



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.



**Keele  
University**

**Investigation of modified forms of  
serum C-reactive protein in diabetes  
mellitus**

**By Sameer Mohammed Mahmood**

Thesis submitted for the degree of Doctor of  
Philosophy

October 2019

Keele University



## Abstract

C-reactive protein (CRP) is known to dissociate to its monomeric subunits (mCRP) *in vivo*. Furthermore, hyperglycaemia associated with diabetes mellitus (DM) results in glycation of serum proteins to a various extent based on their half-life and glycability. The research described here focuses on modified forms (glycated and monomeric) of the acute phase reactant, CRP, in the serum of DM patients with elevated levels of CRP.

The susceptibility of CRP to the glycation was investigated *in vitro*. Human CRP was treated using various glucose concentrations and durations, and the glycation status and putative glycation sites determined by mass spectrometry. Analysis of glucose treated samples shows that CRP can be glycated and the glycation extent is both time- and glucose concentration-dependent. Between one and three glucose molecules were found to be condensed on the protein after a period equivalent to its serum half-life, suggesting the possibility of glycation occurring *in vivo* under high glucose conditions. Among the identified putative glycation sites, K13 was found to be the most susceptible lysine residue in CRP toward glycation.

Human serum samples from participants with DM and elevated CRP levels (>100mg/L, n=24) were investigated for the presence of monomeric and glycated CRP variant. Monomeric CRP was detected in all investigated serum samples, with an average mCRP content of 0.093 mg/L, however no correlation was found between serum levels of mCRP and pCRP. The mCRP showed the ability to reversibly bind phosphocholine, moreover the mCRP displayed a molecular weight of 23,026 Da as determined by mass spectrometry. All 24 diabetic samples were negative for the presence of glycated CRP, suggesting the half-life of the protein is insufficient to cause a detectable level of glycated CRP.

## Table of contents

<b>ABSTRACT</b> .....	<b>I</b>
<b>TABLE OF CONTENTS</b> .....	<b>II</b>
<b>LIST OF TABLES</b> .....	<b>VIII</b>
<b>LIST OF FIGURES</b> .....	<b>X</b>
<b>TABLE OF AMINO ACIDS</b> .....	<b>XIV</b>
<b>ABBREVIATIONS USED</b> .....	<b>XV</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>XVIII</b>
<b>CHAPTER 1: INTRODUCTION</b> .....	<b>1</b>
1.1 HUMAN C-REACTIVE PROTEIN .....	1
1.1.1 <i>C-reactive protein and innate immunity</i> .....	1
1.1.2 <i>Historical background</i> .....	2
1.1.3 <i>Biological functions of C-reactive protein</i> .....	3
1.1.3.1 <i>Activation of the complement system</i> .....	4
1.1.3.2 <i>Binding to FcγR receptors</i> .....	7
1.1.4 <i>C-reactive protein ligands</i> .....	9
1.1.5 <i>Structure of C-reactive protein</i> .....	10
1.1.5.1 <i>The ligand recognition face of CRP</i> .....	11
1.1.5.2 <i>The Effector face of CRP</i> .....	14
1.1.5.3 <i>The structural difference between calcium-bound and calcium-depleted CRP</i> .....	16
1.1.6 <i>Different forms of C-reactive protein</i> .....	17
1.1.6.1 <i>Monomeric C-reactive protein</i> .....	17
1.1.6.2 <i>Fibril-like C-reactive protein</i> .....	18
1.1.6.3 <i>Glycosylated C-reactive protein</i> .....	19
1.1.7 <i>CRP synthesis and gene expression</i> .....	19
1.1.8 <i>The serum concentration of C-reactive protein in health and disease</i> .....	21

1.1.9 Applications of CRP as a clinical marker .....	23
1.1.10 Methods of measurements .....	24
1.1.11 Demographic, behavioural and lifestyle factors that affect baseline CRP .....	25
1.1.12 C-reactive protein as a low-grade inflammatory marker.....	27
1.1.13 CRP and cardiovascular disease .....	27
1.1.13.1 Evidence from epidemiological studies .....	27
1.1.13.2 Role of CRP in atherosclerosis and atherothrombosis .....	29
1.1.13.3 Doubts about the in vitro experiment.....	32
1.1.14 CRP and metabolic syndrome .....	34
1.1.15 CRP and diabetes mellitus.....	35
1.2 PROTEIN GLYCATION .....	36
1.2.1 Introduction .....	36
1.2.2 The chemistry of protein glycation.....	37
1.2.3 Biological effects of protein glycation.....	41
1.3 RESEARCH AIMS AND OBJECTIVES .....	43
<b>CHAPTER 2: PRODUCTION &amp; CHARACTERISATION OF GLYCATED C-REACTIVE PROTEIN .....</b>	<b>45</b>
2.1 INTRODUCTION.....	45
2.1.1 Overview .....	45
2.1.2 Analytical techniques used for detection of glycosylated proteins .....	45
2.1.3 Polyacrylamide Gel Electrophoresis .....	46
2.1.4 Western Blotting .....	49
2.1.5 Mass spectrometry.....	51
2.1.6 Experimental Aims .....	57
2.2 MATERIALS & METHODS.....	58
2.2.1 Reagents and consumables.....	58
2.2.2 Preparation of Glycosylated C-Reactive Protein.....	59
2.2.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis.....	60
2.2.4 Western blotting .....	61

2.2.5 Liquid Chromatography-Mass Spectrometry.....	63
2.2.6 Calculation of glycoforms distribution percentage, the extent of glycation and glucose index ....	65
2.3 RESULTS .....	67
2.3.1 Preliminary characterisation of glycated samples using SDS-PAGE and western blotting .....	67
2.3.2 Characterisation of CRP glycation using LC-MS.....	71
2.3.2.1 LC-MS analysis of CRP samples incubated for 19 hours .....	72
2.3.2.2 LC-MS analysis of samples incubated for 3 days .....	75
2.3.2.3 LC-MS analysis of samples incubated for 6 days .....	77
2.3.3 The percentage distribution of glycoforms and the extent of glycation.....	80
2.3.4 Glucose Index (G-idx) .....	83
2.4 DISCUSSION.....	85
2.4.1 Preliminary characterising of in vitro glycated CRP with SDS-PAGE and western blotting .....	86
2.4.2 Using LC-MS for the intact analysis of glycated proteins .....	88
2.4.3 LC-MS analysis of the in vitro glycated CRP samples.....	90
2.4.4 Conclusion.....	95
<b>CHAPTER 3: REVEALING THE GLYCATION SITES OF <i>IN VITRO</i> GLYCATED C-REACTIVE PROTEIN .....</b>	<b>96</b>
3.1 INTRODUCTION .....	96
3.1.1 Overview.....	96
3.1.2 Proteomic approaches to identify modification sites by mass spectrometry .....	97
3.1.3 Peptide production via enzyme digestion.....	99
3.1.4 Peptide separation.....	101
3.1.5 Peptide mass fingerprinting.....	102
3.1.6 Tandem mass spectrometry .....	104
3.1.7 Identification of peptides.....	109
3.1.8 Fragmentation behaviour of glycated peptides under CID condition .....	110
3.1.9 Other technologies and strategies to characterise glycated peptides.....	112
3.1.10 Experimental Aims.....	113
3.2 MATERIALS & METHODS .....	114

3.2.1 Glycated CRP samples used in this study .....	114
3.2.2 Sample pre-treatment and protein digestion.....	114
3.2.3 LC-ESI-MS/MS Analysis of Protein Digests .....	116
3.2.4 Collection and analysis of mass spectrometry data .....	117
3.2.5 Physicochemical properties of lysine and arginine residues in CRP .....	119
<b>3.3 RESULTS.....</b>	<b>120</b>
3.3.1 The sequence coverage of the non-glycated CRP control .....	120
3.3.2 Tryptic digest of CRP preparations pre-exposed to 1 M glucose for 19 h, 3 and 6 days .....	121
3.3.3 Glu-C digest of CRP preparations pre-exposed to 1 M glucose for 19 h, 3 and 6 days .....	123
3.3.4 Interpretation of MS/MS spectra of the modified peptides .....	125
3.3.4.1 Glycation sites in CRP after 19 hours exposure to the 1 M glucose solution .....	125
3.3.4.2 Glycation sites in CRP after 3 days exposure to 1 M glucose solution .....	132
3.3.4.3 Glycation sites in CRP after 6 days exposure to 1 M glucose solution .....	134
3.3.4.4 Unmatched m/z values after Mascot search .....	138
3.3.5 Determination the most reactive lysine residue .....	139
3.3.6 Characteristics of glycated lysine residues.....	140
<b>3.4 DISCUSSION .....</b>	<b>143</b>
3.4.1 C-reactive protein and enzymatic digestion.....	143
3.4.2 Fragmentation behaviour of glycated peptides.....	144
3.4.3 Glycation sites of the in vitro glycated C-reactive protein .....	147
3.4.4 Factors affecting glycation site specificity .....	150
3.4.5 Conclusion .....	153
<b>CHAPTER 4: ANALYSIS OF HUMAN SERUM SAMPLES FROM DIABETIC PATIENTS .....</b>	<b>154</b>
4.1 INTRODUCTION.....	154
4.1.1 Overview .....	154
4.1.2 Diabetes mellitus.....	155
4.1.3 Monomeric C-reactive protein within the human body .....	156
4.1.4 Experimental aims.....	159



4.2 MATERIALS & METHODS .....	160
4.2.1 Reagents and consumables .....	160
4.2.2 Ethical Application .....	161
4.2.3 Recruitment of Participants .....	161
4.2.4 Sample size calculation .....	163
4.2.5 Acquisition of human serum samples .....	164
4.2.6 Production of monomeric CRP .....	165
4.2.7 Calibration of the size exclusion column .....	165
4.2.8 Purification of the human serum samples by affinity chromatography .....	168
4.2.9 Concentrating the eluted fractions .....	169
4.2.10 Separation of pentameric and monomeric CRP by size exclusion chromatography .....	170
4.2.11 SDS-PAGE of the size exclusion eluted fractions .....	171
4.2.12 Enzyme-linked Immunosorbent assay of the eluted serum fractions .....	172
4.2.13 Intact mass spectrometry of the eluted fractions .....	175
4.2.14 Bottom-up mass spectrometric analysis of pCRP containing samples .....	177
4.2.15 Statistical Analysis .....	179
4.3 RESULTS .....	180
4.3.1 Experimental design .....	180
4.3.2 Calibration of the size exclusion chromatography column .....	181
4.3.3 Purification of in vitro prepared monomeric C-reactive protein .....	185
4.3.4 Purification of the human serum samples by affinity chromatography .....	187
4.3.5 Isolation of CRP isoforms by size exclusion chromatography .....	189
4.3.6 Monitoring size exclusion chromatography by SDS-PAGE .....	192
4.3.7 Enzyme-linked immunosorbent assay of the eluted serum fractions .....	193
4.3.8 Intact mass analysis of mCRP in diabetic serum samples .....	200
4.3.9 Intact mass analysis of pCRP in human serum samples .....	201
4.3.10 Bottom-up tandem mass analysis of pCRP in human serum samples .....	202
4.4 DISCUSSION .....	204
4.4.1 Production and Purification of Monomeric C-reactive Protein in vitro .....	204

4.4.2 Purification and isolation of CRP isoforms from human serum samples .....	206
4.4.3 Enzyme-linked immunosorbent assay of the purified serum samples .....	207
4.4.4 Investigation of the glycated form of CRP using mass spectrometry.....	213
4.4.5 Conclusion .....	216
<b>CHAPTER 5: DISCUSSION, CONCLUSION AND FUTURE WORK.....</b>	<b>218</b>
5.1 OVERVIEW.....	218
5.2 PRODUCTION & CHARACTERISATION OF <i>IN VITRO</i> GLYCATED CRP .....	219
5.3 ANALYSIS OF HUMAN DIABETIC SERUM SAMPLES .....	223
5.4 CONCLUSION AND FUTURE WORK.....	226
<b>REFERENCES .....</b>	<b>228</b>
<b>APPENDIX 1.....</b>	<b>253</b>
<b>APPENDIX 2.....</b>	<b>254</b>
<b>APPENDIX 3.....</b>	<b>257</b>
<b>APPENDIX 4.....</b>	<b>259</b>

## List of Tables

Table 1-1 Functions of CRP according to FcγR class and their effector cells.....	8
Table 2-1 Reagents, solutions and proteins used in the study.....	58
Table 2-2 Components and amounts required for SDS-PAGE experiment. ....	61
Table 2-3 Buffers and solutions used for western blotting transfer and staining.....	62
Table 3-1 Summary of possible early and advanced glycation adducts.....	103
Table 3-2 Peptides identified in the tryptic digest of CRP samples incubated with 1 M glucose in the time course of the experiment.....	121
Table 3-3 Peptides identified in the Glu-C digest of CRP samples incubated with 1 M glucose in the time course of the experiment.....	123
Table 3-4 The Glycated peptides identified in the tryptic and Glu-C digests of the CRP samples incubated for 19 hours with a varying concentration of glucose. ....	140
Table 3-5 Physicochemical characteristics of CRP lysine and arginine residues .....	142
Table 4-1 Reagents, solutions and proteins used in the study.....	160
Table 4-2 A table of the calibration standards and their respective molecular weights. .	167
Table 4-3 A table of the components of the buffers required to purify CRP from human serum samples using the phosphocholine-bound affinity chromatography column. .....	169
Table 4-4 Table of the components of the buffers and solution used in the ELISA experiments for the pCRP and mCRP containing fractions. ....	172
Table 4-5 The amounts of pCRP (ng) from SCRIPPS used in the preparation of the ELISA calibration standards. ....	173

Table 4-6 The amounts of mCRP (ng), from 3 M urea dissociation experiment, used in the preparation of the ELISA calibration standards.....	174
Table 4-7 A table of the calculated Kav values for the molecular weight protein standards..	184
Table 4-8 Protein sequence coverage obtained after trypsin and Glu-C digestion of pCRP purified from diabetic serum samples.....	203

## List of Figures

Figure 1-1 Proposed physiological function of CRP.....	4
Figure 1-2 Schematic presentation of the activation of the complement system.....	6
Figure 1-3 The three-dimensional structure of C-reactive protein .....	10
Figure 1-4 The ligand recognition face of CRP subunit with phosphocholine and the calcium ions.. .....	12
Figure 1-5 Illustrative scheme represents calcium ions coordination with the surrounding amino acid residues .....	13
Figure 1-6 Surface representation of CRP showing the cleft viewed approximately from within the central pore of the CRP pentamer. ....	15
Figure 1-7 Initiation and progression of the atherosclerotic lesion .....	30
Figure 1-8 Protein glycation pathway.....	38
Figure 1-9 Reactive dicarbonyl species.....	39
Figure 1-10 AGEs Structures. ....	40
Figure 2-1 Schematic representation of the components of a mass spectrometer. ....	52
Figure 2-2 Schematic description of a TOF instrument equipped with a reflectron.....	54
Figure 2-3 Diagram of Quadrupole Mass Analyser.....	55
Figure 2-4 SDS-PAGE of <i>in vitro</i> glycated CRP Samples .....	69
Figure 2-5 Western blot of the <i>in vitro</i> glycated samples .....	70
Figure 2-6 Mass spectrum of native non-glycated CRP as measured by nESI-QTOF-MS in positive ion modes.....	71
Figure 2-7 The deconvoluted mass spectra of CRP samples incubated with different concentrations of glucose for 19 h. ....	74

Figure 2-8 The deconvoluted mass spectra of CRP samples incubated with different concentrations of glucose for 3 days .....	76
Figure 2-9 The deconvoluted mass spectra of CRP samples incubated with different concentrations of glucose for 6 days .....	79
Figure 2-10 The extent of glycation and the percentage of glycoform signal distributions of <i>in vitro</i> glycated samples .....	82
Figure 2-11 The G-idx of the <i>in vitro</i> glycated samples versus the incubation time. ....	83
Figure 3-1 Schematic presentation for the general steps in bottom-up proteomics.....	98
Figure 3-2 Peptide ion fragmentation in tandem mass spectrometry .....	105
Figure 3-3 Tandem mass spectrometry scan types .....	108
Figure 3-4 Fragmentation of Amadori compounds under CID conditions .....	111
Figure 3-5, Sequence coverage of the control CRP sample obtained from the tryptic and Glu-C digest.....	120
Figure 3-6 Assignment of product ion data for peptide sequence KAFVFPKESDTSYVSLK.	126
Figure 3-7 Assignment of product ion data for peptide sequence AFVFPKESDTSYVSLK...	127
Figure 3-8 Assignment of product ion data for peptide sequence GYSIFSYATKR. ....	128
Figure 3-9 Assignment of product ion data for peptide sequence ILIFWSKDIGYSFTVGGSE .....	130
Figure 3-10 Assignment of product ion data for peptide sequence FWVDGKPRVRKSLKKGTVGAE.. .....	131
Figure 3-11 Assignment of product ion data for peptide sequence ALKYEVQGEVFTKPQLWP. ....	133
Figure 3-12 Assignment of product ion data for peptide sequence YEVQGEVFTKPQLWP.. .....	135

Figure 3-13 Assignment of product ion data for peptide sequence LSSTRGYSIFS YATKRQDNE .....	136
Figure 3-14 Assignment of product ion data for peptide INTIYLGGPFSPNVLNWRALKYE.	137
Figure 3-15 Assignment of product ion data for peptide sequence ESDTSYVSLKAPLTKPLK. .....	139
Figure 3-16 Crystal structure of CRP highlighting putative glycation sites.....	148
Figure 4-1 A flow diagram of the methodology used for the analysis of human serum samples. ....	181
Figure 4-2 A size exclusion chromatography trace of the molecular weight standards ...	182
Figure 4-3 A size exclusion chromatography trace for Blue Dextran. ....	184
Figure 4-5 A Size exclusion chromatography trace of CRP treated with 3 M urea in the presence of EDTA for 10 weeks. ....	186
Figure 4-6 A chromatography trace of human serum passed through PC-bound affinity column. ....	188
Figure 4-7 Size exclusion chromatography trace of the proteins from human serum sample retained by the PC-bound affinity column.....	191
Figure 4-8 SDS-PAGE analysis of pCRP containing fractions after size exclusion chromatography separation. ....	194
Figure 4-10 Pentameric CRP concentrations purified from serum samples as determined by ELISA.....	194
Figure 4-11 A comparison of the pCRP concentrations obtained from the ELISA analysis of purified pCRP against the data provided from UHNM as whole serum CRP levels. .....	195
Figure 4-12 Calibration graph for mCRP ELISA assays. ....	196

Figure 4-13 Monomeric CRP concentrations purified from human serum samples as determined by ELISA.....	197
Figure 4-14 Scatter plot illustrating the correlation between the pCRP and mCRP levels purified from human serum samples as determined by ELISA. ....	199
Figure 4-15 Mass spectra of mCRP purified from human serum sample and measured by nESI-QTOF-MS in positive ion mode.....	200
Figure 4-16 The deconvoluted mass spectrum of pCRP isolated from a diabetic serum sample.....	202



## Table of Amino acids

<b>Amino Acid</b>	<b>Single Letter Code</b>	<b>Three Letter Code</b>
<b>Alanine</b>	A	Ala
<b>Cysteine</b>	C	Cys
<b>Aspartic acid</b>	D	Asp
<b>Glutamic acid</b>	E	Glu
<b>Phenylalanine</b>	F	Phe
<b>Glycine</b>	G	Gly
<b>Histidine</b>	H	His
<b>Isoleucine</b>	I	Ile
<b>Lysine</b>	K	Lys
<b>Leucine</b>	L	Leu
<b>Methionine</b>	M	Met
<b>Asparagine</b>	N	Asn
<b>Proline</b>	P	Pro
<b>Glutamine</b>	Q	Gln
<b>Arginine</b>	R	Arg
<b>Serine</b>	S	Ser
<b>Threonine</b>	T	Thr
<b>Valine</b>	V	Val
<b>Tryptophan</b>	W	Trp
<b>Tyrosine</b>	Y	Tyr

## Abbreviations Used

ACN	Acetonitrile
AGE	Advanced glycation end
AMI	Acute myocardial infarction
AP	Alternative pathway
APP	Acute phase proteins
BMI	Body mass index
CEL	Carboxyethyl lysine
CHD	Coronary heart disease
CI	Chief investigator
CID	Collision-induced dissociation
CML	carboxymethyl lysine
CP	Classical pathway
CRP	C-reactive protein
CVD	Cardiovascular diseases
DM	Diabetes mellitus
DTT	Dithiothreitol
DW	Deionised water
ECD	Electron-capture dissociation
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-Linked Immunosorbent Assay
ESI	Electrospray ionisation
ETD	Electron-transfer dissociation
F.A	Formic acid
FAS	Fractional solvent accessible surface areas
FPLC	Fast Protein Liquid Chromatography
FTICR	Fourier transform ion cyclotron resonance
HCD	Higher energy collisional dissociation
HDL	High-density lipoprotein
HPLC	High-performance liquid chromatography

HRP	Horseradish peroxidase
HTA	Human Tissue Authority
IAA	Iodoacetamide
IEF	Isoelectric focusing
IL	Interleukin
IPRC	Independent Peer Review Committee
IRAS	Integrated Research Application System
LC	Liquid chromatography
LDL	Low-density lipoprotein
LMW	Low Molecular Weight
LP	Lectin pathway
LPC	Lysophosphatidylcholine
LPS	Lipopolysaccharide
mAB	Monoclonal Antibody
MAC	Membrane attack complexes
MALDI	Matrix-assisted laser desorption/ ionisation
MBL	Mannose-binding lectin
mCRP	Monomeric C-reactive protein
ME	Margin of error
MG	Methylglyoxal
MP	Micro-particles
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MSA	multi-stage activation tandem mass analysis
MWCO	Molecular weight cut-off
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PC	Phosphocholine
PCF	Participant consent form
pCRP	Pentameric C-reactive Protein
PI	Isoelectric point

PIS	Patient information sheet
PMF	Peptide mass fingerprinting
PTM	Post-translational modifications
PVDF	Polyvinylidene difluoride
QqQ	Triple quadrupole
RA	Rheumatoid arthritis
RAGE	Receptor for AGE
REC	Research Ethics Committees
RF	Radio frequency
RT	Room temperature
SDS	Sodium dodecyl sulfate
SEM	Standard error of mean
TFA	Trifluoroacetic acid
TIC	Total ion chromatograms
TMB	3,3',5,5' tetramethylbenzidine
TNF	Tumour necrosis factor
TOF	Time of flight
UHNM	University Hospital of North Midlands
VADAR	Volume, Area, Dihedral Angle Reporter

## Acknowledgements

It is difficult to know from where to begin when it comes to the writing of the acknowledgements section of the thesis, as not only have so many assisted during the research and completing the thesis writing, but also all the others who have provided the support, guidance and motivation in the past four years.

