same method was used in the Luna C18(2) column for comparison.

Figure 8 displays the results of the gradient elution of each peptide run in the Luna C18(2) column. The elution volumes are significantly greater than in the Spectra 75 column. This large difference in retention capacity between the columns can be attributed primarily to the significant difference in surface area. As found through our material characterisation study, the specific surface area of the Spectra 75 fibre is less than 1 m²/g compared to 400 m²/g in the Luna C18(2)

column. Another factor of note is the greater hydrophobicity of the stationary phase in the Luna column, as demonstrated by the previously conducted Tanaka test, and, therefore, stronger interactions with the injected peptides. The resolutions between adjacent peaks were calculated for both columns and listed in Table V.4.

The Spectra column improved in resolution with each pair of peptides. Between P4 and P5, there is sufficient resolution for a separation process, for which resolutions greater than 1.5 are acceptable to achieve good separation.

Table V.4 | Resolution between adjacent chromatographic peaks for Spectra 75 and Luna C18(2) columns.

Peptide pair	Resolution	
	Spectra 75	Luna C18(2)
P2/P1	0.10	4.32
P3/P2	0.46	15.54
P4/P3	0.97	13.41
P5/P4	2.00	14.94

5.3 Conclusions

In this study, the pure hydrophobic retention within the Spectra 75 fibre column was demonstrated. Compared to silica-based columns, the Spectra column displayed no capacity for polar interactions, and the column behaviour was notably unaffected by changes in pH. However, issues of low separation capability arose due to the significantly lower surface area relative to conventional columns. The potential of an RPLC column unaffected by the issues arising from free silanol groups is worth further study to improve separation power by increasing surface area within the column. This may be achieved by reducing the fibre diameter, developing capillary channels, and increasing column size and packing fraction.

There was sufficient retention with gradient elution to establish a hydrophobicity scale for peptides based on elution concentration and correlate it with a GRAVY scale. This relationship could be used to estimate the optimal mobile phase composition for the separation of a peptide with a known GRAVY value. In addition, this novel scale could replace more conventional scales based on modelling, contact angle and silica-based RPLC columns. This study could be expanded by investigating peptides with a wider hydrophobicity range and performing isocratic elution. To investigate the latter, the separation power of the column should be improved, as mentioned previously.

Although the separation power of Spectra 75 is low relative to conventional columns, it could be utilised in processes with large hydrophobicity differences between analytes with the benefit of high column throughput. In addition, implementing alternative organic modifiers, such as ethanol, could be investigated to make the separations more environmentally friendly.

5.4 References

1. Kimata, K., et al., *Chromatographic Characterization of Silica C18 Packing Materials. Correlation between a Preparation Method and Retention Behavior of Stationary Phase*. Journal of Chromatographic Science, 1989. **27**(12): p. 721-728.
2. Fowkes, F.M., *Additivity of Intermolecular Forces at Interfaces. I. Determination of the Contribution to Surface and Interfacial Tensions of Dispersion Forces in Various Liquids*. The Journal of Physical Chemistry, 2002. **67**(12): p. 2538-2541.
3. ACDLabs. *ACD Column Selection Database*. 2021 [cited 2021 February 20]; Available from: <https://www.acdlabs.com/resources/freeware/colsel/>.
4. Petersson, P. and M.R. Euerby, *An evaluation of the robustness of the Tanaka characterization protocol for reversed-phase liquid chromatography columns*. Journal of Separation Science, 2005. **28**(16): p. 2120-9.
5. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein*. Journal of Molecular Biology, 1982. **157**(1): p. 105-32.