I should perhaps first thank my sponsor, The Higher Committee for Education Development in Iraq who provided me with a postgraduate scholarship without which, quite simply, I would not be writing these words now.

I obviously owe many thanks to Dr Annette Shrive my supervisor, Professor Trevor Greenhough and Dr Sarah Hart my co-supervisors, and Dr Frederic Tripet my adviser, for their combined help, advice and encouragement. Annette, I am truly grateful for your scientific guidance, encouragement and supports through this work. You gave me the space I needed to accomplish my work but were around when I needed your help. Thank you for believing in me, for the motivation which always reached me when I needed it most. Sarah, when I started my PhD I had but a vague idea about mass spectrometry, your mentoring and guidance along the way has been invaluable. You have been a continuous source of valuable advice, encouragement and motivation; not only in scientific matters.

To the past and present members of the Structural Biology Research group, thanks for the friendship and always being ready to help when needed. I am very grateful for the excellent help from Rob for the knowledge and expertise he shared with me, special thanks for Ian Burns, for his experimental help and advice. Many thanks as well for Jenny for help with ELISA assay.

There are so many thank go to Professor Tony Fryer, members of R&D and the pathology department at UHNM their help, assistance, and extensive clinical knowledge.

I am infinitely thankful to my brothers, sister and my parents for giving me their unconditional love, support and encouragement, lastly to Saba, my wife and best friend, I do not know what I have done to deserve your love and unqualified support. You are always in my corner, thanks for everything. I cannot thank you enough.



## Chapter 1: Introduction

### 1.1 Human C-reactive protein

#### 1.1.1 C-reactive protein and innate immunity

Innate immunity is the first line of host defence against invading microorganisms and tissue injury, it plays a crucial role in the activation of acquired immunity and the maintenance of tissue repair. The innate immune system has a cellular and a humoral arm and comprises all mechanisms that do not have immunological memory including: physical barriers, phagocytes, inflammatory mediators releasing cells and molecules such as pathogen recognition receptors, complement proteins, acute phase proteins, and cytokines. All these mechanisms serve to prevent pathogens infections, eliminate damaged host cells, and initiate the inflammatory process (Cray *et al.*, 2009). Later aspects of induced innate responses are dependent on cytokines and chemokines, which are generated by activated cells (as monocytes, macrophages, endothelium, and T-cells). These proinflammatory mediators, including interleukin-6 (IL-6), IL-1 and tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ), have numerous effects throughout the body including initiating the acute phase response which is a core part of the innate immune response (Cray *et al.*, 2009). One of the most interesting features of the acute phase response is the change in the concentrations of many plasma proteins, known as the acute phase proteins (APP) which are proteins whose plasma concentration increases or decreases by at least 25% during inflammation (Cray *et al.*, 2009).

C-reactive protein (CRP) was the first APP described (in the early 1930s). It is of significant interest in humans as its plasma concentration may increase up to 1000 fold in response to infection or inflammation. Therefore, it is extensively used as a marker of disease activity



(Black *et al.*, 2004). CRP is a part of the humoral arm of the innate immune system and plays a vital role as a pattern recognition molecule, it involved in the differentiation of self (host molecules) from non-self molecules (antigens) by recognising pathogen-associated molecular patterns (CRP ligand on pathogen surfaces). However, CRP may also have impacts on adaptive immunity (Janeway & Medzhitov, 2002, Szalai, 2002).

### 1.1.2 Historical background

In 1930, William S. Tillett and Thomas Francis noticed a precipitation reaction in the serum of patients acutely infected with pneumococcus pneumonia when pneumococcal C-polysaccharide fraction was added to it (Tillett & Francis, 1930). The precipitation reaction did not occur in serum obtained from patients who had recovered from the disease, nor in clinically healthy individuals. The researchers further demonstrated this reaction occurred in the serum of patients in the early phases of certain acute infectious diseases, such as bacterial endocarditis and acute rheumatic fever (Tillett & Francis, 1930). In 1933, Ash confirmed and extended the observations of Tillett and Francis to involve diseases resulted from gram-negative bacteria. Ash used the term "C-reactive substance" to describe the substance responsible for the precipitation reaction with the pneumococcal polysaccharide C fraction in the serum (Ash, 1933). Several years later, Abernathy and Avery (1941) established that the so-called C-reactive substance was a protein, and they were the first who used the phrase "C-reactive protein" in describing its properties, they also illustrated that the precipitation reaction with C-polysaccharide of the pneumococcus pneumonia was a calcium-dependent reaction (Abernathy & Avery, 1941). In 1947, human C-reactive protein (CRP) was first crystallised, whilst the first report of the medical applicability of a CRP test in monitoring rheumatic activity in rheumatic patients appeared in 1950 (Anderson & McCarty, 1950, McCarty, 1947). In 1955, Roantree and Rantz stated that the

CRP test rarely gave a false positive result, and the most important feature of the test was that there was no need for interpretation of the normal range, as any positive reaction was considered abnormal (Roantree & Rantz, 1955). Several years later, Parker *et al.* (1962) concluded that CRP is of limited importance in viral infections (Parker *et al.*, 1962). In 1971, Volanakis and Kaplan identified the ligand for CRP in C-polysaccharide to be a phosphocholine moiety in teichoic acid which is part of the pneumococcal cell wall (Volanakis & Kaplan, 1971). In the 1990s, investigators recognised that inflammation contributes to the development of cardiovascular diseases (CVD) and suggested the possibility of using CRP as a marker to predict the risk for development of CVD in patients (Lagrand *et al.*, 1999; Ridker, 1998; Ross, 1993).

### 1.1.3 Biological functions of C-reactive protein

The primary biologic functions of CRP are determined by its ability to recognise pathogens and damaged cells of the host and to mediate their elimination by activating the complement system and phagocytic cells (Szalai, 2002). The capacity of human CRP to recognise and opsonise pathogens and damaged cells is mediated by the ligand-binding sites that present on the recognition face of the molecule; this face attaches with high avidity to the ligand decorated surfaces, figure 1-1 (Pepys & Hirschfield, 2003). After binding the opposite side of the molecule (the effector face) can bind C1q (a subunit of the C1 enzyme complex that activates the complement system) or possibly Fc-gamma receptors (FcγRs, a group of receptors found on the surface of certain cells and bind IgG and other ligands to induce phagocytosis of the opsonised cell ; figure 1-1; Pepys & Hirschfield, 2003). Activation of complement system via binding to C1q, and the binding to FcγRs are the two functions that assumed to be mediated by the effector face of the ligand-complexed CRP molecule (Du Clos & Mold, 2011). The ability of CRP to increase the

clearance of host apoptotic or/and necrotic cells and nuclear antigens suggests another function, prevention of autoimmunity (Du Clos & Mold, 2011).

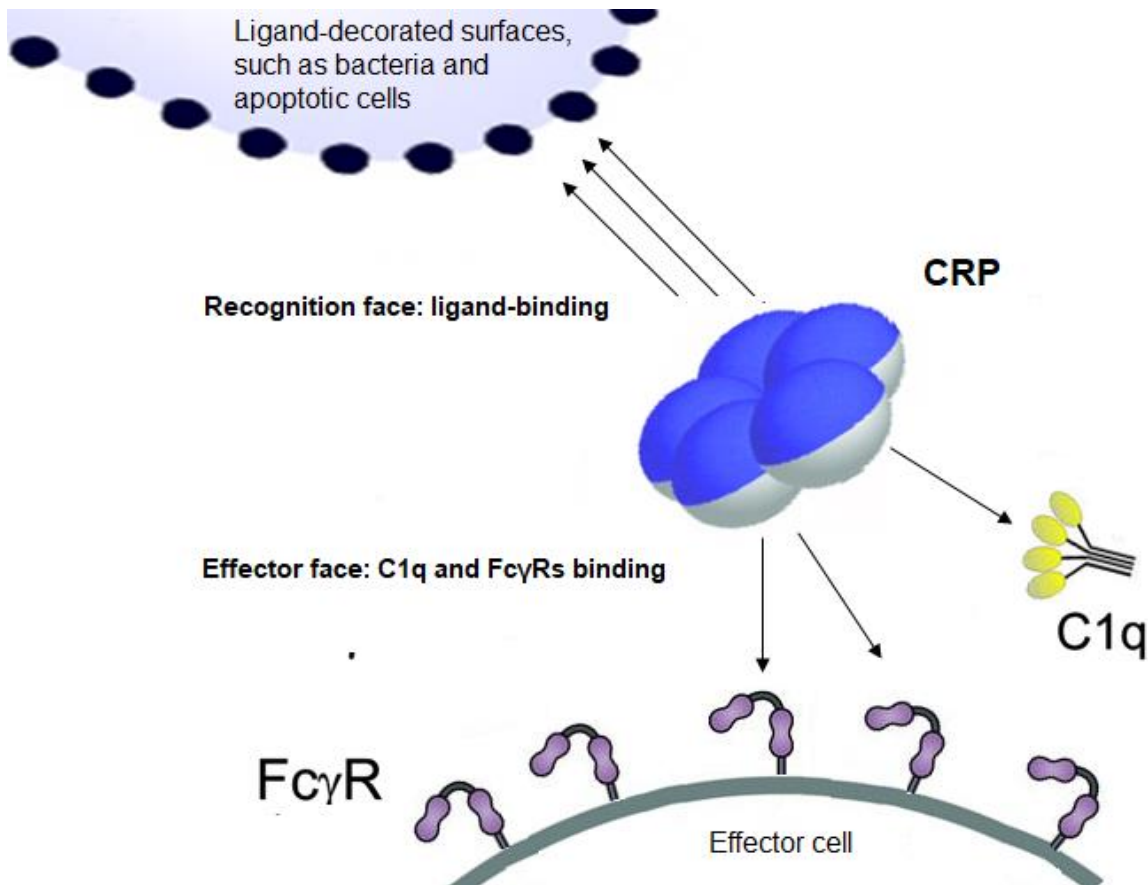


Figure 1-1 Proposed physiological function of CRP. One face of the CRP binds to ligand-decorated surfaces, the other face binds C1q to activate the complement system or to FcγR to facilitate phagocytosis; reproduced from Hage & Szalai, 2007.

### 1.1.3.1 Activation of the complement system

The complement system is a set of plasma proteins that can be activated either by pathogens or by a pathogen-bound antibody, leading to a cascade of reactions that occurs on the surface of pathogens, and generates active components with various effector functions (Janeway *et al.*, 2005). Complement can be activated via three pathways: the

classical pathway (CP), the lectin pathway (LP), and the alternative pathway (AP); (Janeway *et al.*, 2005). In all three pathways, the major complement component C3 is cleaved and activated by C3-convertase into C3a and C3b; see figure 1-2 (Janeway *et al.*, 2005). The CP is initiated by C1q binding to antigen-antibody complexes or other molecules (including  $\beta$ -amyloid and CRP; Orsini *et al.*, 2014). After the binding, the C1r and C1s proteases cleave the C4 and C2, which subsequently result in the formation of the C3-convertase (C4b2b) as shown in figure 1-2; (Orsini *et al.*, 2014). The LP is initiated by mannose-binding lectin (MBL), ficolins, or collectin-11 binding to carbohydrates or other molecules. After the binding, the MBL-associated serine proteases (MASPs), MASP-1 and MASP-2, cleave C4 and C2 forming the same C3-convertase as the CP; see figure 1-2 (Orsini *et al.*, 2014). The C3-convertase generated from CP or LP cleaves the C3 into C3a, which acts as potent anaphylatoxin, and C3b, which binds to the complex forming the C5 convertase (C4b2a3b) (Orsini *et al.*, 2014). The activation of the alternative pathway occurs by spontaneous hydrolysis of C3 to form the C3 (H<sub>2</sub>O). This molecule then associates factor B and factor D to create the AP C3-convertase (C3bBb) similar to that generated from CP and LP which in turn splits the C3 into C3a and C3b, with the latter creating a new C3 convertase (AP amplification loop, dotted line in figure 1-2) and creating more C3bBb3b complex (AP C5 convertase) (Orsini *et al.*, 2014). Increased levels of C3b results in the formation of the C5-convertase, which cleaves the C5 into C5a, a powerful anaphylatoxin, and C5b. The C5b binds to the cell surface and interact with C6, C7, C8 and C9 to generate the complex C5b-9 which is called the membrane attack complexes (MAC); see figure 1-2 (Orsini *et al.*, 2014).

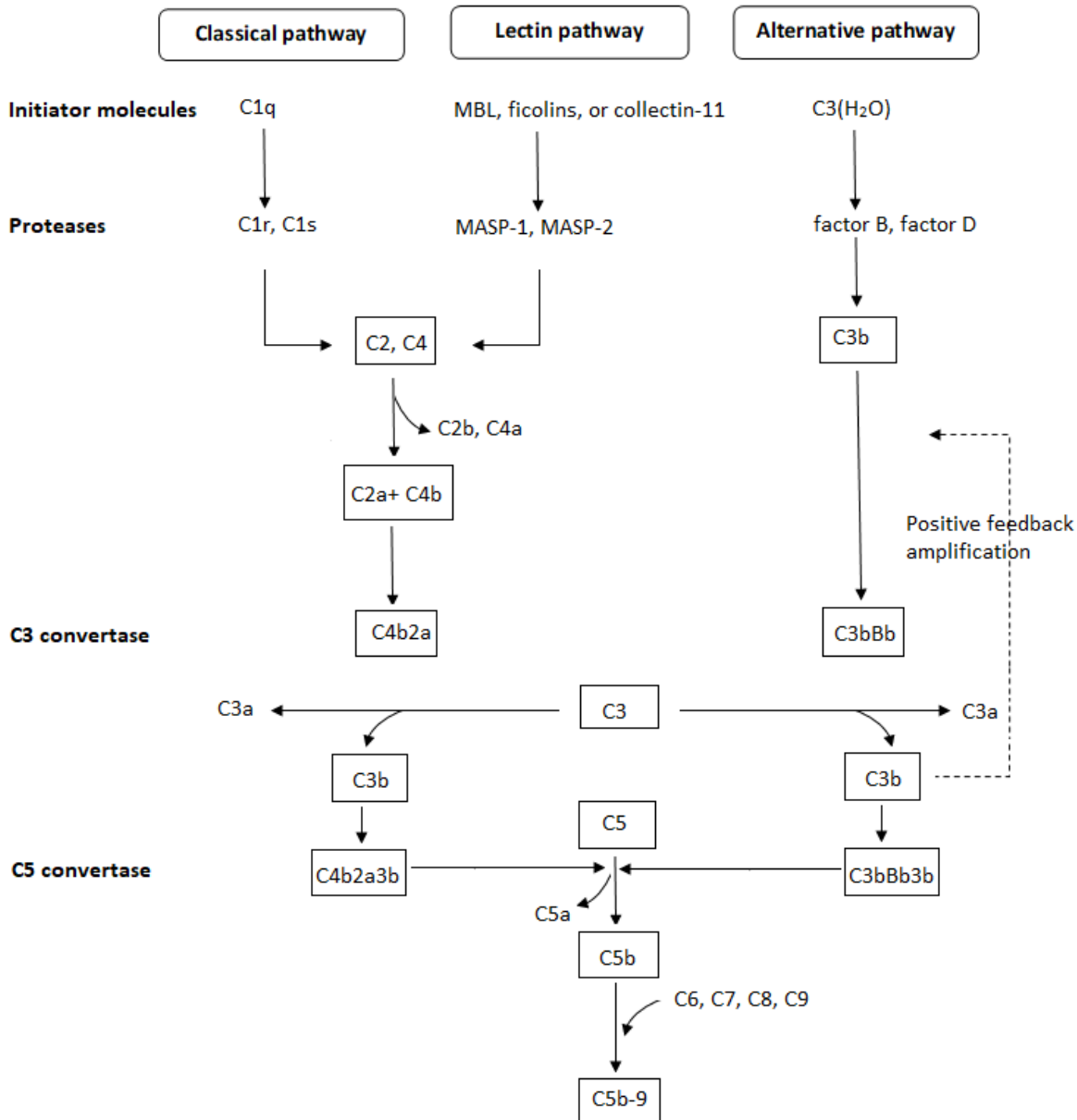


Figure 1-2 Schematic presentation of the activation of the complement system. The complement system can be activated via three pathways; the classical pathway, the lectin pathway, and the alternative pathway; Orsini *et al.*, 2014.

Complement plays a key role in both innate and acquired immunity. Its proper activation is essential for defence against pathogens and elimination of apoptotic and necrotic cells.

Nevertheless, excessive or inappropriate activation of this system contributes to the

pathogenesis of many chronic inflammatory diseases, including atherosclerosis (Salazar *et al.*, 2014).

CRP complexed with its ligand (ligand-CRP) on the surface of the pathogen or damaged host cell can activate the classical complement system, through recognition of this complex by C1q (Agrawal & Volanakis, 1994, Volanakis, 2001). The resulting ligand-CRP-C1q interaction leads to the formation of C3 convertase, which in turn leads to the decoration of the ligand surface with the opsonising C3 and C4 complement fragments (Volanakis, 2001). CRP-mediated activation appears to be limited to this initial stage involving C1-C4 (i.e. classical complement pathway), with limited formation of C5 convertase and MAC, due to inhibition of the alternative complement pathway (Agrawal, 2005). It has been postulated that the activation of the classical complement pathway and attenuation of the alternative complement pathway by CRP is due to the direct binding of CRP to the Factor H, an alternative complement pathway inhibitor (Sjöberg *et al.*, 2006, Jarva *et al.*, 1999). Activation of the classical complement pathway by ligand-bound CRP, therefore, mediates the anti-inflammatory functions of CRP (such as phagocytosis of apoptotic cells and protection from bacterial infections) but not the proinflammatory functions that require C5 convertase (Agrawal, 2005, Black *et al.*, 2004, Du Clos, 2013).

#### *1.1.3.2 Binding to FcγR receptors*

The effector face of CRP is also assumed to bind FcγRs on the surface of numerous cell types including macrophages, leukocytes, mast cells and platelets (Salazar *et al.*, 2014). FcγRs plays a co-ordinating role in immune responses; they are essential in phagocytosis, antibody-dependent cellular cytotoxicity, antigen presentation and inflammation (Rosales & Uribe-Querol, 2013). There are three families of these receptors, FcγRI, FcγRII, and FcγRIII

(Rosales & Uribe-Querol, 2013). Pentameric CRP (pCRP) primarily binds to the FcγRIIIa (CD32) (Bharadwaj *et al.*, 1999) and to a lesser extent to the FcγRI (CD64), whereas monomeric CRP (mCRP) binds FcγRIII (CD16), while FcγRIIb (CD32b) recognises both pCRP and mCRP (Salazar *et al.*, 2014). The differences in the expression and function of the FcγR isoforms may explain the diverse activities of CRP as shown in table 1-1 (Salazar *et al.*, 2014). Binding of CRP to FcγR receptors activates the effector cells, resulting in phagocytosis of CRP-opsonised particles and the release of inflammatory cytokines such as IL-1 $\alpha$ , IL-6 and TNF- $\alpha$  (Salazar *et al.*, 2014).

Table 1-1 Functions of CRP according to FcγR class and their effector cells. Adapted from Salazar *et al.*, 2014.

Receptor-type	CRP isoform	Effector cell-Functions
<b>FcγRI (CD64)</b>	pCRP	Monocyte- induce the release of cytokines Macrophages- mediates phagocytosis
<b>FcγRIIIa (CD32A)</b>		Monocyte- induce the release of cytokine Macrophages- induce expression of lipoprotein lipase and mediate phagocytosis Platelets- inhibit platelet binding to neutrophil
<b>FcγRIIb (CD32B)</b>	pCRP or mCRP	Endothelial cells- inhibits bradykinin and insulin-mediated activation of eNOS
<b>FcγRIII (CD16)</b>	mCRP	Monocyte- induce the release of cytokines Macrophages- induce expression of lipoprotein lipase Platelets- promotes binding of neutrophils and Induces synthesis of IL-8 and MCP-1 Endothelial cells- induces expression of endothelial adhesion molecules

CRP can bind to these receptors even in the absence of its ligand. However, this binding will produce a relatively a limited release of TNF- $\alpha$ , IL1- $\beta$  and IL-6 from peripheral blood mononuclear cells (lymphocytes, monocytes or macrophages) while binding with ligand-

complexed CRP increase the plasma concentration of these cytokines significantly (Salazar *et al.*, 2014). In addition to phagocytosis and cytokine release, the interaction of CRP with peripheral blood mononuclear cells has been reported to induce tumoricidal activity, respiratory burst and hydrogen peroxide production (Volanakis, 2001).

#### 1.1.4 C-reactive protein ligands

Human CRP binds mainly to phosphocholine-containing ligands in a calcium-dependent manner (Volanakis & Kaplan, 1971). Phosphocholine is commonly found in phospholipids of cell membranes, native and modified plasma lipoproteins and complex phospholipids of capsular and somatic components of bacteria, fungi, and parasites (Thompson *et al.*, 1999). In healthy cell membranes, the phosphocholine residues are inaccessible to CRP, and the binding only occurs in cases of damaged and apoptotic cells when disruption of the cell membrane exposes the phosphocholine to CRP (Marnell *et al.*, 2005, Volanakis & Wirtz, 1979). Moreover, the binding affinity of CRP to phosphocholine-containing ligand increases when there are multiple phosphocholine residues on the membrane, as there are multiple phosphocholine-binding sites in each CRP pentamer (Szalai, 2002).

Other CRP ligands that depend on calcium for binding include histones, chromatin, small ribonucleoproteins, fibronectin (an extracellular matrix protein), laminin (basement membrane protein), and phosphorylated polysaccharide (Black *et al.*, 2004, Marnell *et al.*, 2005, Shrive *et al.*, 1996). CRP can also bind to polycationic species such as poly-L-lysine, poly-L-arginine, poly-L-ornithine, and protamine sulfate in a calcium-independent manner (Ablij & Meinders, 2002).



### 1.1.5 Structure of C-reactive protein

CRP is a member of short pentraxin family proteins which is characterised by a cyclic pentameric structure (Shrive *et al.*, 1996). CRP is made up of five identical globular subunits (protomers) as shown in figure 1-3, each subunit has 206 amino acids and molecular mass of 23 kDa (Shrive *et al.*, 1996). These amino acids form a single polypeptide chain that folded into two antiparallel  $\beta$ -strands with a flattened jelly roll topology; see figure 1-3 (Salazar *et al.*, 2014, Shrive *et al.*, 1996). The protomers are noncovalently associated and arranged in a symmetrical cyclical pattern around a central pore, to form a pentamer with an outside diameter of about 102 Å and a central pore diameter of 30 Å and protomer diameter of 36 Å (Volanakis, 2001).

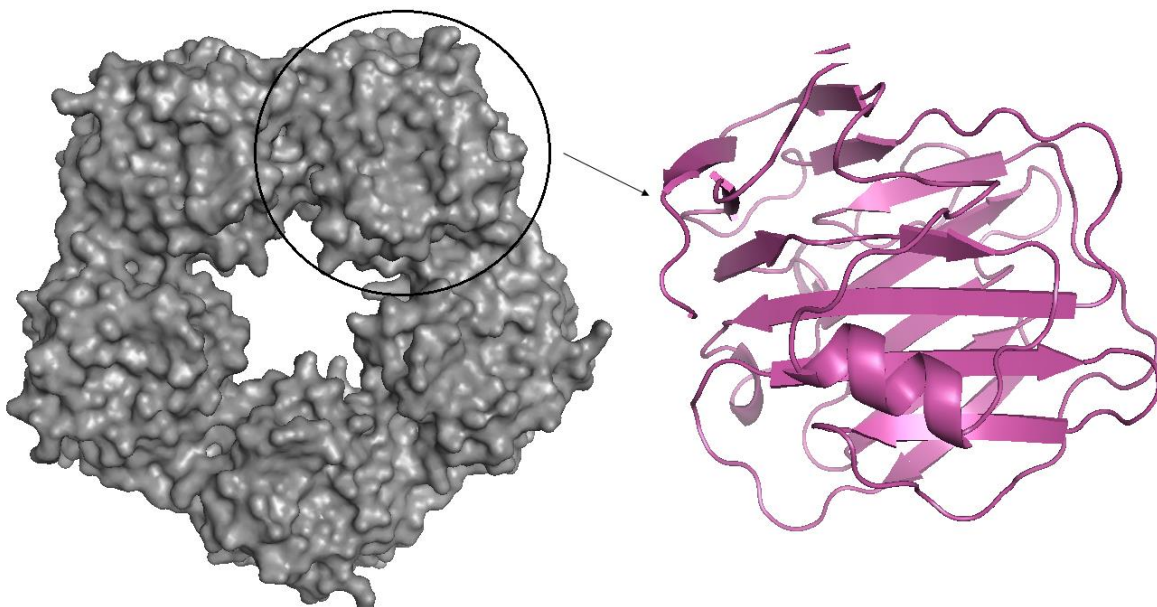
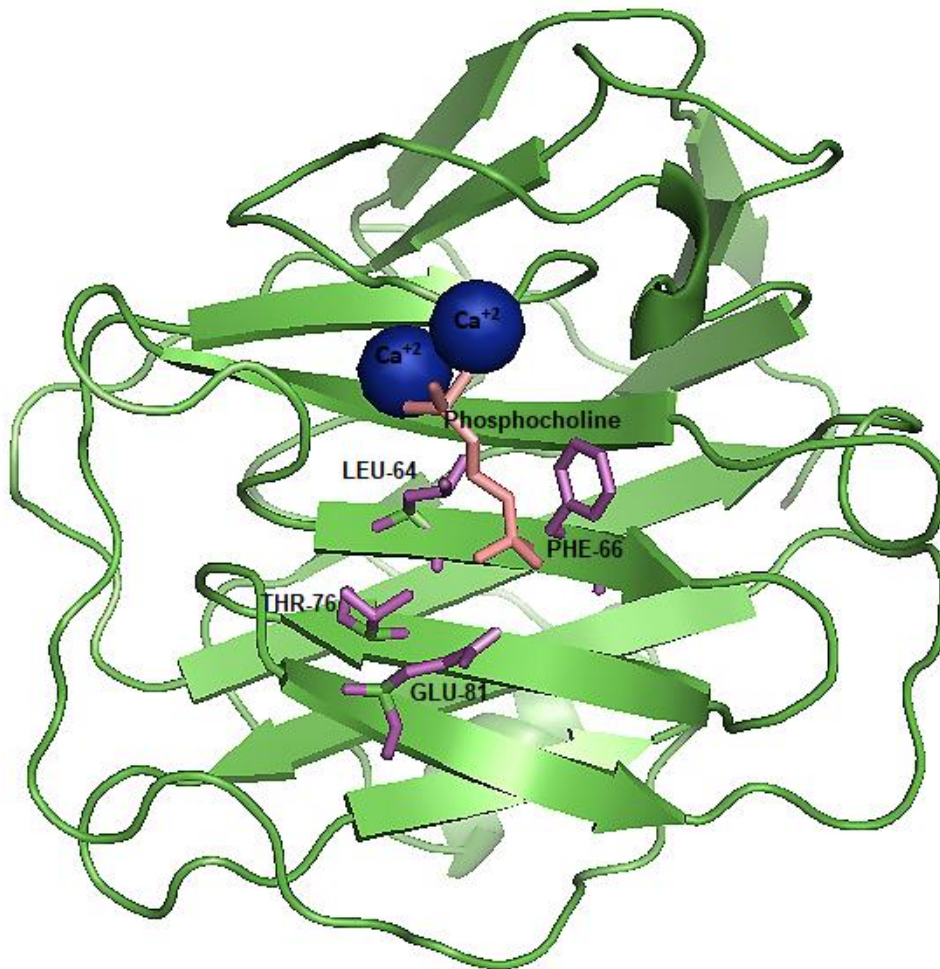


Figure 1-3 The three-dimensional structure of C-reactive protein. CRP consist of five identical protomers arranged around a central pore (left). The protomer made from single polypeptide chain folded into two antiparallel  $\beta$ -strands forming a flattened jelly roll topology (right). 1B09 (Thompson *et al.*, 1999) created with PyMol (DeLano, 2002).

The crystallographic data indicate that in the assembled CRP pentamer, all protomers have the same orientation (Shrive *et al.*, 1996), and the molecule has two faces: a 'recognition' face that binds to the ligands and an 'effector' face which mediate CRP functions (Marnell *et al.*, 2005, Szalai, 2002).

#### *1.1.5.1 The ligand recognition face of CRP*

The ligand recognition face of CRP binds ligands such as phosphocholine in the presence of calcium ions (Shrive *et al.*, 1996). Each CRP subunit has a single phosphocholine-binding site, and the assembled CRP pentamer can, therefore, bind with high avidity to surfaces that present multiple ligands in proximity (Culley *et al.*, 2000). The phosphocholine binding site of the CRP protomer has a hydrophobic pocket formed by residues Leu64, Phe66, and Thr76, in addition to Glu81 which is located on the other side of the hydrophobic pocket as shown in figure 1-4 (Shrive *et al.*, 1996). Adjoining to this pocket, there are two calcium ions close to each other, see figure 1-4 (Shrive *et al.*, 1996).



*Figure 1-4* The ligand recognition face of CRP subunit with phosphocholine and the calcium ions. The phosphocholine is displayed in raspberry and the calcium ions displayed as blue spheres. The residues Leu64, Phe66, Thr76 and Glu81 of the hydrophobic pocket are displayed in magenta. Two of the oxygens of the phosphate group of phosphocholine directly coordinates with the two calcium ions while the third oxygen group is pointed away. The opposite side of phosphocholine molecule, the methyl groups of choline interact with Phe66 residue, while the nitrogen of choline interacts with the side chain of Glu8. Image of 1B09 (Thompson *et al.*, 1999), created with PyMol (DeLano, 2002).

The first calcium ion is coordinated by residues Asp60, Asn61, Glu138, Asp140 and the main chain carbonyl of Gln139 (Shrive *et al.*, 1996), these residues provide one oxygen ligand with calcium ion (Shrive *et al.*, 1996). The second calcium ion is 4 Å distance from the first one and coordinated by Glu138, Asp140 and Gln150 and Glu147, these residues provide

one oxygen ligand with the calcium ion except for Glu147 which provide two coordination. Therefore, both calcium ion is equally coordinated with protein chelation being five each (Shrive *et al.*, 1996, Thompson *et al.*, 1999). Figure 1-5 illustrates the calcium ions coordination with the surrounding residues.

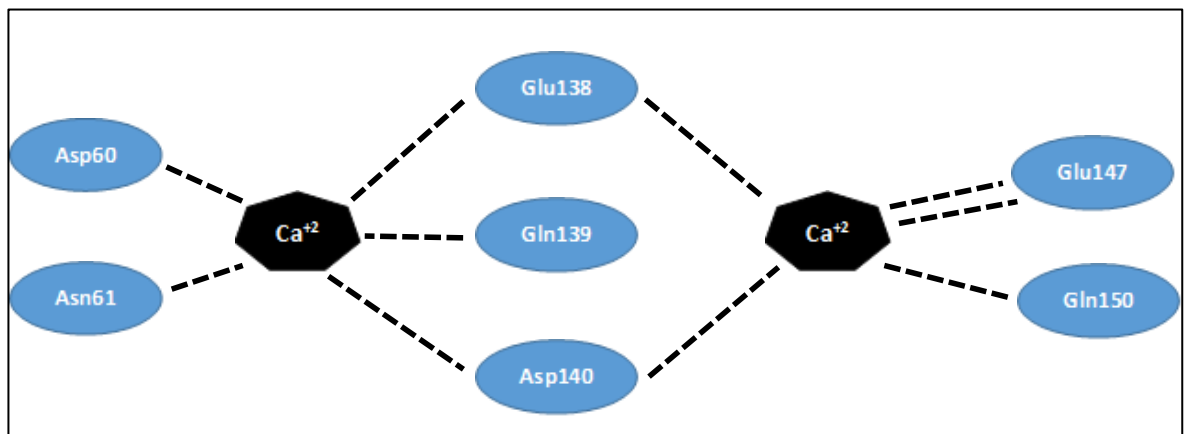


Figure 1-5 Illustrative scheme represents calcium ions coordination with the surrounding amino acid residues. Each dashed line represents single coordination.

A crystallographic analysis of CRP complexed with phosphocholine demonstrated that the choline moiety of phosphocholine lies within the hydrophobic pocket. The methyl groups of choline interact with the exposed face of the hydrophobic Phe66 residue, whilst the positively charged quaternary nitrogen of choline interacts with the side chain of Glu81. The opposite side of phosphocholine molecule, two of the oxygens of the phosphate group of phosphocholine directly coordinates with the two calcium ions whilst the third oxygen group is pointed away as shown in figure 1-4 (Black *et al.*, 2004, Shrive *et al.*, 1996, Thompson *et al.*, 1999). Mutational studies demonstrated the importance of the hydrophobic pocket in CRP-phosphocholine binding. The Phe66 residue appears to be the

major determinant of CRP-phosphocholine binding, whilst Thr76 plays a role in the topology of the hydrophobic pocket because its replacement with Tyr results in a significantly reduced avidity for phosphocholine-containing ligands, ostensibly because the bulky side chain of Tyr partially obstructs the pocket (Volanakis, 2001).

#### 1.1.5.2 *The Effector face of CRP*

Opposite the recognition face of CRP is the effector face, where complement C1q bind and FcγRs assumed to bind (Black *et al.*, 2004). This face is characterised by the presence of an  $\alpha$ -helix and a deep groove as shown in figure 1-6 (Shrive *et al.*, 1996). The  $\alpha$ -helix is formed from 168–176 residues, and found to be folded against one of the two  $\beta$ -sheets of the subunit. The groove is formed on one side by the carboxyl-terminal end of the  $\alpha$ -helix along with the 177–182 loop (figure 1-6), the other side of the groove is formed by parts of the amino and carboxyl termini of the protomer (Shrive *et al.*, 1996). The groove is deep and narrow at its origin in the centre of protomer, but becomes wide and shallow as it extends toward the pore of the pentamer (Shrive *et al.*, 1996). It has been speculated that the C1q-binding site and perhaps the FcγR binding site is associated with the wide shallow end of the groove (Volanakis, 2001).

Mutagenesis experiments have demonstrated that the Asp112 residue is the major determinant for recognition of C1q by CRP (Agrawal & Volanakis, 1994, Agrawal *et al.*, 2001). Asp112 is located within the shallow end of the groove toward the centre of the pentamer pore (see figure 1-6); therefore the five Asp112 residues of CRP pentamer will form a negative ring lining the pentamer pore (Thompson *et al.*, 1999). It is proposed that the positively charged head of C1q engage with the predominantly negatively charged central pore that formed by the five clefts present one on CRP pentamer. The C1q head

extends across the central pore of CRP and interacts with two of the five protomers of the pentamer. The optimal C1q binding is accompanied by slight conformational changes in the CRP structure, these conformational changes appear to differ depending on the ligand to which CRP is bound (Agrawal, 2005, Black *et al.*, 2004). The C1q unit has six globular heads suggesting that more than one CRP molecule is required to be in close proximity for complement activation, this may suggest that ligands decorated surfaces facilitate bringing of CRP pentamers close together to bind to the C1q heads (Agrawal, 2005).

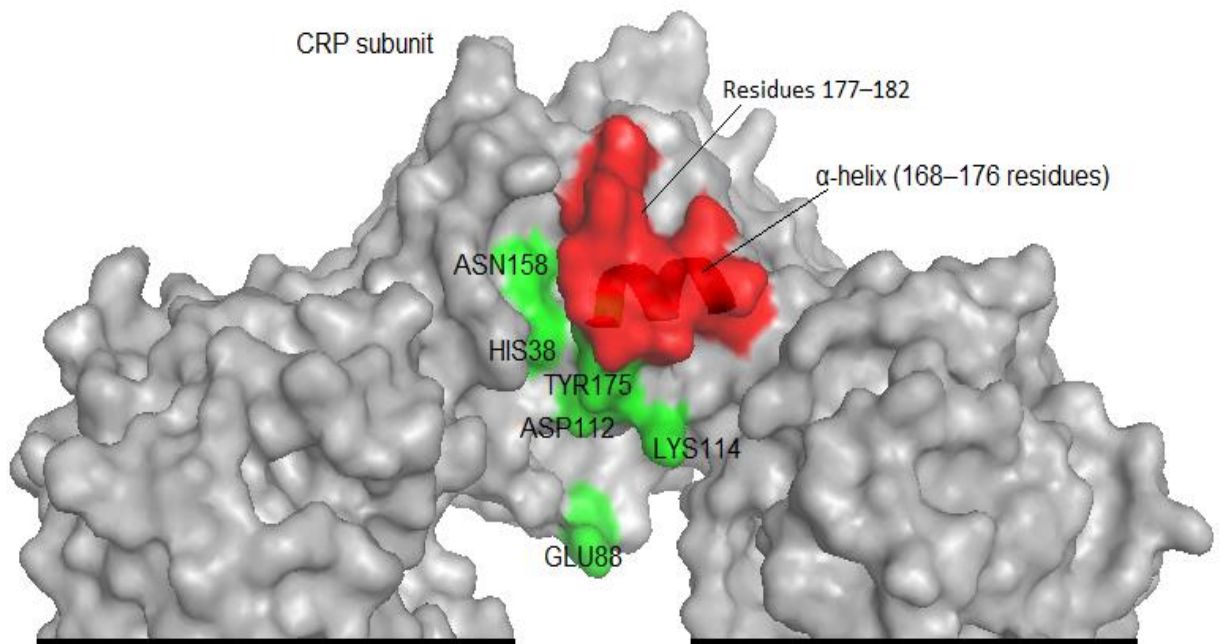


Figure 1-6 Surface representation of CRP showing the cleft viewed approximately from within the central pore of the CRP pentamer. The  $\alpha$ -helix and the 177-182 loop that forming one side of the groove are displayed in red. The residues involved in C1q binding are illustrated in green. Image of 1B09 (Thompson *et al.*, 1999), created with PyMol (DeLano, 2002).

In addition to Asp112, Tyr175 is also an important determinant for C1q binding to complexed CRP, and both of them (Asp112 and Tyr175) are assumed to be the contact

residue. Substitution of Tyr175 with Ala results in reduced or complete loss of C1q-binding activity depending on the ligand to which CRP was bound (Agrawal *et al.*, 2001). Glu88 appears to influence the conformational change of C1q necessary for activation of the complement, while Asn158 and His38 probably contribute to the geometry of the site (Agrawal *et al.*, 2001). In contrast, Lys114 appears to reduce binding of C1q to CRP because substitution of this residue with negatively-charged, neutral, or nonpolar amino acids resulted in greatly improved C1q binding and complement activation, this indicates that the extended positively-charged side chain of Lys114 interferes with the interaction between C1q and Asp112 (Agrawal & Volanakis, 1994). The configuration of the CRP groove is not changed by the presence or absence of calcium ions at the recognition face (Ramadan *et al.*, 2002). This in agreement with the findings that both calcium-independent CRP complexes and artificial trimers of CRP pentamers are capable of binding C1q (Agrawal *et al.*, 2001).

#### *1.1.5.3 The structural difference between calcium-bound and calcium-depleted CRP*

The presence of calcium in CRP protomers is not only crucial for ligand binding, but also increases the stability and maintains the integrity of the protein (Ramadan *et al.*, 2002). Early studies show that calcium-bound CRP resists denaturation caused by high temperatures and high urea concentrations (Heaton *et al.*, 1999). In calcium-bound protomers, the 140-150 loop is folded inward and held tightly against the main body of the protein, mainly by the Glu147 residue. Furthermore, the Arg47 residue further stabilises the 140-150 loop by forming two hydrogen bonds to the main chain carbonyls of Ser151 and Ser149, which surround calcium binding Gln150; this conformation protects the 145-147 site, a protease cleavage site, from proteolytic degradation (Ramadan *et al.*, 2002, Shrive *et al.*, 1996). In the absence of calcium, the 140-150 loop becomes mobile, and CRP

can be cleaved at two sites: Asn145-Phe146 by nagarse protease, and Phe146-Glu147 by pronase (Kinoshita *et al.*, 1989). The movement of the 140-150 loop away from the CRP protomer may also give rise to an increased binding affinity for polycations which are calcium-independent ligands (Ramadan *et al.*, 2002, Thompson *et al.*, 1999).

#### 1.1.6 Different forms of C-reactive protein

##### 1.1.6.1 Monomeric C-reactive protein

The native pentameric CRP (pCRP) can be dissociated irreversibly into monomeric CRP or modified CRP (mCRP) under conditions of heat, altered pH, and urea with calcium chelation (Potempa *et al.*, 1987). These modified CRP were found to run faster in gel electrophoresis, and were characterised by a reduced solubility, and lower isoelectric point than the native CRP (Potempa *et al.*, 1987). The mCRP expresses new epitopes (neoepitopes) which are typically hidden in the native pCRP, and therefore mCRP has different biological features (Ablij & Meinders, 2002, Potempa *et al.*, 1987).

There is now strong evidence to demonstrate that pCRP dissociation can take place under physiological conditions after binding to cell membranes, and the mCRP has been detected in various tissues throughout the body (Salazar *et al.*, 2014, Thiele *et al.*, 2015). The mCRP is commonly considered as a tissue-based protein rather than as a soluble circulating protein, it is seldom found in circulation by the common detection and quantification methods possibly because of its limited solubility (Salazar *et al.*, 2014, Thiele *et al.*, 2015). The pCRP dissociation has been observed on the membranes of apoptotic cells and activated platelets in atherosclerotic plaque, and the dissociation appears to be initiated by binding of pCRP to the exposed lysophosphatidylcholine (LPC) on the cell surface of injured tissue (Salazar *et al.*, 2014, Thiele *et al.*, 2015). LPC, which is produced by the action



of phospholipase A2 on cell membrane phosphatidylcholine, appears to be important in this process (Salazar *et al.*, 2014, Thiele *et al.*, 2015). The primary step in this dissociation is the formation of a hybrid intermediate molecule with partial structural change termed mCRPm, mCRPm possess mCRP antigenicity but retains the native pentameric conformation, this molecule (mCRPm) rapidly detaches from the cell membrane and finally dissociates into mCRP (Ji *et al.*, 2007, Salazar *et al.*, 2014). Independent mCRP synthesis is also another possible mechanism, CRP mRNA has been reported in numerous extra-hepatic tissues including adipocytes, smooth muscle cells, and inflammatory cells within atheromatous plaques. However, whether the synthesised CRP is assembled into pCRP or expressed as the mCRP isoform is still unclear (Salazar *et al.*, 2014). In fact, there is evidence that CRP mRNA in U937 macrophages of atherosclerotic lesions is processed into mCRP isoform rather than pCRP (Salazar *et al.*, 2014).

The mCRP receptors are not well characterised, but it has been suggested that mCRP can bind to the FcγRIIIb (CD16) receptor on the neutrophil surface. However, when the binding of mCRP to FcγRIIIb is inhibited by anti-CD16 antibodies, full inhibition of mCRP actions was not obtained suggesting the existence of additional receptors (Salazar *et al.*, 2014).

#### *1.1.6.2 Fibril-like C-reactive protein*

Freshly purified CRP in a neutral buffer solution (pH 7.4) form short strand fibrils after a few days of storage, probably by the face-to-face stacking of pCRP and electrostatic interactions between pCRP molecules; these strands continue to grow in length with increasing duration of storage (Wang *et al.*, 2002). Wang *et al.* (2002), reported fibril-like structure in both calcium-containing and calcium-free buffers, the ionic strength of

solutions appear to be a key determinant as the solutions with high ionic strength result in less fibril-like structures (Wang *et al.*, 2002).

#### 1.1.6.3 Glycosylated C-reactive protein

Das *et al.*, (2003), identified a glycosylated version of CRP in several pathological conditions including systemic lupus erythematosus, acute lymphoblastic leukaemia, tuberculosis, visceral leishmaniasis, Cushing's syndrome, osteogenic sarcoma, and osteogenic sarcoma. The glycosylated version of CRP was identified in the serum samples collected from patients with the above conditions by utilising digoxigenin kits, neuraminidase treatment and binding with lectins. The presence of N-linked sugar moiety was confirmed by N-glycosidase F digestion. The presence of sialic acid, glucose, galactose and mannose has been demonstrated by gas-liquid chromatography, mass spectroscopic and fluorimetric analysis (Das *et al.*, 2003).

The researchers were also demonstrated that the glycosylated CRP shows a variation in glycosylation pattern in each studied pathological condition, they also suggested two potential glycosylation sites on the CRP groove floor. The researchers demonstrated that these glycosylated CRP have variable binding characteristics with plasma glycoprotein and antibodies, including human IgG and anti-CRP antibody and plasma fibronectin, due to minute structural differences in the CRP as a result of glycosylation (Das *et al.*, 2003).

#### 1.1.7 CRP synthesis and gene expression

CRP is primarily produced by hepatocytes, and is mainly controlled at the transcriptional level through IL-6 signals, and enhanced synergistically by IL-1 $\beta$  signalling (Black *et al.*, 2004, Marnell *et al.*, 2005). The CRP gene located on the short arm of chromosome 1 (Black *et al.*, 2004). Both IL-6 and IL-1 $\beta$  control expression of several acute phase proteins through

activation of the transcription factors STAT3 (signal transducer and activator of transcription 3), Rel proteins, and C/EBP family members C/EBP $\beta$  and C/EBP $\delta$ . These transcription factors bind to response elements in the CRP promoter (Black *et al.*, 2004, Marnell *et al.*, 2005). Clinical trials demonstrated the importance of IL-6 as a principal inducer of CRP synthesis as the treatment with an anti-IL-6 receptor antibody cause a dramatic decrease in CRP levels (Lazzerini *et al.*, 2015, Nishimoto *et al.*, 2000). After transcription, CRP mRNA is translated to protomers that assemble to form CRP pentamer in the endoplasmic reticulum. The pentamers are largely retained in the endoplasmic reticulum-bound by two resident carboxylesterases. During the acute phase of the inflammatory response, the half-time for the release of CRP from the endoplasmic reticulum is largely accelerated. The pronounced acceleration of CRP secretion is apparently due to the reduced affinity of one of the carboxylesterases to CRP, thus result in increased serum CRP levels (Ablij & Meinders, 2002, Volanakis, 2001).

Apart from hepatocytes, CRP synthesis has been reported to occur in many extra-hepatic cell including: lymphocytes, macrophages, adipose tissue, smooth muscle cells, aortic endothelial cells, epithelial cells from nasal polyps, epithelial cells of the upper respiratory tract, and tubular epithelial cells in the kidney (Ablij & Meinders, 2002, Black *et al.*, 2004). The specific mechanisms by which the synthesis at these sites is regulated are not known. Although extra-hepatic CRP may mediate local effects, it is unlikely that it substantially affects the serum levels of CRP (Ablij & Meinders, 2002, Black *et al.*, 2004).

### 1.1.8 The serum concentration of C-reactive protein in health and disease

Many studies have reported the serum concentration of CRP in apparently healthy adults of the general population. Shine *et al.* (1981) studied 468 healthy adult volunteers using radioimmunoassay, the volunteers showed median CRP concentration of 0.8 mg/L, with 90% of the values being less than 3.0 mg/L and 99% being less than 10 mg/L (Shine *et al.* 1981). These results are similar to those obtained from a more recent study carried out on 13,000 healthy individuals across Europe, in which the median concentration ranged from 0.6-1.7 mg/L, and the 90th percentiles from 3.2-8.0 mg/L (Imhof *et al.*, 2003). Another study on 22,000 healthy adults from the United States show median values of less than 1.52 mg/L; and the 90th percentiles less than 6.61 mg/L (Rifai & Ridker, 2003). While studies on Asian population give a much lower result (Kao *et al.*, 2006, Yamada *et al.*, 2001). The slight difference in CRP values between populations may be due to genetic as well as environmental factors (physical activity, lifestyle, and dietary factors); (Kao *et al.*, 2006).

CRP is the most sensitive among the APPs. During the acute phase response to infection or inflammation, CRP concentration increases rapidly within the first 4-6 hours, doubling at least every 8 hours and reaching its peak after 36-50 h. With very intense stimuli, the CRP concentration can rise to more than 500 mg/L, i.e. 1000 times the average value; other APPs rarely increase by more than fourfold (Hogarth *et al.*, 1997, Jaye & Waites, 1997, Vigushin *et al.*, 1993).

The half-life of serum CRP is 19 hours independent of its concentration in serum and the pathophysiological circumstances that cause the elevation. Therefore the only determinant of circulating CRP concentration is the rate of synthesis, which in turn reflects the intensity of the pathological process that stimulated the CRP production (Vigushin *et al.*, 1993).

Therefore measuring CRP can provide information about the presence and intensity of an inflammatory process, and help to differentiate inflammatory from non-inflammatory conditions (Póvoa, 2002). If the stimulus that causes increased production completely ceases, the circulating CRP concentration will fall rapidly. However, CRP serum concentration can remain high, even for very long periods of times if the underlying pathology of the elevation persists (Vigushin *et al.*, 1993).

Elevation in serum concentration of CRP is seen in infections (bacterial, viral, and fungal infection), inflammation and tissue damage. In infectious diseases, the highest elevations of CRP concentrations are seen in bacterial infections, with levels exceeding 200 mg/L in case of severe conditions. In contrast, viral infections are less likely to cause a substantial elevation. Therefore, CRP can be used to differentiate between viral and bacterial infection (Chew, 2012). However, this generalisation is not always true, as sometimes uncomplicated infections with adenovirus, measles, mumps and influenza are associated with high serum CRP levels while some chronic bacterial infection as tuberculosis and leprosy, cause only a modest elevation (Póvoa, 2002). There is limited information about CRP response in parasitic infections, but some protozoan parasitic diseases such as malaria, pneumocystosis and toxoplasmosis characterised by a marked increase in serum CRP level (Póvoa, 2002). Serum CRP concentration is also markedly elevated in most of the inflammatory diseases and measuring CRP concentration in these diseases can give valuable information about the extent of activity and severity of disease state. However, the behaviour of CRP is not the same with all inflammatory diseases such as SLE, dermatomyositis, Sjögren's syndrome, or ulcerative colitis have a normal or only slightly elevated CRP despite the presence of a severe disease state. The reason for this contradiction remains speculative (Vermeire *et al.*, 2004). Even in inflammatory conditions that are not associated with elevated CRP levels,

the measurement may still be useful to differentiate between disease exacerbation and complication with infection, as these patients are still able to mount a major CRP response in the presence of infection (Póvoa, 2002).

There are several other conditions that lead to a substantial increase in CRP concentrations, these include trauma, surgery, burns, tissue necrosis, and advanced cancer (Gabray & Kushner, 1999, Vigushin *et al.*, 1993), even vigorous exercise and some psychiatric diseases are associated with mild CRP changes (Póvoa, 2002).

#### 1.1.9 Applications of CRP as a clinical marker

CRP has many analytical properties that support its use as a clinical marker, particularly, the rapid change follows the onset of the inflammation, the long half-life compared to other markers of inflammation, a broad range of abnormal values and availability of easy and sensitive methods for its measurement (Pepys & Hirschfield, 2003). CRP can be measured in plasma, serum and detected in various body fluids (including cerebrospinal, synovial, pleural and ascitic fluids), and there is no diurnal variation in its serum concentration; therefore the measurement can be done without concern for the time of collection (Pepys & Hirschfield, 2003).

Elevation of CRP is not specific of any particular disease as it happens in all diseases involving tissue damage or inflammation; however, because of its extreme sensitivity CRP can be used as an adjunct to clinical assessment in many situations including the followings (Aguiar *et al.*, 2013, Hirschfield & Pepys, 2003, Reeves, 2007):

1. Screening for the organic disease to support the clinical diagnosis
2. Monitoring response to treatment of infection and inflammation

3. Detection of intercurrent infection in inflammatory diseases characterised by modest or absent CRP response
4. Prediction in CVD.

#### 1.1.10 Methods of measurements

CRP levels have been measured utilising a variety of methods. The most common methods of measurement include; latex agglutination, immunonephelometry, immunoturbidimetry, and enzyme immunoassay (Deodhar, 1989, Urdal *et al.*, 1992). These immunological methods of measurement depend on the formation of antigen-antibody complexes when anti-CRP antiserum reacts with its antigen (CRP in patient's serum). These complexes (which proportionally reflect the concentration of CRP) can scatter incident light; in turn, this scattering of the light can be measured by different methods including turbidimetry (measuring the reduction of light passing through a reaction mixture) or nephelometry (measuring the light scattered by a reaction mixture); (Deodhar, 1989, Urdal *et al.*, 1992). Nephelometry is the most widely used method as it is stable, reproducible, rapid (results can be obtained in less than 30 min), sensitive ( $\leq 0.04$  mg/L) and relatively inexpensive (Enguix *et al.*, 2001).

There are two major types of CRP assay; standard or high-sensitivity CRP (hs-CRP) assay. The standard assay has a limit of detection of approximately 10 mg/L and can be used for any clinical condition with a serum CRP concentration above this limit. The hs-CRP assay range is from  $< 1.0$  mg/L to  $\leq 10.0$  mg/L; this assay can be used for an apparently healthy individual with serum CRP concentrations of less than 10 mg/L (Faraj & Salem, 2012). The advent of the hs-CRP assay led to numerous studies that showed utility in individuals at risk

for cardiovascular disease, metabolic syndrome, and other chronic diseases associated with a low level of inflammation (Du Clos, 2013).

#### 1.1.11 Demographic, behavioural and lifestyle factors that affect baseline CRP

Baseline values of CRP are influenced by a wide variety of factors including: age, obesity, smoking, and other factors. Increasing age was associated with increased serum level of CRP in several studies of different gender groups. A large population study, including individuals aged from 25 to 74 years by Hutchinson *et al.* (2000) found the median CRP concentration is approximately doubled with age, from ~1 mg/L in the younger age groups to ~2 mg/L in the older age group and the levels tended to be higher in females (Hutchinson *et al.*, 2000). Further evidence from a 20 year follow-up study in men aged 60-79 showed that 57% of the median serum CRP levels increased between the 60-64 year age group and the 75-79 year age group (Rumley *et al.*, 2006). Caparevic and Kostic also showed that healthy women over 56 years have higher levels of serum CRP compared with healthy women younger than 45 years of age (Čaparević & Kostić, 2007).

Many studies have demonstrated a strong positive correlation between serum CRP concentration and obesity (Timpson *et al.*, 2005, Visser *et al.*, 1999). As obesity is associated with a higher CRP concentration compared to lean control subjects, weight reduction results in a decrease in CRP concentration regardless of the type of intervention, with a roughly linear relationship (Dietrich & Jialal, 2005, Selvin *et al.*, 2007). Elevated serum CRP levels are associated with abdominal obesity, which is a measure of general obesity, independent of body mass index (BMI). CRP concentration is higher in individuals with abdominal obesity, compared to control subjects with the same BMI (Lapice *et al.*, 2009). It has been hypothesised that a chronic low-grade inflammatory state is associated with



obesity; however, the origin of this inflammation and the underlying molecular mechanisms that explain its occurrence are not completely understood. Pro-inflammatory cytokines are postulated to play a central role (Rodríguez-Hernández *et al.*, 2013). Infiltrating macrophages in expanded adipose tissue and adipocyte secrete a number of pro-inflammatory cytokines including IL-6 and TNF- $\alpha$ . It found that adipocytes may produce as much as one-third of the total IL-6 in circulation. IL-6, in turn, activates CRP expression from hepatocytes, as a consequence, obese individuals tend to have slightly elevated serum CRP levels (Rodríguez-Hernández *et al.*, 2013).

Smoking also results in a high serum CRP concentration as evidenced by many studies, for example, a study of 2,920 British men aged 60-79 showed that current smokers have significantly higher CRP levels (2.53 vs 1.35 mg/L) than to non-smokers (Wannamethee *et al.*, 2005). The level of serum CRP in smokers is dose-dependent, such that CRP concentration is positively correlated with the number of cigarettes smoked per day (Melbye *et al.*, 2007, O'Loughlin *et al.*, 2008). Furthermore, the relationship between smoking and the high CRP concentration appears to be temporal, as smoking cessation for long periods ( $\geq 5$  years) can significantly improve CRP concentration (Ohsawa *et al.*, 2005, Wannamethee *et al.*, 2005). This finding suggests that the risk reduction of CVD by smoking cessation could be reflected by the reduction in CRP level (Ohsawa *et al.*, 2005).

Other factors that may cause a small elevation in CRP levels include ethnicity, low socioeconomic status, high fat and low fibre consumption, hormonal replacement therapy, pregnancy, poor physical fitness or low levels of physical activity (Kushner *et al.*, 2006, Nanri *et al.*, 2007).

### 1.1.12 C-reactive protein as a low-grade inflammatory marker

Historically, only serum CRP values above 10 mg/L were considered clinically important (Morley & Kushner, 1982), today a mildly elevated serum CRP levels higher than those in most normal subjects is now believed to represent low-grade inflammation, and may underlie chronic complex diseases such as diabetes, dyslipidaemia, atherosclerosis, and metabolic syndrome. Recent studies have shown the importance of minor CRP elevation in the prediction of CVD. Ridker and colleagues have suggested clinical categories for baseline serum CRP values in predicting the risk of future cardiovascular events (Ridker, Paul M. *et al.*, 2003).

- Low-risk group: less than 1.0 mg/L
- Intermediate risk group: 1.0 to 3.0 mg/L
- High-risk group: above 3.0 mg/L

Thus, CRP is now also considered as a marker of low-grade inflammation alongside its traditional role as an acute phase protein.

### 1.1.13 CRP and cardiovascular disease

#### 1.1.13.1 Evidence from epidemiological studies

Since the 1990s, several epidemiological studies have shown a link between the circulating CRP and incidence of a variety of CVD (Lagrand *et al.*, 1999). Arima *et al.* (2008) followed 2589 adult aged 40 years and older for 14 years in a population-based prospective cohort study. The researchers found that the risk of coronary heart disease (CHD) in the highest CRP quartile group was three times higher than that in the lowest CRP group even after adjusting the other cardiovascular risk factors such as age, gender, BMI, smoking habits,

alcohol intake and regular exercise (Arima *et al.*, 2008). Zieske *et al.* also found that elevated CRP levels are independently associated with atherosclerosis in young people (Zieske *et al.*, 2005). Furthermore, Williams *et al.* (2008) showed that the elevated CRP levels in patients with CHD can predict hospitalisation due to heart failure independent of a prior history of heart failure, medication use, and severity of CHD; therefore CRP can add prognostic information and predict morbidity and mortality in CHD patients. Ridker *et al.*, (2002), stated that serum CRP was the stronger predictor of future cardiovascular events (myocardial infarction, ischaemic stroke, or death from cardiovascular causes) than plasma low-density lipoprotein cholesterol (LDL-C). Ridker *et al.*, (2002), followed 27,939 healthy women for eight years for the occurrence of the first cardiovascular events. They found that the relative risk of a first cardiovascular event according to increasing quintiles of CRP level as compared to the lowest quintile, were 1.4, 1.6, 2.0, and 2.3, whereas the corresponding relative risks in increasing quintiles of LDL-C level compared with the lowest were 0.9, 1.1, 1.3, and 1.5 (Ridker, *et al.*, 2002).

CRP is not only a predictor of CVD, but also can be a therapeutic target, as demonstrated by JUPITER trial (Justification for the Use of Statins in Primary Prevention), which showed a significant reduction in the incidence of major cardiovascular events through lowering serum CRP levels by using statin therapy (Rosuvastatin) among people without hyperlipidemia. This finding emphasises the vital role that CRP plays in CVD prevention (Ridker, 2003). Some researchers in the field consider that CRP is more than just an inflammatory marker, but also a direct cause of CVD (Li, Jian-Jun & Fang, 2004, Stumpf & Hilgers, 2009).

### *1.1.13.2 Role of CRP in atherosclerosis and atherothrombosis*

Atherosclerosis is a multifactorial disease affecting the endothelial lining of medium and large-sized arteries, and usually develops several years before the manifestation of any clinical symptoms. The main risk factors for atherosclerosis include dyslipidaemia, smoking, hypertension and diabetes mellitus, all these factors are associated with low-grade inflammation (Packard & Libby, 2008). In atherosclerosis, the LDL-C which is transmigrated across the vascular endothelium and deposited in the intima undergo oxidation by local reactive oxygen species (Heinecke, 2006). Oxidised LDL (Ox-LDL) stimulates endothelial cells to express adhesion molecules (vascular cell adhesion molecule 1 (VCAM-1), intercellular adhesion molecule 1 (ICAM-1) and E-selectin), cytokines (IL-1, IL-8, TNF- $\alpha$ ) and chemokines (monocyte chemoattractant protein 1 (MCP-1)). These molecules make vascular endothelial surfaces 'sticky', so that white blood cells such as monocytes adhere to the endothelium. Once the monocytes adhere to the activated endothelium, they migrate through it into the subendothelial space. Within the intima, the monocytes mature into macrophages where many of them can differentiate into macrophages, which express scavenger receptors enabling them to ingest Ox-LDLs, and subsequently transform into foam cells as shown in figure 1-7 (Heinecke, 2006, Packard & Libby, 2008). Foam cells accumulate and form 'silent' fatty streaks. If the process continues, smooth muscle cells from the vascular intima can migrate into the subendothelial intima, proliferate and form a fibrous cap which surrounds the atherosclerotic plaque (see figure 1-7). The plaque may grow large enough to reduce the blood flow or the fibrous cap may weaken and rupture, exposing the underlying thrombogenic tissues causing the contents of the plaque to spill out into the lumen of the blood vessel, impeding blood flow, and resulting in severe clinical manifestation according to the site of occlusion (Packard & Libby, 2008). The clinical

manifestation may be in the form of angina pectoris, myocardial infarction, sudden cardiac death, cerebrovascular disease (transient ischaemic attacks, stroke) or peripheral vascular disease.

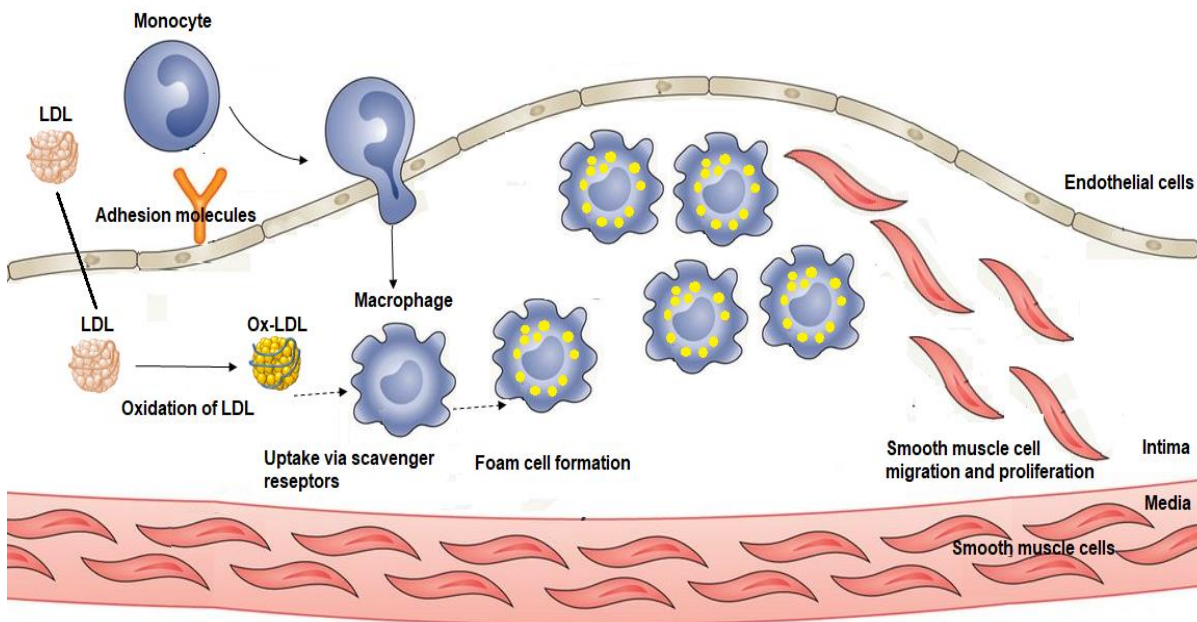


Figure 1-7 Initiation and progression of the atherosclerotic lesion. Atherosclerosis begins with dysfunctional endothelial cells and the retention of LDL-C in the subendothelial space. The retained LDL-C are modified (e.g. by oxidation) and promotes activation of endothelial cells to express monocyte interaction/adhesion molecules leading to attachment and transmigration of monocytes into the intimal space. The monocytes differentiate into macrophages, which express receptors that mediate the uptake of modified LDL and converting it into foam cells. Subsequently, smooth muscle cells migrate into the subendothelial space and form a fibrous cap which surrounds the atherosclerotic plaque. Reproduced from Heinecke, 2006.

Inflammation plays a fundamental role in all phases of atherosclerosis, as evidenced by the continuous presence and accumulation of monocyte-derived macrophages and T lymphocytes in fatty streaks and advanced atherosclerotic lesions (Ablij & Meinders, 2002). Both activated complement and CRP are present in atherosclerotic lesions (Ablij & Meinders, 2002). CRP may be more than just an inflammatory marker of increased cardiovascular risk, several *in vitro* and *in vivo* studies claim that CRP has an active role in

the endothelial dysfunction, atherosclerotic plaque formation, progression and rupture that result in atherothrombotic events. CRP actively participates in atheroma formation by binding modified LDL of the type that accumulates in atherosclerotic plaques (ox-LDL and partly degraded LDL) and facilitates their uptake by macrophages by opsonisation thus enhancing foam cell formation, a typical feature of atherosclerotic plaques (Osman *et al.*, 2006). Aggregated and/or ligand complexed CRP activates complement, thus promoting inflammation in the plaques; therefore, CRP can be proinflammatory. CRP and activated complement are found in all acute myocardial infarction lesions, and they can significantly increase the infarct size in experimental models (Pepys & Hirschfield, 2003). Furthermore, CRP up-regulates the expression of endothelial adhesion molecules (ICAM-1, VCAM-1, and E-selectin) through binding of FcγRII and activation of the NF-κB pathway (Pasceri *et al.*, 2000, Verma *et al.*, 2003). CRP also stimulates MCP-1 expression from endothelial cells. MCP-1 and other chemoattractant factors cause monocytes that have adhered to endothelial cells to migrate across the endothelial cells (Pasceri *et al.*, 2001). Elevated serum CRP levels are also associated with reduced expression and activity of endothelial nitric oxide synthase (eNOS), thus reducing nitric oxide production in endothelial cells. Nitric oxide plays crucial roles in many biological processes including regulation of vascular tone, inhibition of smooth muscle cells proliferation, leukocyte adhesion, and platelet aggregation. Inhibition of nitric oxide synthesis by CRP, therefore, leads to vasoconstriction and promotes endothelial cell dysfunction and apoptosis (Osman *et al.*, 2006, Salazar *et al.*, 2014). CRP can also promote the synthesis of matrix metalloproteinases (MMPs), proteolytic enzymes responsible for remodelling the extracellular matrix, which is implicated in plaque rupture and thrombosis. Elevated MMP levels may contribute to plaque instability, which may explain the heightened risk of coronary events in subjects

with elevated serum levels of CRP (Cimmino *et al.*, 2013, Montero *et al.*, 2006). CRP has been also linked to thrombosis by stimulating the release of tissue factor (a potent stimulus for the extrinsic pathway of coagulation) from peripheral blood monocytes, endothelial cells and smooth muscle cells (Cermak *et al.*, 1993, Cirillo *et al.*, 2005). In addition, CRP increases the expression and activity of plasminogen activator inhibitor-1 (PAI-1), a marker of atherothrombosis and impaired fibrinolysis in human aortic endothelial cells (Devaraj *et al.*, 2003). These postulated mechanisms support the rule of CRP in clinical atherothrombosis in addition to its role in atherosclerosis progression. CRP has been shown to induce the production of endothelin-1 (a potent endogenous vasoconstrictor) and IL-6 from endothelial cells. IL-6, in turn, stimulates CRP production by the liver (Verma, Li *et al.*, 2002). CRP inhibits bone marrow-derived endothelial progenitor cell survival and differentiation, these cells play a role in promoting angiogenesis in ischaemic tissue, therefore, their inhibition results in impaired maintenance of vascular integrity (Verma *et al.*, 2004). CRP also assumed to induce overexpression of angiotensin type 1 receptor in vascular smooth muscle cells *in vitro* and *in vivo* and stimulates vascular smooth muscle migration, proliferation, neointimal formation, and reactive oxygen species production (Verma *et al.*, 2004).

#### 1.1.13.3 Doubts about the *in vitro* experiment

There are some doubts about the validity of the *in vitro* investigation of CRP effects on endothelial cells. Commercial preparations of recombinant human CRP produced in *E. coli* may contain contaminating bacterial endotoxins such as lipopolysaccharide (LPS), as well as sodium azide as a bacteriostatic preservative. LPS is known to activate endothelial cells, while azide is generally cytotoxic; the presence of these substances may interfere with *in vitro* investigation of CRP effects on endothelial cells. For this reason, all the published work

that has been done using commercially-sourced CRP of incompletely defined provenance and purity were criticised since results may be attributed to LPS or azide rather than CRP itself (Pepys & Hirschfield, 2003, Hirschfield & Pepys, 2003). Taylor and associates examined the literature up to 2005, and found that only two studies detailed the use of azide-free CRP was used. One of these two studies demonstrated a modulatory effect of CRP on monocyte adhesion to human endothelial cells (Taylor *et al.*, 2005). The LPS contamination aspect has been better investigated, with at least four studies employing a Detoxi-Gel™ column for endotoxin removal, with SDS-polyacrylamide gel electrophoresis and/or Limulus assay for endotoxin level estimation. The results of these studies support the role of CRP in atherogenesis (Pasceri *et al.*, 2001, Venugopal *et al.*, 2002, Venugopal *et al.*, 2003, Verma, Wang *et al.*, 2002). The remainder of the previous studies were however potentially affected either by LPS or azide contamination. Many researchers tried to prove the results of previous studies, Dasu and co-workers used azide-free CRP preparations purified by Detoxi-Gel™ column in Toll-like receptor 4 (TLR4) knockdown endothelial cells to abrogate the effect of LPS as TLR4 is the receptor for sensing LPS and consequent signalling. CRP incubation induced secretion of pro-inflammatory cytokines (IL-8, IL-6, IL1- $\beta$ ) and PAI-1 from knockdown endothelial cells and inhibited eNOS activity (Dasu *et al.*, 2007). Other studies using heat-inactivated CRP (which has no effect on endotoxin activity), abolishes the pro-atherogenic effects of CRP (Devaraj *et al.*, 2009, Singh *et al.*, 2005). Bisioendial *et al.* have stated that the trace amount of endotoxin present in recombinant human CRP following additional purification steps is insufficient to contribute to the pro-atherogenic effects (Bisioendial *et al.*, 2005). These studies may confirm that the biological effects of CRP on endothelial cells are independent of endotoxin, a common contaminant in commercial CRP preparations.



#### 1.1.14 CRP and metabolic syndrome

Metabolic syndrome (Mets) can be defined as the presence of three or more of the following criteria: Central obesity, Dyslipidemia (elevated triglycerides and/or reduced levels of HDL cholesterol), High blood pressure, and High fasting glucose (Grundy *et al.*, 2004). Mets associated increased risk of mortality as a result of the systemic inflammatory state, increased risk of CVD or type two diabetes mellitus that found in Mets patients. Therefore, measurement of inflammatory markers such as CRP in patients with Mets can help in the prediction of the hazardous outcome of this syndrome (Huffman *et al.*, 2009, Vidyasagar *et al.*, 2013).

Increased serum CRP levels (more than 3 mg/L) have been found in patients with Met, and a positive correlation has been found between serum CRP level and each of the component of Mets (Huffman *et al.*, 2009, Vidyasagar *et al.*, 2013). It is suggested that the elevated CRP concentrations seen in patients with severe Mets may reflect IL-6 production by adipocytes (Fröhlich *et al.*, 2000). Large scale, population-based studies have concluded that CRP concentration >3 mg/L in patients with Mets confers greater risk for developing cardiovascular disease (Freeman *et al.*, 2002, Ridker *et al.*, 2008). This led to the recommendation of considering CRP as a clinical criterion for metabolic syndrome definition because it adds prognostic information with regard to the risk of CVD, and improves the early detection of risk for future diabetes and CVD in these individuals (Devaraj *et al.*, 2009, Ridker *et al.*, 2004).

### 1.1.15 CRP and diabetes mellitus

Numerous studies have hypothesised that chronic low-grade inflammation plays a role in the pathogenesis of both type 1 and type 2 diabetes, and a moderately increased levels of CRP were found in both types of diabetes (Schalkwijk *et al.*, 1999, Kilpatrick *et al.*, 2000, Pickup & Crook, 1998, Duncan *et al.*, 2003). The increased levels of CRP in diabetes are not caused by the development or presence of diabetic complications but might precede these (Schalkwijk *et al.*, 1999, Chase *et al.*, 2004). High CRP levels were found to be an independent predictor of risk for the development of type 2 diabetes mellitus even after adjustment for coronary risk factors, BMI, fasting insulin levels, smoking and other factors (Freeman *et al.*, 2002, Pradhan *et al.*, 2001). Prolonged exposure to hyperglycaemia is known to stimulate the release of the pro-inflammatory cytokines from adipocytes, monocytes and other cells, which can increase the production of CRP in the liver (Ingle & Patel, 2011, Goh & Cooper, 2008). Moreover, diabetic individuals are at high risk of developing other inflammatory conditions including atherosclerosis and heart disease, which associated with high levels of CRP (Messier *et al.*, 2004, Lee *et al.*, 2015). CRP may also give information about glycaemic control; CRP found to be higher in type 2 diabetic patients with high haemoglobin A1c (HbA1c) as compared to those with recommended HbA1c level (Capuzzi & Freeman, 2007, King *et al.*, 2003).

## 1.2 Protein glycation

### 1.2.1 Introduction

Protein glycation is a process that involves non-enzymatic attachment of a reducing sugar or sugar derivative to a reactive protein residue (Rabbani & Thornalley, 2012). The reducing sugars are sugars with a free aldehyde or ketone groups, this includes the common monosaccharides such as glucose, galactose, glyceraldehyde, fructose, and ribose. In the human body, glucose, the most important source of fuel, is the most abundant monosaccharide making it the main glycating sugar present in the body (Aragno & Mastrocola, 2017). Glucose exists almost entirely in the pyranose (ring-closed) form, with less than 0.4% in the furanose ring form, and less than 0.02% in the open-chain aldehyde form (Bubb, 2003). Only the acyclic form of glucose can react with the reactive protein residues during glycation reaction, this is because the open-chain form possesses the reactive carbonyl group (Aragno & Mastrocola, 2017). Other sugars and sugar derivatives such as galactose, fructose, ribose, sialic acid, mannose, glucose 6-phosphate, and glyceraldehydes are more reactive than glucose and have been used *in vitro* glycation experiments, sometimes to hasten the process of glycation (Arasteh *et al.*, 2014). The main targets of glycation in the proteins are the  $\epsilon$ -amino groups of lysine residue side chains and N-terminal amino acid residues (Ansari & Ali, 2011, Rabbani & Thornalley, 2012). Glycation of arginine residues is also seen as a major feature of protein glycation in physiological systems, and there is also evidence for cysteine glycation (Rabbani & Thornalley, 2012, Münch *et al.*, 1999). Protein glycation reactions are relatively non-specific, not mediated by enzymes and usually resulting in protein damage and reducing functionality, it differs from the enzymatic glycosylation, an essential post-translational modification required for the molecule to function (Spiro, 2002). Glycation was first identified by Louis Camille

Maillard in 1912; therefore, glycation reactions is also referred to as 'Maillard reactions' (Maillard, 1912). Protein glycation has been studied most extensively in foods; the formation of glycation reaction products is responsible for the brown colour formation and loss of protein quality in cooked and stored foods (Ulrich & Cerami, 2001). The realisation of the importance of the glycation in the human body began after the studies on haemoglobin A 1c (HbA1c), a naturally occurring minor fraction of human haemoglobin that is elevated in diabetics (Koenig *et al.*, 1976). HbA1c is glycated haemoglobin in which the glucose molecule is condensed on N-terminal of valine residue of the beta chain of haemoglobin (Welsh *et al.*, 2016). The measurement of the elevation of HbA1c in diabetics allowed the assessment of the glycaemic control integrated over several weeks (Koenig *et al.*, 1976). In the human body, protein glycation can take place under normal physiological conditions, but the reaction is accelerated in hyperglycemia as seen in diabetes especially in an uncontrolled state (Arasteh *et al.*, 2014). In addition to the haemoglobin, *in vivo* glycation has been described in other proteins such as; serum albumin, ferritin, apolipoprotein, collagen, insulin, histone, IgM, IgG, and the lens protein crystallin (Arasteh *et al.*, 2014, Ansari & Ali, 2011).

### 1.2.2 The chemistry of protein glycation

The process of protein glycation can be divided into three distinguishable phases: the initial stage, the intermediate stage and the late stage (Wang & Ho, 2012).

During the initial stages, the carbonyl group of the reducing carbohydrates reacts with the nitrogen atom of the  $\epsilon$ -amino group on lysine or arginine or the N-terminal amino group by the nucleophilic attack (Wang & Ho, 2012). The product of this reaction is a reversible glycosylamine which dehydrates to form unstable Schiff base (aldimine) as shown in figure

1-8, the formation of the Schiff base is fast but highly reversible (Rabbani & Thornalley, 2012). The Schiff base then undergoes a rearrangement to form a more stable ketoamine named Amadori product (such as fructosamine if the reaction is between glucose and lysine residues) as shown in figure 1-8 (Rabbani & Thornalley, 2012). Formation of Amadori product from the Schiff base (the forward reaction) is much faster than the reversed reaction so that the Amadori adducts tend to accumulation on the protein (Rabbani & Thornalley, 2012). This process (Amadori rearrangement) is thought to be facilitated by localised acid-base catalysis if there is a basic residue in the proximity from the target amino residue on which the Schiff base has formed (Ulrich & Cerami, 2001).

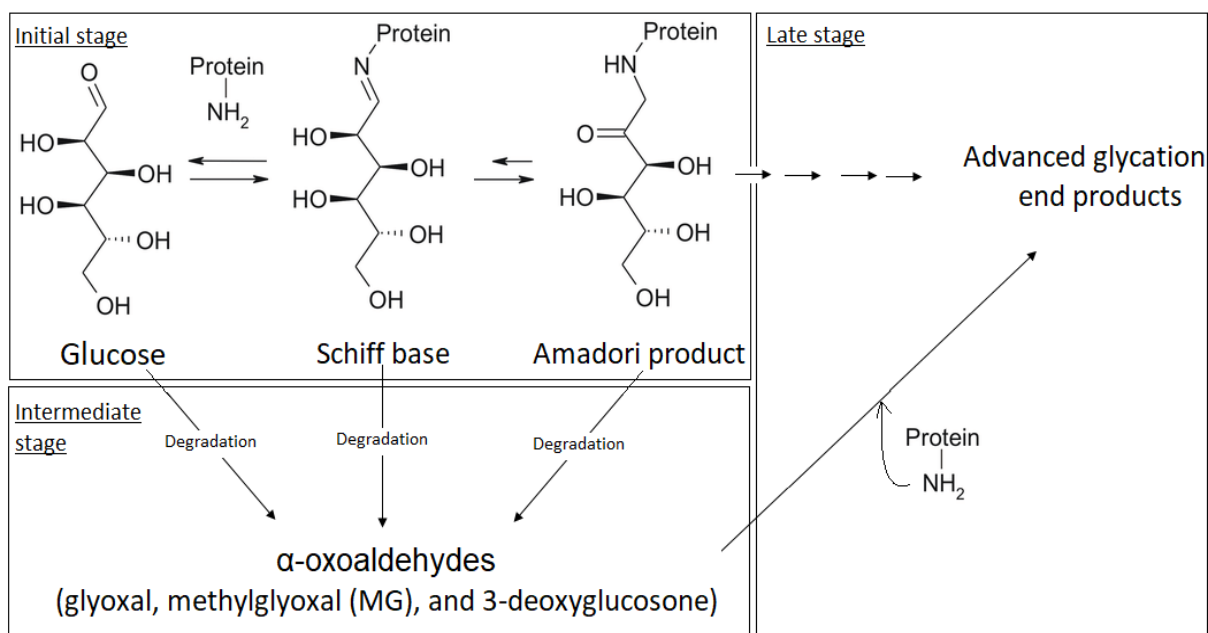


Figure 1-8 Protein glycation pathway. Protein glycation reactions can be divided into three stages: the initial stage involves condensation of glucose with primary amines to give a Schiff base, which rearranges to an Amadori product, the intermediate stage during which the α-oxoaldehydes are formed from the degradation of glucose, Schiff base or Amadori product, and the late stage that result in the formation of the advanced glycation end products via several pathways, including direct degradation of Amadori products or protein modification by dicarbonyl compounds.

During the Intermediate stage, the Amadori products or the Schiff base can degrade via non-Amadori rearrangement reactions to  $\alpha$ -oxoaldehydes (dicarbonyl compounds) as shown in figure 1-8, these compounds more potent glycating agents than glucose (Rabbani & Thornalley, 2012). The most important  $\alpha$ -oxoaldehydes are glyoxal (GO), methylglyoxal (MG), and 3-deoxyglucosone (3-DG), see figure 1-9; (Rabbani & Thornalley, 2012). GO is formed from several reactions including the oxidative fragmentation of Schiff base via the Namiki pathway (Wang & Ho, 2012), the 3-DG can be produced from non-oxidative fragmentation and hydrolysis of Amadori product (Wang & Ho, 2012). GO and 3-DG can be also produced as a result of metabolic degradation of D-glucose; the auto-oxidation of glucose in the presence of metal can produce GO (Wolff pathway); (Jakuš & Rietbrock, 2004), while 3-DG can also be produced from fructose-3-phosphate, an intermediate of the Polyol pathway in which glucose is reduced to sorbitol by aldose reductase (Lal *et al.*, 1995). Methyl glyoxal can be formed from further degradation of 3-DG generates by retro-aldol condensation (Wang & Ho, 2012).  $\alpha$ -oxoaldehydes are more reactive glycating agents than glucose due to their lack of cyclic structure and comparatively smaller size (Rabbani & Thornalley, 2012).

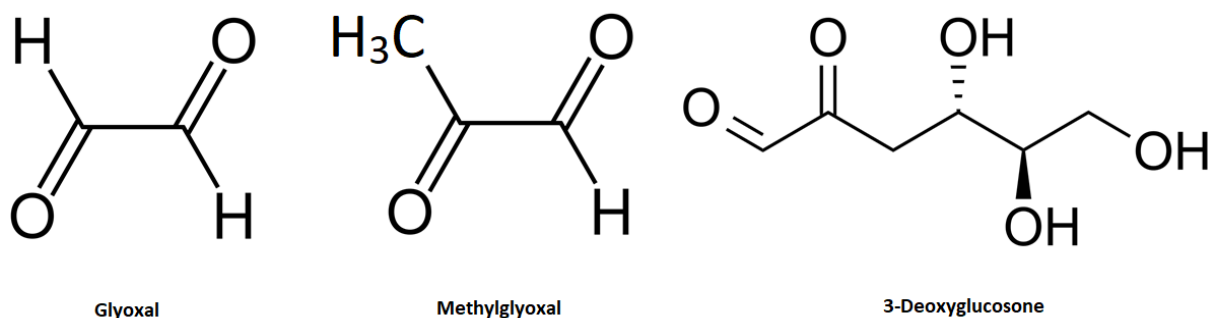


Figure 1-9 Reactive dicarbonyl species. Glyoxal, methylglyoxal and 3-deoxyglucosone are highly reactive  $\alpha$ -oxoaldehydes and precursors for the formation of advanced glycation end products.

During the late stage, heterogeneous group of compounds, collectively termed the advanced glycation end products (AGEs), are formed (Rabbani & Thornalley, 2012). The Amadori product or Schiff base can slowly undergo further dehydration, cyclisation, oxidation and rearrangement to form AGEs (Rabbani & Thornalley, 2012). Dicarbonyl compounds, formed from the intermediate stage of glycation, potent precursors for AGEs, can interact with proteins or other amino-containing molecules to produce AGEs at a much higher reaction rate than does the reducing sugar itself (Rabbani & Thornalley, 2012, Thorpe & Baynes, 2003). Examples of these AGEs are N $\epsilon$ -carboxymethyl lysine (CML), carboxyethyl lysine (CEL), pentosidine, glucosepane, and others such as methylglyoxal lysine dimer (MOLD), glyoxal-lysine dimer (GLOD), 3-deoxyglucosone-derived imidazolium cross-link (DOGDIC), methylglyoxal-derived imidazolium cross-link (MODIC) and glyoxal-derived imidazolium cross-link (GODIC), see figure 1-10 (Cho *et al.*, 2007, Sell *et al.*, 2005, Ulrich & Cerami, 2001).

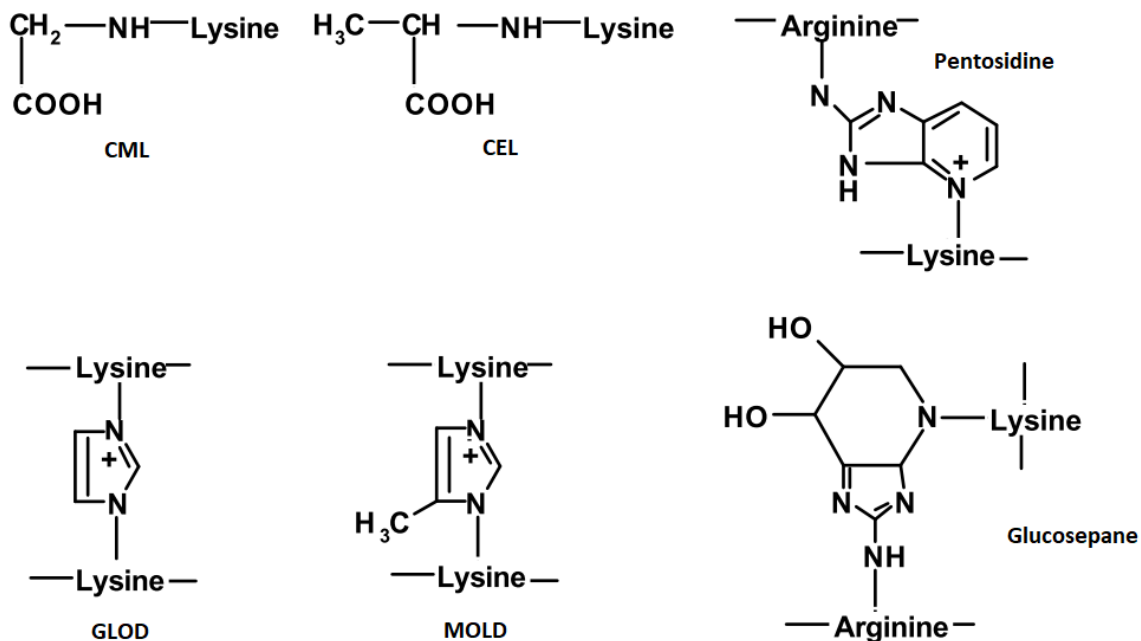


Figure 1-10 AGEs Structures. The structures of some of the known AGEs

Except for CML these are all cross-links between two amino acid residue side chains within proteins. GOLD and MOLD are lysine-lysine cross-links, whereas pentosidine, glucosepane, DOGDIC, MODIC and GODIC are lysine-arginine cross-links (Cho *et al.*, 2007, Sell *et al.*, 2005).

### 1.2.3 Biological effects of protein glycation

Protein glycation and the resulting AGEs have varied deleterious effects on the proteins and biological systems. Modification of the side chains of lysine and arginine residues by glycation is likely to affect the protein interaction of with other molecules such as ligands, substrate or the cell receptors (Avery & Bailey, 2005). Moreover, the cross-links between amino acid residues within the protein as a result of AGEs formation could alter the shape, restrict the movement, and affect the mechanical properties of the protein (Avery & Bailey, 2005). Furthermore, the interaction between AGEs and the receptor for AGEs (RAGE, non-specific cellular receptors belong to the immunoglobulin super-family of cell surface molecules) can increase oxidant stress and induce a cascade of inflammatory reactions, these interactions (AGEs-RAGE) plays a significant role in diabetic complications (Sparvero *et al.*, 2009, Stern *et al.*, 2002).

The rate of glycation of any protein is dependent on the glucose concentration and the duration of exposure (the biological half-life) of the protein to the glycating agent (Olufemi *et al.*, 1987). Therefore, the degree of glycation of proteins in diabetic patients is significantly greater than in non-diabetic patients because of the persistent hyperglycaemia in those patients (Negre-Salvayre *et al.*, 2009). Many studies indicate that the increased rate of protein glycation in diabetic patients and the accumulation of the AGEs play an important role in diabetic complications such as diabetic nephropathy, retinopathy, neuropathy and cardiovascular diseases (Negre-Salvayre *et al.*, 2009, Thornalley *et al.*, 1999). The proteins with a longer half-life, such as collagen and lens crystalline, are more implicated with glycation mediated effects than the short-lived proteins because of the longer exposure time (Thornalley *et al.*, 1999). The accumulation of the AGEs on the long-lived proteins are implicated in several chronic complications, for example cataract formation as a result of accumulation of AGEs on lens crystalline (Ulrich & Cerami, 2001),



while glycation of collagen may contribute to atherosclerosis and coronary disease, poor peripheral circulation, and other lesions (Ulrich & Cerami, 2001). AGEs are involved not only in diabetic complications but also in pathophysiology of other conditions such as neurodegenerative diseases (Alzheimer's and Parkinson disease); (Grillo & Colombatto, 2008), renal failure (Thornalley, 2006), some types of liver disease (Hyogo & Yamagishi, 2008), and the normal ageing (Grillo & Colombatto, 2008). The accumulation of AGEs in tissues and their interaction with RAGE are the main mechanisms that explain glycation-related pathophysiology (Fournet *et al.*, 2018, Nawale *et al.*, 2006).

### 1.3 Research aims and objectives

The high levels of circulating blood sugar found in diabetic patients can cause protein glycation, however, the effect of the glycation reaction on CRP remaining unexplored. In addition, recent studies have shown that a monomeric isoform of CRP, formed as a result of a dissociation process of CRP, plays a pro-inflammatory harmful role in the progression of diseases.

The overall aims of this research were first to provide insight, via the production and characterisation of *in vitro* produced glycated CRP, into the susceptibility of CRP to the glycation reaction. Secondly, this research aims to clarify whether glycated and monomeric modified forms of CRP are present, detectable and quantifiable within the serum of diabetes patients (type 1 and 2) with elevated levels of CRP, with a view to assessing the relationship between pentameric, monomeric and glycated CRP in diabetes mellitus.

The specific objectives required to fulfil these aims are as follows:

1. Preparation of *in vitro* glycated samples of CRP conjugated with glucose through the Maillard reaction, and study the effects of the concentration and exposure time to glucose on the extent of glycation of the target protein using electrospray mass spectrometry and other biochemical methods.
2. Identification of the potential glycation sites of CRP treated with various concentrations of glucose for a varying time interval, and determination of the favourable glycation site using the tandem mass spectrometry, this will also involve optimisation of the enzymatic digestion method and glycated CRP samples preparation for tandem mass analysis.

3. Investigation of the microenvironment of the potential glycation sites including; the distance to the nearby acidic and basic residues in the 3D structure of the protein, the exposed surface area of the potential glycation sites and their pKa values.
4. Identification, then quantification of the modified (monomeric and glycated) isoforms of CRP in the serum of diabetes patients (type 1 and 2) with elevated CRP levels and assessing the relationship between pCRP, mCRP and the glycated isoform. This will involve separation of CRP isoforms using affinity chromatography, identification of the glycated form and the site(s) of glycation using the mass spectrometry, separation of the monomeric form of the protein using size exclusion chromatography followed by quantification by ELISA.

## Chapter 2: Production & characterisation of glycated C-reactive protein

### 2.1 Introduction

#### 2.1.1 Overview

Glycation involves the non-enzymatic addition of reducing sugars and/or reactive sugar derivatives to the protein reactive residues. This modification alters the three-dimensional structure of proteins and consequently their physical and functional properties. Glycation affects a wide variety of proteins in human serum at a rate greatly accelerated during hyperglycaemia and derived disorders. C-reactive protein (CRP) is a classic acute phase protein which plays a vital role in infection, inflammation, and tissue damage. CRP is a clinically important molecule. However, the effect of glycation of this protein remains unexplored. We aim to shed light on this subject, and this chapter discusses modifying CRP with various concentrations of glucose for different intervals, and identification of the post-translational modifications of CRP resulting from non-enzymatic glycation, determined via electrospray mass spectrometry.

#### 2.1.2 Analytical techniques used for detection of glycated proteins

Glycation reactions are a complicated group of processes, with primary and secondary reactions potentially resulting in a heterogeneous group of compounds. The end products depend on the glycating agent involved, its concentration, the exposure time and reaction conditions (Eble *et al.*, 1983, McPherson *et al.*, 1988, Austin *et al.*, 1987). The complex nature of the reaction and the diversity in the reaction endpoint makes glycation reaction monitoring a challenging task (Priego Capote & Sanchez, 2009). Multiple analytical methods have been utilised for the detection and characterisation of glycation products,

including colourimetric methods (such as the thiobarbituric acid, nitroblue tetrazolium and hydrazine-based), immunological assays (enzyme-linked immunosorbent assays, Western blotting or radioimmunoassays), electrophoretic methods (single and two dimensional gel electrophoreses), chromatographic (boronate affinity chromatography), and mass spectrometric-based methods (Lacinová *et al.*, 2010, Priego Capote & Sanchez, 2009). The choice of method depends on the nature of information required; however, a combination of more than one method may be needed to derive the required information. The focus will be limited to the analytical methods utilised in this chapter.

### 2.1.3 Polyacrylamide Gel Electrophoresis

Polyacrylamide gel electrophoresis (PAGE) is an analytical method used to separate proteins in a sample based on their mobility through a sieving gel matrix under the influence of an applied electrical field (Hames, 1998). PAGE is relatively simple, fast and highly reproducible. PAGE can provide information about the molecular weight, subunit structures, electrical charges, amount and purity of polypeptides or proteins in the sample depending on the particular technique used (Hames, 1998).

The critical element in the PAGE system is the gel itself. Polyacrylamide gels are polymerised from monomeric acrylamide and *N,N'*-methylene-bisacrylamide (Hames, 1998). In the presence of suitable initiator compounds (ammonium persulfate, tetramethylethylenediamine), the acrylamide polymerises to form long polyacrylamide chains, cross-linked by *N,N'*-methylene bisacrylamide to form a three-dimensional network with uniform pores (Hames, 1998). The size of the pores within the gel is governed by the total amount of acrylamide used per unit volume and the cross-linker ratio (Bass *et al.*, 2017). Gels with higher acrylamide concentration (e.g., 20%) have smaller pores, meaning

the gel impedes the movement of larger proteins to a greater degree than those of smaller molecular weight but better resolves those of lower molecular weights (Bass *et al.*, 2017).

To obtain an optimal resolution of a sample component a “stacking” gel is cast over the top of the “resolving” gel (Bass *et al.*, 2017). The stacking gel (where the samples are loaded) has a lower pH and reduced acrylamide content, hence larger pores (Bass *et al.*, 2017). This allows the sample component to be concentrated into a tight ‘band’ during the first few minutes of electrophoresis prior to entering the resolving portion of a gel, where proteins are separated according to their molecular weight (Bass *et al.*, 2017).

PAGE comprises a number of different techniques each with its uses, advantages and limitations. Native-polyacrylamide gel electrophoresis (native-PAGE) separates proteins, in their native form, based on their charge, shape and size (Hames, 1998). In native-PAGE the protein is loaded in non-reducing and non-denaturing conditions to preserve the secondary and tertiary structure of the protein in the sample; thus native-PAGE preserves the proteins’ function, structure and activity and can be used, for example, alongside detection of the proteins biological activity (Hames, 1998).

In contrast to native-PAGE, where three independent molecular properties influence the electrophoretic mobility of the proteins, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins almost exclusively based on their size (Hames, 1998). In this technique, sodium dodecyl sulfate (SDS), an anionic detergent, is incorporated in sufficient excess to denature the protein and produce a linear polypeptide chain covered with a uniform negative charge of SDS molecules (Hames, 1998). Samples are heated to a high temperature (typically 95 °C for 5 minutes) in the presence of denaturants such as beta-mercaptoethanol to further denature the protein, removing the

three-dimensional structure. The denatured proteins migrate through SDS gel based primarily on their size, and hence SDS-PAGE is used to estimate the molecular weights of proteins (Hames, 1998).

Isoelectric focusing (IEF) is an electrophoretic technique used to separate proteins across a pH gradient based on their isoelectric point (PI); (Friedman *et al.*, 2009). During IEF, proteins migrate under the influence of an electric field through an immobilised pH gradient until a given protein reaches a pH within the gel at which their net charge is zero; this is the PI of this protein. Determination of PI by IEF can separate minor charge variants resulting from protein post-translational modifications (Friedman *et al.*, 2009).

Protein glycation increases the mass of the protein based on the net number of sugar molecules incorporated to a given protein subunit. Protein glycation could result in the formation of further cross-linked structures in the later reaction stages, with significantly higher mass than their non-cross-linked counterparts. SDS-PAGE has the potential to reveal changes in the molecular weight of glycated and cross-linked species due to the net increase in mass (Ghosh *et al.*, 2013, Galván-Moroyoqui *et al.*, 2016, Bai *et al.*, 2012, Ledesma-Osuna *et al.*, 2008, Chen *et al.*, 2012). Native-PAGE can reveal the changes in the charge properties of glycated proteins, where glycated proteins will migrate more quickly toward the positive pole of the gel under native conditions (Wijetunge & Perera, 2014, Btissam *et al.*, 2017, Handa & Kuroda, 1999). The increased electrophoretic mobility of glycated proteins has been attributed to the increase in the net negative charge of glycated proteins due to the attachment of sugar residues to the amino groups of the protein (Wijetunge & Perera, 2014, Handa & Kuroda, 1999). This change in the net charge is more significant when proteins are glycated with phosphorylated sugars such as glucose-6-

phosphate and fructose bisphosphate (Yaylayan *et al.*, 1994), or with carboxyl-containing sugars such as glucuronic acid (Aoki *et al.*, 1999). Incorporation of negatively charged sugars derivatives adds extra negative charges and increase the net negative charge of the protein; this will increase their electrophoretic mobility through Native-PAGE. IEF also has been used to separate glycosylated proteins. The increase in the net negative charge of glycosylated proteins results in a decrease in the pI, which supports the occurrence of glycosylation (Luthra & Balasubramanian, 1993).

#### 2.1.4 Western Blotting

Western blotting is an analytical technique to identify a particular antigen in the midst of a complex mixture of proteins, based on the interaction between an antibody and its specific antigen. Western blotting has diverse applications including; detection and quantification of specific proteins, estimation of protein molecular weight, and in some cases the determination of post-translational modifications (Blancher & Jones, 2001, Bass *et al.*, 2017).

The first step in western blotting is to separate the proteins in the sample by gel electrophoresis, commonly SDS-PAGE. Separated proteins are transferred to a membrane by electroblotting; this immobilises the separated proteins and provides a durable surface for probing with antibodies (Bass *et al.*, 2017). Electroblotting is typically achieved by placing the membrane in direct contact with the gel and immersing this sandwich in a buffer solution (wet transfer). After applying a lateral electrical current, the negatively charged protein will transfer from the gel to produce a mirror image on the membrane (Bass *et al.*, 2017, Mahmood & Yang, 2012).



The two most commonly used membranes in western blotting are nitrocellulose and polyvinylidene difluoride (PVDF). The nitrocellulose membranes have the advantage of the low background regardless of the detection method, lower cost, and ease of use as it does not require an activation step. However, it is fragile when dried and not suitable for stripping and re-probing. On the other hand, PVDF membranes are more durable, allowing stripping and re-probing the blots and more convenient to store. Their hydrophobic nature results in high protein binding capacity, however, as a consequence the background noise is also higher. PVDF require pre-soaking with methanol to hydrating the membrane and improve transfer and binding (Bass *et al.*, 2017, Mahmood & Yang, 2012).

After transferring the proteins to the membrane, the unoccupied surface of the membrane is 'blocked' to prevent nonspecific interaction of antibodies with the membrane. Non-fat dry milk is commonly used as it is cheap and widely available. Blocking reduces background noise and improves signal-to-noise and hence the sensitivity of the assay (Mahmood & Yang, 2012, Bass *et al.*, 2017). Blocked membranes are probed using a primary antibody that specifically recognises the target protein. This primary antibody is produced by immunising a host with purified target protein (or fragments thereof), and it should be thoroughly assessed for specificity and sensitivity for the target protein (Bass *et al.*, 2017). Both polyclonal and monoclonal primary antibodies can be used for target protein detection. Polyclonal antibodies are more cost-effective, and they provide a greater level of sensitivity as they bind to multiple epitopes in the target protein; however, a polyclonal antibody can give higher background and have a greater chance of non-specific binding (Bass *et al.*, 2017). In contrast, the monoclonal antibody is raised against single specific epitope; therefore, they can provide higher specificity for the target antigen with less background (Bass *et al.*, 2017). The binding affinity of the monoclonal antibody can be

influenced if the target epitope is affected in any way through denaturing or modification (Bass *et al.*, 2017).

The primary antibody-target protein complex is then detected using a secondary antibody, typically conjugated to a reporter, usually an enzyme such as horseradish peroxidase (HRP). The importance of the secondary antibody is to amplify the detectable signal as several molecules of secondary antibodies can bind to the primary antibody (Bass *et al.*, 2017). The secondary antibody must be species-specific and isotype-specific (Bass *et al.*, 2017). Finally, a suitable substrate is added to react with the secondary antibody's reporter to produce a signal. This signal locates the target protein on the membrane with an intensity proportional to the amount of the target protein (Mahmood & Yang, 2012).

Western blot can be used for the detection of the glycated form of proteins. Monoclonal and polyclonal antibodies raised against early glycation adduct (1-deoxyfructosyl lysine) (Wu & Steward, 1991), and advanced glycation end products (N<sup>ε</sup>-carboxymethyl-lysine and N<sup>ε</sup>-carboxyethyl-lysine) have been produced and found to be specific for the corresponding glycation products (Choi & Lim, 2009). On the other hand, protein glycation may result in alterations in protein structure which in turn could result in steric hindrance or epitope modification and thereby affecting antigen-antibody recognition (Seo *et al.*, 2013).

#### 2.1.5 Mass spectrometry

Mass spectrometry is a powerful analytical technique in the field of proteomics and related studies (Aebbersold & Mann, 2003). It relies on the conversion of analytes into gaseous ions, separating them according to their mass to charge ratio ( $m/z$ ) and recording the abundance of each separated ion. Increased information yield is accessed via fragmentation of selected precursor ions (De Hoffmann, 2000). Mass spectrometry has a myriad of applications

including protein identification and quantification, detection of post-translational modifications, determination of protein structure, function, folding and interactions (Aebersold & Mann, 2003).

Different kinds of mass spectrometry instruments are available; with the newest mass spectrometers aimed to suit multiple experimental goals. Generally, a mass spectrometer consists of an ion source to produce gas phase ions from the sample under investigation, one or several mass analysers to separate the generated ions according to their  $m/z$  ratio, and a detector to count the ions at each  $m/z$  value, see Figure 2-1 (De Hoffmann, 2000).

The most common ionisation methods employed in protein mass spectrometry are Matrix-assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI; Aebersold & Mann, 2003). Both are capable of “soft” ionisation and volatilisation of biomolecules into gas phase ions without excessive fragmentation (De Hoffmann, 2000)

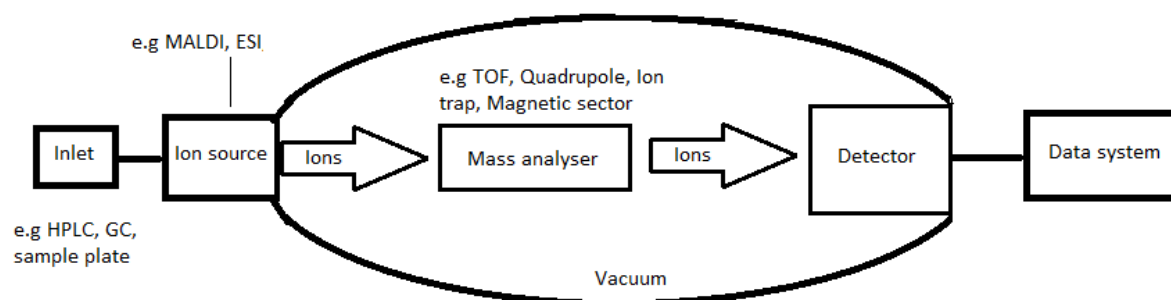


Figure 2-1 Schematic representation of the components of a mass spectrometer. Mass spectrometers can be divided into three fundamental parts; the ion source, the analyser, and the detector. The sample can be inserted directly into the ion source, or can undergo some type of chromatography en route to the ion source e.g. a high-pressure liquid chromatography (HPLC) or gas chromatography (GC); (De Hoffmann, 2000).

MALDI ionises analytes from the solid phase, where the analyte is typically mixed with a solvent containing organic matrix molecules then dried to form a thin layer in which the analyte is embedded within the matrix crystals. Within the source, the analyte is irradiated with a pulsed laser beam causing sublimation of the analyte as singly or doubly charged ions (De Hoffmann, 2000). The principally singly charged nature of MALDI-generated precursor ions results in easy to interpret spectra with limited need for post-analysis processing.

ESI ionises solutes; sample solution is pumped through a capillary with a high voltage applied to create a spray of charged droplets; these droplets evaporate, usually by means of heat and drying gas, eventually releasing the ions into the gas phase (De Hoffmann, 2000). ESI typically generates multiply charged ions and is compatible with solution-based separation techniques such as liquid chromatography (LC), hence allowing online separation of complex samples. A minimised-flow version of ESI, named nano-spray (nESI), may be utilised with a low flow rate LC. nESI produces droplets with a smaller diameter compared to ESI. This enhances droplet desolvation and thus increases the ionisation efficiency of analyte molecules beside minimising the volume of sample needed (Juraschek *et al.*, 1999).

After the ions are produced and transferred into a gas phase, they have to be separated in a mass analyser according to their  $m/z$ . There are many types of mass analysers each has its own characteristics; however, in the proteomic field time of flight (TOF), quadrupoles, ion traps and the orbitraps are the commonly used analysers (Aebbersold & Mann, 2003). More than one analyser type may be joined together in tandem in a single instrument platform. This offers the advantage of mass filtering and subsequent ions activation for

tandem mass analysis through which additional information about specific ion species can be obtained. In tandem mass spectrometry analysis (also called MS/MS) the ionised sample is separated in the first mass analyser then ions with a particular  $m/z$  value are selected. The selected ions are subjected to collisions, commonly with inert gas, the resulting fragment ions are separated and their  $m/z$  ratios recorded (De Hoffmann, 2000).

TOF analysers are conceptually the simplest mass analyser; this can be considered as a velocity metre. TOF measures the time that ions need to pass through a tube of a fixed length under a vacuum (De Hoffmann, 2000). To improve mass resolution and obtain the best performance from TOF analyser; an electrostatic mirror (called a reflectron) is used to turn the ions around and sends them back a second flight distance to the detector as shown in figure 2-2 (Glish & Vachet, 2003). The function of the reflectron is to compensate for small differences in the velocities of ions with the same  $m/z$ . Therefore, the reflectron increases the resolution of TOF spectrometry (Glish & Vachet, 2003). TOF analysers are typically fast, sensitive, and theoretically have no upper mass limit, furthermore, increasing the path length of a reflectron increase available resolving power (De Hoffmann, 2000).

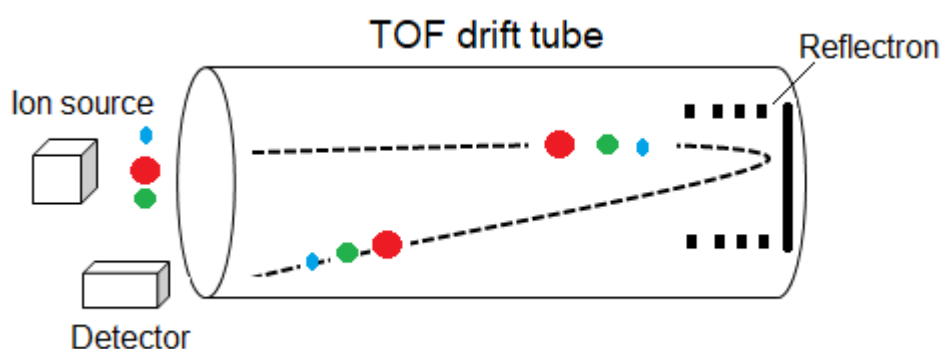


Figure 2-2 Schematic description of a TOF instrument equipped with a reflectron. Following ionisation, the sample ions will pass through TOF drift tube, the larger ions will travel slower than smaller ions and the time required for the ions to reach the detector is used to determine mass to charge ratio ( $m/z$ ) of the sample ions. A reflectron is used to focus the ions onto the detector and improves the resolution of the instrument.

Quadrupole mass analysers consist of four rods assembled in a symmetrical and parallel structure as shown in figure 2-3, opposite rods being paired, across which radio frequency (RF) and direct current (DC) voltage are applied (De Hoffmann, 2000).

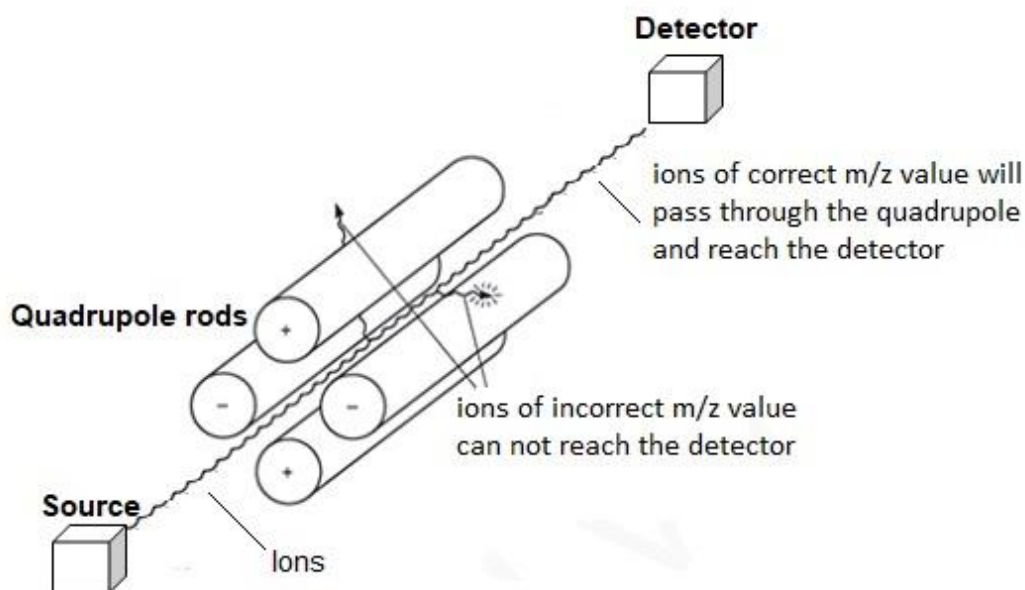


Figure 2-3 Diagram of Quadrupole Mass Analyser. At a particular RF and DC voltages, only the ion with a given  $m/z$  passes through the quadrupole field, reproduced from De Hoffmann, 2000.

The applied voltages influence ion trajectories. For a given DC and RF voltages, only ions of a particular  $m/z$  ratio will pass through the quadrupole while the other ions will be ejected. Manipulating the RF and DC voltages, while keeping their ratio constant, allows selection of a particular range of  $m/z$  ions to the detector, thus the quadrupole act as a 'mass filter'. Moreover, if the RF is the only voltage applied, then a wide range of oscillating ions will traverse through the quadrupole and in this case the quadrupole can be used as an ion focusing component for the next analyser in tandem mass spectrometers (De Hoffmann,

2000). Quadrupole mass analysers are relatively cheap and have high reproducibility. However, they have limited resolving power compared to TOF (Aebersold & Mann, 2003).

Ion traps are another type of analyser, they capture (hold) the ions in a physical volume for a period of time, the desired ions are selected, fragmented, and scanned all in the same component. Hence, they offer the advantage of tandem MS analysis without having multiple analysers (Aebersold & Mann, 2003).

Many instruments combine mass analysers of different types, referred to as hybrid instruments, to gain the strengths of each analyser and/or provide new or enhanced capabilities. The quadrupole ion trap-time of flight QIT-TOF, and quadrupole-time of flight (Q-TOF) mass spectrometer are two examples of hybrid instruments (Glish & Burinsky, 2008).

All the of MS work presented in this thesis was generated using a hybrid Q-TOF mass spectrometer; therefore it will be discussed in a little more detail. Q-TOF hybrid mass spectrometers were first introduced commercially in the late 1990s and remain popular today; typically these are equipped with an ESI or nESI source, although other ionisation sources can also be used (Chernushevich *et al.*, 2001, Ens & Standing, 2005). The instruments consist of ion focusing and collisional damping devices, a quadrupole mass filter, a gas-filled collision cell (which can be of quadrupole or hexapole geometry), and a TOF analyser (Chernushevich *et al.*, 2001). The Q-TOF can operate in MS and MS/MS mode; in MS mode the ions pass through a series of ion focusing devices and are separated in the TOF analyser. In MS mode, the quadrupole operates in RF only mode and functions as an ion guide (Chernushevich *et al.*, 2001). In MS/MS, ions of a particular  $m/z$  value are selected using the quadrupole filter; the selected ion population then passes through the collision

cell (RF only q2) that promotes ion fragmentation by collisions with an inert gas. Product ions are then separated in the TOF analyser (Chernushevich *et al.*, 2001). Thus, the Q-TOF instruments combine the ion guiding and isolating properties of a quadrupole mass filter with the high accuracy and performance of the TOF analyser. This combination results in accurate mass measurement, high resolution, and the ability to perform tandem mass spectrometry MS/MS experiments (Chernushevich *et al.*, 2001). Q-TOF instruments have been widely used for qualitative analyses and in the identification of post-translational modifications (Aebersold & Mann, 2003, Chernushevich *et al.*, 2001).

#### 2.1.6 Experimental Aims

CRP is a clinically important molecule; however, no prior work in the literature describes the glycation of CRP *in vitro* nor *in vivo*. Exploring CRP glycation in a complex biological sample (serum, plasma) without prior knowledge about the susceptibility of CRP to glycation is a difficult task. Consequently, CRP glycation study was first carried out by *in vitro* experiments with the aim to extrapolate the generated information to the *in vivo* situation, taking into account the factors that affect the glycation process under physiological conditions.

The aim of the work described this chapter was to prepare *in vitro* glycated samples of CRP, and to investigate the susceptibility of the target protein to glycation using electrospray mass spectrometry, and other biochemical methods to monitor CRP glycation as a function of concentration and exposure time to the glucose, with the aim to extrapolate the generated information to *in vivo* studies.



## 2.2 Materials & Methods

### 2.2.1 Reagents and consumables

Reagents, solutions and proteins used in this chapter and their suppliers are summarised in Table 2-1.

Table 2-1 Reagents, solutions and proteins used in the study. All substances were purchased in the highest possible purity from the following suppliers

Chemicals/proteins	Supplier
<b>30% T acrylamide/bisacrylamide</b>	Sigma-Aldrich
<b>Acetonitrile (HPLC grade, 99.9%)</b>	VWR International
<b>Ammonium bicarbonate</b>	Sigma-Aldrich
<b>Anti-mouse IgG (whole molecule) peroxidase antibody</b>	Sigma-Aldrich
<b>Ammonium persulfate (APS)</b>	VWR International
<b>Bromophenol blue</b>	VWR International
<b>Clarity western ECL blotting substrates</b>	Bio-Rad
<b>C-reactive protein (CRP, &gt;99% pure)</b>	SCRIPPS Laboratories
<b>D-glucose (99.5%)</b>	Sigma-Aldrich
<b>Formic acid (96%)</b>	Sigma-Aldrich
<b>Glycerol</b>	Sigma-Aldrich
<b>Glycine</b>	Sigma-Aldrich
<b>InstantBlue™ protein stain</b>	Expedeon
<b>Low range molecular weight markers</b>	Bio-Rad
<b>Marvel dried milk</b>	Local supermarket
<b>Monoclonal anti-C-reactive protein antibody (Clone-8)</b>	Sigma-Aldrich
<b>Ponceau S solution</b>	Sigma-Aldrich
<b>Sodium dodecyl sulfate SDS</b>	Sigma-Aldrich
<b>Sodium chloride</b>	Sigma-Aldrich
<b>Sodium phosphate dibasic (99%)</b>	Sigma-Aldrich
<b>Sodium phosphate monobasic (99%)</b>	Sigma-Aldrich
<b>Tetramethylethylenediamine (TEMED)</b>	Sigma-Aldrich
<b>Trifluoroacetic acid (98%)</b>	Sigma-Aldrich
<b>Tris</b>	Sigma-Aldrich
<b>Triton-X-100</b>	Sigma-Aldrich
<b>Water (HPLC grade)</b>	VWR International

### 2.2.2 Preparation of Glycated C-Reactive Protein

C-reactive protein (CRP, 99% pure) was obtained from SCRIPPS Laboratories and supplied as a solution in 20 mM Tris, 280 mM sodium chloride, 5 mM calcium chloride, pH  $8.0 \pm 0.1$ , with sodium azide added as a preservative. Protein concentration was determined by UV absorbance at 280 nm (Nanodrop 1000, Thermo Scientific). An extinction coefficient of 1.7 was used to give the concentration values in mg/mL (De Beer & Pepys, 1982).

*In vitro* glycation of CRP was carried out by exposure of the protein at a varying concentration of glucose up to 6 days. Aliquots of CRP were diluted to 0.2 mg/mL (8.6 pmol/ $\mu$ L) using 0, 30, 100, 300, 500 and 1000 mM glucose in 0.2 M phosphate buffer, pH 7.4, respectively. The aliquot diluted in the 0.2 M phosphate buffer alone was used as a control.

The samples were incubated at 37 °C with continuous shaking (TS-100 Thermo-Shaker, bioSan), and a portion was removed from each incubated sample after 19 hours, 3 days and 6 days. Unreacted glucose was separated from protein by ultrafiltration using Amicon ultra centrifugal filter devices (Millipore) fitted with a 10 kDa molecular weight cut-off membrane. The Amicon devices were pre-washed with 500 $\mu$ L of deionised water at 1000 x g for 5 minutes. After the washing step, the sample was then concentrated down at 14000 x g for 15 minutes, discarding the filtrate, then reconstituting the concentrate to 500  $\mu$ L with 50 mM ammonium bicarbonate, pH 7.4. This process (“washing out”) was repeated three times. The ultrafiltrate was removed, and the volume adjusted to 1  $\mu$ g/ $\mu$ L (43.4 pmol/ $\mu$ L) of the protein using ammonium bicarbonate buffer. The samples were stored at 4 °C until analysis.

### 2.2.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out to see if we could detect the glycosylated forms of CRP based on its mobility through the gel compared to that of non-glycosylated CRP standard. SDS-PAGE was performed according to the method of Laemmli (Laemmli, 1970), using a mini format 1-D electrophoresis system (Bio-Rad Laboratories, Inc.).

Gels were hand cast with acrylamide concentration of 4% and 15% (v/v) in the stacking and resolving gels, respectively (see Table 2-2). Low range molecular weight markers, purchased from Bio-Rad, were used for estimation of molecular weight. 5 $\mu$ L of each sample (containing 5  $\mu$ g protein) as prepared in section 2.2.2 was mixed 50% (v/v) with non-reducing loading buffer containing bromophenol blue (see Table 2-2) to visualise the dye front while the gel was running. All samples were heated at 95 °C for 5 minutes before being run on the gel. After loading the samples, the gel was placed vertically in a Mini-PROTEAN Tetra cell attached to a PowerPac Basic electrophoresis apparatus (Bio-Rad Laboratories, Inc.). The gel was run at 200V in Tris-glycine-SDS running buffer (see Table 2-2) for 45 minutes, or until the bromophenol blue tracking dye reached the lower edge of the gel. Gels were then removed from the tank, and the proteins in the gel were either transferred to a nitrocellulose membrane (Bio-Rad Laboratories) for western blot analysis or visualised using Coomassie InstantBlue™ stain.

The gels were soaked in a sufficient volume of InstantBlue™ Protein Stain to cover the membrane for a minimum of 30 minutes at room temperature. The excess dye was washed off with deionised water (destaining) once a sufficient level of staining had been achieved. Gels were photographed on a FluorChem imager (Alpha Innotech Corporation).

Table 2-2 Components and amounts required for SDS-PAGE experiment. The gel, sample buffer and the running buffers components are listed, the amounts are also given.

Gel/ buffer	Components
<b>Resolving Gel 15%</b> (given amounts are enough to cast two identical 0.75 mm SDS-PAGE gels)	<ul style="list-style-type: none"> <li>• 30 % T acrylamide/bisacrylamide 4.013 mL</li> <li>• Deionised Water 1.867 mL</li> <li>• Tris 1.5 M (pH 8.8) 2 mL</li> <li>• SDS 10% (w/v) 80 uL</li> <li>• TEMED 6 uL</li> <li>• APS 10% (w/v) 60 uL</li> </ul>
<b>Stacking Gel 4%</b> (given amounts are sufficient for two identical 0.75mm stacks)	<ul style="list-style-type: none"> <li>• 30 % T acrylamide/bisacrylamide 0.67 mL</li> <li>• Deionised Water 3 mL</li> <li>• Tris 0.5M (pH 6.8) 1.25 mL</li> <li>• SDS 10% (w/v) 60 uL</li> <li>• TEMED 6 uL</li> <li>• APS 10% (w/v) 60 uL</li> </ul>
<b>Running Buffer 1L</b>	<ul style="list-style-type: none"> <li>• Tris Base 3.03 g</li> <li>• Glycine 14.42 g</li> <li>• SDS 1 g</li> </ul>
<b>Non-Reducing sample Buffer</b>	<ul style="list-style-type: none"> <li>• Tris 0.5 M (pH 6.8) 1.25 mL</li> <li>• Glycerol 2.5 mL</li> <li>• Deionised Water 3.55 mL</li> <li>• SDS 10% (w/v) 2 mL</li> <li>• Bromophenol Blue 0.5% 0.2 mL</li> </ul>

#### 2.2.4 Western blotting

The samples prepared as described in section 2.2.2 were first subjected to SDS-PAGE as described in section 2.2.3 but without staining the gel. After electrophoresis, the gel, a nitrocellulose membrane blotting paper (0.2  $\mu$ m, 7x8.4 cm, pre-cut, Bio-Rad) and filter pads (Bio-Rad) were soaked for 1 hour in equilibration buffer (Table 2-3) and then assembled in cassette case of the Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad). The assembled cassette was placed in the transfer apparatus with an ice pack to prevent overheating during the transfer. The running buffer (Table 2-3) were then added, and a magnetic stirrer was used to maintain a consistent flow of running buffer throughout the tank. The cell was connected to a PowerPac Basic power supply (Bio-Rad). Proteins were transferred from the

gel to the nitrocellulose membrane by electroblotting at 350 mA for 1 hour at 4 °C. Post-transfer, the membrane was stained for 1-5 min with Ponceau S solution (0.1% w/v Ponceau S, 5% v/v acetic acid) to determine the quality of the transfer. When bands are visible, the membrane was destained by washing with deionised water to attain a clear background.

The membrane was then incubated for 5 hours at room temperature with agitation in a blocking solution of Tris-buffered saline (TBS), 2.5 % (w/v) defatted milk powder (see Table 2-3). Following blocking, membranes were washed sequentially five times for 5 minutes each with BLOTTO, (Table 2-3).

Table 2-3 Buffers and solutions used for western blotting transfer and staining. The components and their corresponding amounts are given.

Buffer/Solution	Components/Amount
<b>Running/Equilibration Buffer</b>	<ul style="list-style-type: none"> <li>• SDS 0.5 g</li> <li>• Tris 7.21 g</li> <li>• Glycine 1.515 g</li> </ul>
<b>TBS pH 7.4 with HCl</b>	<ul style="list-style-type: none"> <li>• Sodium chloride 150 mM</li> <li>• Tris 50 mM</li> </ul>
<b>Blocking Solution</b>	<ul style="list-style-type: none"> <li>• Marvel dried milk 1 g</li> <li>• TBS 40 mL</li> </ul>
<b>BLOTTO</b>	<ul style="list-style-type: none"> <li>• Marvel dried milk 0.5 g</li> <li>• Triton-X-100 1 mL</li> <li>• TBS 50 mL</li> </ul>
<b>Primary Antibody Dilution</b>	<ul style="list-style-type: none"> <li>• Monoclonal anti-CRP antibody 6 µL</li> <li>• BLOTTO up to 24 mL</li> </ul>
<b>Secondary Antibody Dilution</b>	<ul style="list-style-type: none"> <li>• Anti-mouse IgG peroxidase antibody 1 µL</li> <li>• BLOTTO up to 40 mL</li> </ul>
<b>ECL Reagent</b>	<ul style="list-style-type: none"> <li>• Clarity peroxidase reagent 7 mL</li> <li>• Clarity western luminol/enhancer reagent 7 mL</li> </ul>

The primary antibody solution was added in a 1:1000 (v/v) dilution in BLOTTO to the membrane and incubated overnight at 4 °C with agitation. The primary antibody used was a monoclonal mouse anti-C-reactive protein antibody (Clone CRP-8), provided in ascites fluid. The antibody concentration reading was 9.1 mg/mL; stocks were stored at -20 °C and thawed to room temperature before use. Following incubation with primary antibody, membranes were washed as previously with BLOTTO (5 minutes x five washes). Polyclonal rabbit anti-mouse IgG antibody, conjugated with HRP, was added to the membrane in a 1:40,000 dilution in BLOTTO and incubated for 1h at room temperature with agitation. The secondary antibody was decanted, and the membrane washed.

Chemiluminescence (Clarity Western ECL Substrate, Bio-Rad) enabled visualisation of immunoreactive bands. HRP peroxide reagent and luminol/enhancer were mixed 1:1 (v/v) immediately prior to being added to membranes, then incubated for a minimum of 5 minutes with gentle agitation. Membranes were wrapped in Saran Wrap (protein side up) to prevent drying out and visualised using a FluorChem imager (Alpha Innotech Corporation), with exposure times of up to 30 minutes, depending on observed band intensity.

#### 2.2.5 Liquid Chromatography-Mass Spectrometry

Liquid chromatography-mass spectrometry (LC-MS) of intact protein was carried out to detect protein glycation and monitor the reaction as a function of the net gain in the molecular weight of the protein following glucose conjugation. The molecular masses of the intact proteins in the samples were determined by using Ultimate 3000 nano-flow high-performance liquid chromatography (HPLC) system (Dionex Ultimate 3000, Thermo Fisher Corporation) coupled with a Premier hybrid quadrupole-time of flight mass spectrometer

Q-TOF (Waters Corporation). The HPLC system consists of a pump, microflow controller and autosampler maintained at 4 °C. Proteins were separated using a C4 analytical column (Acclaim PepMap300, ID 75 µm, length 15 cm, particle size 5 µm, pore size 300 Å, Dionex) preceded by a C4 guard/trapping column (Acclaim PepMap300, ID 0.3 mm, length 5 mm, particle size 5µm, pore size 300 Å, Dionex) to concentrate protein and wash away buffer salt. The HPLC solvents were solvent A, consisting of 0.1% formic acid (FA), 5% acetonitrile (ACN) and solvent B, consisting of 0.1% FA, 95% ACN. The Q-TOF Premier mass spectrometer was equipped with a nano-flow electrospray ionisation source (nESI) which produces the electrospray through a PicoTip needle (10 µm i.d., New Objectives, Woburn, MA, USA), the electrospray was optimised by adjusting both capillary voltage and the x-y-z position of the nESI emitter relative to the front of the Q-TOF sample cone. The software used to operate the Ultimate 3000 HPLC system, and the Q-TOF Premier were Chromeleon Client version 6.8 (Thermo Fisher Corporation) and MassLynx 4.1 (Waters Corporation), respectively.

For intact protein mass spectrometric analysis, the samples described in section 2.2.2 were diluted with solvent A to a final concentration of 5 pmol/µL. Samples were centrifuged for 3-5 minutes to remove any precipitate, then transferred to autosampler vials (Waters Corporation). The vials were loaded onto autosampler, and an aliquot of solvent A was also loaded to be used as a blank.

Sample runs were bracketed by blank runs to ensure full equilibration of columns and to minimise system carryover and consequent column contamination. For each HPLC separation, 10 pmol (2 µL) of protein solution was injected into the 20 µL loop without prior purification. Protein samples were loaded to the guard column over 5 minutes using 95%

solvent A, 5% solvent B (30  $\mu\text{L}/\text{min}$ ). After the trap/wash period, the sample was loaded onto the reversed phase column at a split flow rate of 0.35  $\mu\text{L}/\text{min}$ . The gradient consisted of 2 ramps; 15 minutes (5% - 60% solvent B), then 5 minutes (60-95% solvent B). The column was then washed at 95% solvent B for 5 minutes before re-equilibration to the starting conditions.

The HPLC column was coupled online to the Q-TOF mass spectrometer via nESI, with typical capillary voltages of 1-2 kV being applied. The Q-TOF was operated in MS mode. The Q-TOF parameters were as follow; sample cone and the extractor cone voltages were 45 V and 5 V, respectively, source and desolvation temperatures 80 and 180  $^{\circ}\text{C}$ , cone gas flow and the desolvation gas flow were 50 and 400 L/h, respectively. The collision energy setting was set at 5-V. The scan  $m/z$  range started at 500 and ended at 2500 in positive ion mode, scan time 1 s, 0.1 s interval. Data were collected and processed using MassLynx V4.1 (Waters). The multiply charged electrospray spectra were deconvoluted (after background subtraction) using the Maximum Entropy 1 package of the MassLynx software, with a resolution of 1 Da/channel, a uniform Gaussian width at half height of 0.5 Da, and minimum intensity ratios of 33%.

#### 2.2.6 Calculation of glycoforms distribution percentage, the extent of glycation and glucose index

Depending on peak intensities, the percentage of each glycosylated form of CRP was estimated as an approach to determine the relative quantity of each form of the protein. LC-MS analysis of each sample was carried out in triplicate, and the average intensities of the peaks in the deconvoluted LC-MS spectra corresponding to different species of CRP were calculated. From the calculated average intensities, the percentage of glycosylated CRP species



was determined by dividing the average intensity for each peak corresponding to CRP with  $i$  number of attached glucose, by the sum of average intensities of all peaks in the spectrum that corresponds to the glycosylated and non-glycosylated forms. The obtained ratio was multiplied by 100.

$$\% \text{ of each CRP glycoform } (i * \text{ glucose}) \text{ in the sample} = \frac{I (i * \text{ glucose})}{N} * 100\%$$

where ( $I$ ) is the average intensity of the CRP peak with  $i$  number of attached glucose,  $N$  is the summation of the average intensities of CRP peaks that correspond to different species in the spectrum.

The extent of CRP glycosylation for each sample was calculated by summing up the individual glycoforms percentages.

$$\text{Extent of glycosylation} = \sum_{i=1}^n \% \text{ of CRP glycoforms } (i * \text{ glucose})$$

To further illustrate the tendency of CRP glycosylation over incubation time, the average glucose index was calculated.

Glucose Index (G-idx), can be defined as the average number of glucose residues per molecule of CRP. It was calculated according to the following equation (Gadgil, Bondarenko, Pipes *et al.*, 2007, Miller *et al.*, 2011a):

$$G - \text{idx} = \frac{\sum_{i=0}^n i * I (\text{intact CRP} + i * \text{ glucose})}{\sum_{i=0}^n I (\text{intact CRP} + i * \text{ glucose})}$$

where  $I$  is the average intensity of various glycosylated CRP species, and  $i$  is the number of glucose units attached to in each peak of the deconvoluted spectrum.

## 2.3 Results

*In vitro* glycation of CRP was carried out with various concentrations of glucose and incubation time to understand the effects of glycating agent and the exposure time *in vitro*. The protein was treated with 30 mM to 1 M glucose (the primary glycating agent in human serum) in phosphate buffer at pH 7.4. Phosphate buffer is a commonly used buffer for *in vitro* glycation reactions (Priego Capote & Sanchez, 2009). Since the glycation process relies on several factors including the functional lifetime protein (Schleicher & Wieland, 1986), the protein was exposed to the glycating agent for a period equivalent to its half-life in the human serum at 37 °C. Longer incubation times were used to enable observation of higher order glycation products.

### 2.3.1 Preliminary characterisation of glycated samples using SDS-PAGE and western blotting

Glycated CRP samples prepared as per section 2.2.2 were subjected to SDS-PAGE and western blotting to assess the impact of *in vitro* glycation of CRP on its apparent molecular weight or interaction with antiserum.

Figure 2-4 shows the SDS-PAGE mobility of CRP samples incubated at 30, 100, 300, 500 and 1000 mM glucose for 19 hours (Figure 2-4, A), three days (Figure 2-4, B) and for 6 days (Figure 2-4, C). An untreated CRP sample was included as a control (lane 2 in figure 2-4). The control sample produced a single band at a molecular weight of ~23 kDa, corresponding to the molecular weight of monomeric protein. In all glucose incubated samples, a single band was also observed. Following 19 h incubation of CRP with different concentrations of glucose, no difference in electrophoretic mobility was observed compared to control (Figure 2-4, A). However, after three days of incubation, CRP incubated with 500 and 1000 mM of glucose tend to produce a broad and diffused band

compared to the control (lanes 6 and 7 in figure 2-4, B). Further increase of incubation time to 6 days results in a minor difference in migration patterns of the glucose treated samples, where the glycosylated samples show a slower migration through the gel in parallel with the increase in glucose concentration (Figure 2-4, C). The difference in electrophoretic mobility was more clear with the samples incubated with 500 and 1000 mM compared to control.

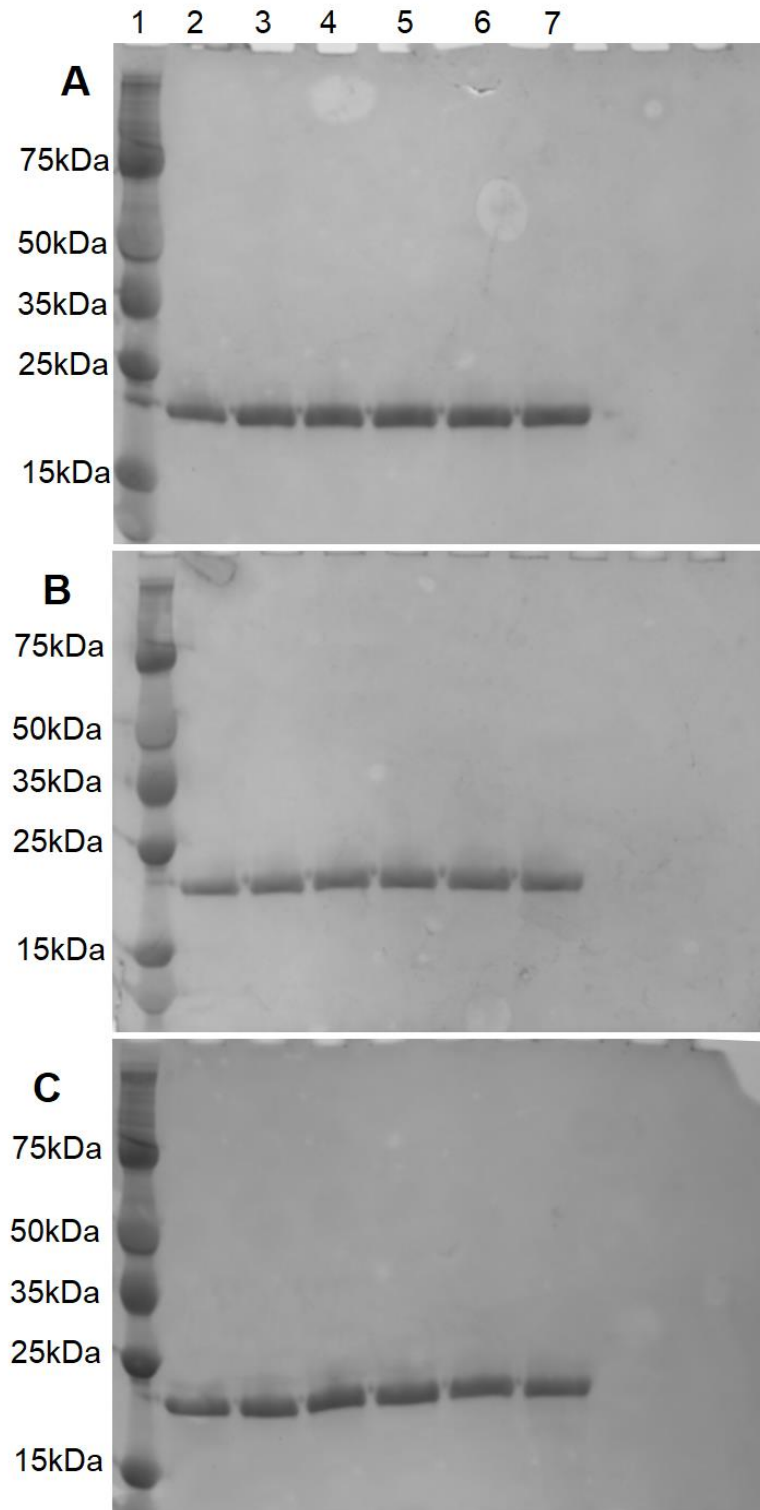


Figure 2-4 SDS-PAGE of *in vitro* glycosylated CRP Samples. Native CRP (0.2 mg/mL) were incubated with 0 mM glucose (Lane 2), 30 mM glucose (Lane 3), 100 mM glucose (Lane 4), 300 mM glucose (Lane 5), 500 mM glucose (Lane 6), or 1000 mM glucose (Lane 7), n=6. The sample incubations were carried out at 37 °C in 0.2 M phosphate buffer, pH 7.4, for 19 hours (A), 3 days (B), and for 6 days (C). Lane 1 represents the molecular weight markers. 5  $\mu$ L of each sample (containing 5  $\mu$ g of the protein) were loaded and analysed by 4-15% SDS- PAGE, the Gels were stained with Coomassie brilliant blue.

A single band was also seen when the glycosylated samples were subjected to western blotting, figure 2-5 is the result of western blotting of CRP samples incubated at 0, 30, 100, 300, 500 and 1000 mM glucose for 19 hours (figure 2-5, A), three days (figure 2-5, B) and for 6 days (figure 2-5, C). The intensity of the bands produced by glycosylated CRP samples were relatively comparable to that produced by non-glycosylated CRP control suggesting the reactivity of the monoclonal anti-CRP antibody toward the glycosylated samples was comparable to that toward the control.

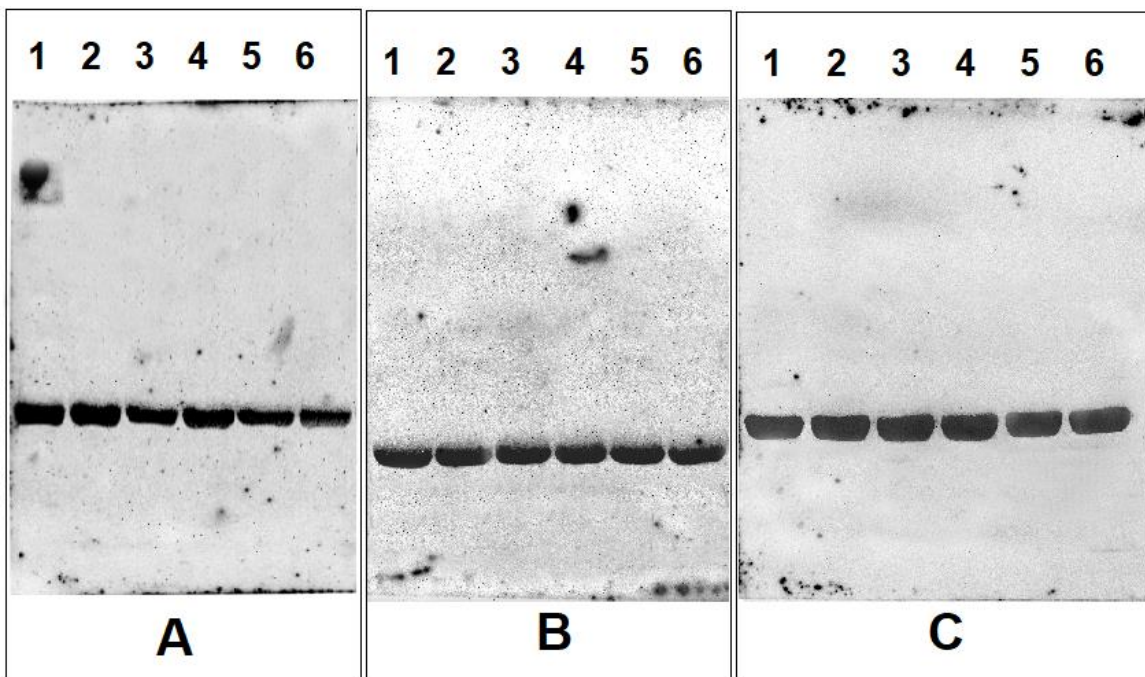


Figure 2-5 Western blot of the *in vitro* glycosylated samples. Native CRP (0.2 mg/mL) were incubated with 0 mM glucose (Lane 1), 30 mM glucose (Lane 2), 100 mM glucose (Lane 3), 300 mM glucose (Lane 4), 500 mM glucose (Lane 5), or 1000 mM glucose (Lane 6), for a period of 19 hours (A), 3 days (B), and for 6 days (C). The primary mAb dilution was 1:4,000. The secondary mAb dilution was 1:40,000. Samples were exposed for up to 30 minutes.

### 2.3.2 Characterisation of CRP glycation using LC-MS

A typical spectrum of untreated CRP as observed by nESI-Q-TOF-MS is shown in figure 2-6, A. The charge states of the protein range from 10+ to 21+, after a software-assisted deconvolution of the spectrum, the mass spectrum depicted in figure 2-5, B was obtained, the deconvoluted spectrum shows a major mean species at 23027 Da, which corresponds to the expected mass of a single CRP subunit.

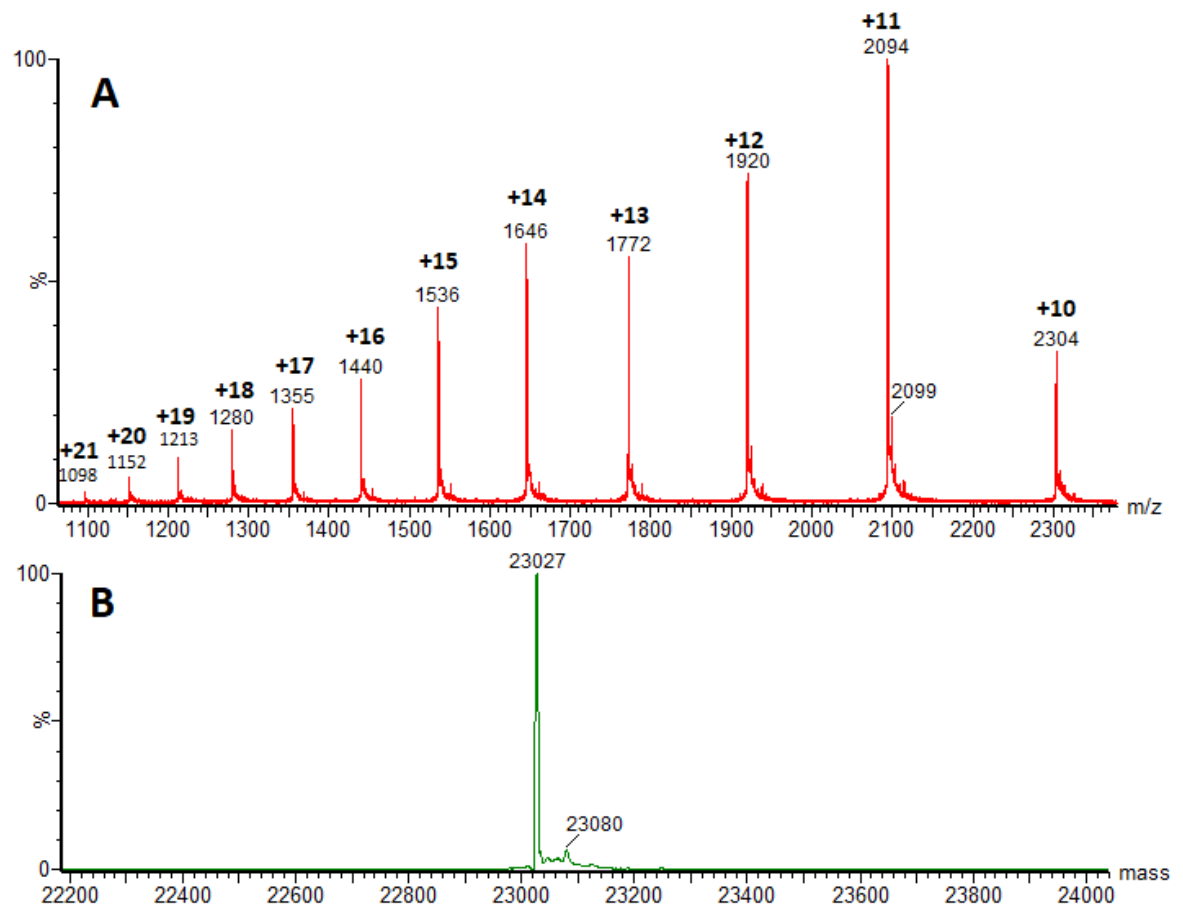


Figure 2-6 Mass spectrum of native non-glycosylated CRP as measured by nESI-QTOF-MS in positive ion modes. A, Combined  $m/z$  spectrum after LC-MS analysis of 10 pmole of standard non-glycosylated CRP. The charge states are indicated in the spectrum, and the intensities are scaled to the highest peak, B: Deconvoluted LC-MS mass spectrum of CRP based on the  $m/z$  spectrum presented in figure A.

### 2.3.2.1 LC-MS analysis of CRP samples incubated for 19 hours

The deconvoluted mass spectra of CRP samples incubated with different glucose concentration for 19 hours at 37 °C, pH 7.4. are shown in Figure 2-7. The deconvoluted spectrum generated following glycation of CRP with 30 mM glucose, Figure 2.7, A, demonstrate the emergence of a new low intensity signal, observed mass 23,189 Da in addition to the unmodified protein observed at 23,027 Da. The mass increment of 162 Da between emergent and original peak corresponds to the addition of a single dehydrated glucose molecule to the protein, suggesting that the new peak is the representative of a mono-glycosylated form of CRP obtained via the Maillard reaction. This mono-glycosylated CRP signal was observed in all other samples incubated for the same duration; increasing peak intensity corresponded with increasing glucose concentration (Figure 2-7).

Incubating the sample with 100 mM glucose resulted in increased intensity of the mono-glycosylated form to the double compared to 30 mM sample (Figure 2-7, B). while the spectra obtained from the sample glycosylated by 300 mM of glucose, indicate an additional signal at 23,351 Da in addition to the mono-glycosylated signal at 23,189 Da and the unmodified peak at 23,027 Da (Figure 2-7, C). This new peak represents a mass increase of 324 Da compared to the unmodified protein molecular weight and this equal to the net mass of two glucose residues, suggesting that this peak is a di-glycosylated form of the protein. This di-glycosylated peak is more evident in the sample incubated with 500 mM glucose (Figure 2-7, D), and its intensity increased an additional fourfold when the glucose concentration was increased to 1 M (Figure 2-7, E). At 1 M glucose, the intensity of the mono and di-glycosylated forms increased significantly, with the signal intensity of the mono-glycosylated form reaching >70% the intensity of the residual unmodified form. Furthermore, a trace signal at 23,513 Da, representing conjugation of three glucose molecules, was also observed (Figure 2.7, E).

With increasing glucose concentration in the sample, the intensity of the unmodified protein signal decreases (Figure 2-7) in parallel with the increasing intensities of the glycosylated forms of the protein.



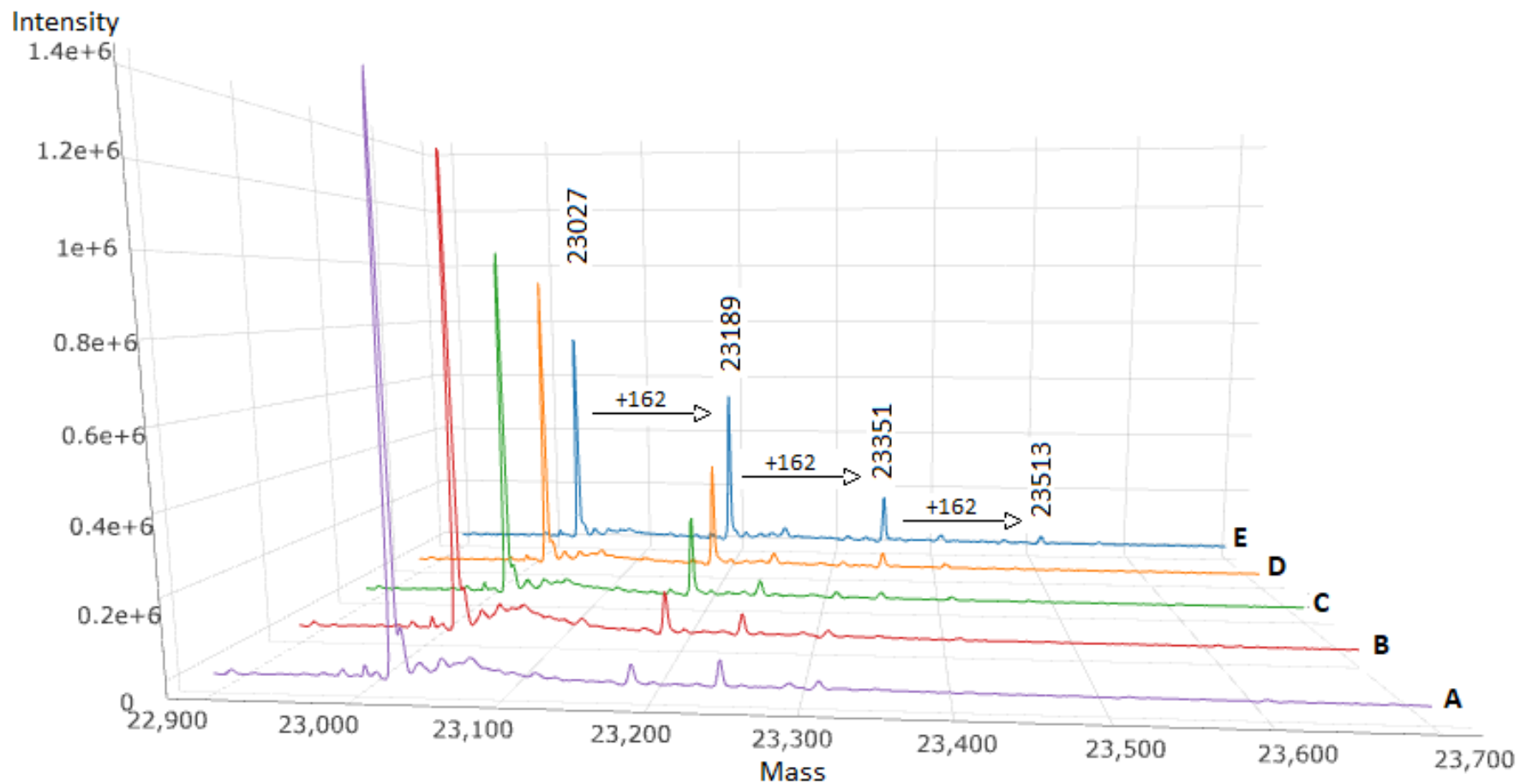


Figure 2-7 The deconvoluted mass spectra of CRP samples incubated with different concentrations of glucose for 19 h. A; with 30 mM glucose, B; with 100 mM glucose, C; with 300 mM glucose, D; with 500 mM glucose, and E; with 1000 mM glucose. The sample incubations were carried out at 37°C in 0.2 M phosphate buffer, pH 7.4 for 19 hours.

### 2.3.2.2 LC-MS analysis of samples incubated for 3 days

After three days of exposure of CRP to different concentrations of glucose, the deconvoluted spectra in Figure 2-8 were obtained.

At 30 mM glucose level (Figure 2-8, A), the peak at 23,189 Da is identified; however, its relative intensity is double compared to that from the sample incubated with the same glucose concentration but for 19 hours (Figure 2-7, A). On the other hand, a trace signal for the di-glycosylated CRP peak was observed using 100 mM glucose, (Figure 2-8, B). Moreover, the sample incubated with 300 mM glucose for 3 days (Figure 2-8, C) showed a peak distribution similar to that produced by 1000 mM glucose after 19 h (Figure 2-7, C), with conjugation of up to three glucose molecules to the protein. At 500 mM glucose (Figure 2-8, D), the mono-glycosylated CRP peak intensity was almost identical to that of the non-modified protein. Furthermore, a trace signal at 23,675 Da, with a mass difference of 648 Da, indicates condensation of four glucose molecules.

At 1 M glucose, the base peak is for the mono-glycosylated CRP, with the di-glycosylated form being the second most intense peak in the spectrum while the tri-glycosylated CRP peak is three times higher compared to the 500 mM glucose level after same incubation time. The tetra-glycosylated CRP is more intense, and a trace signal is observed at 23,837 Da, consistent with the addition of five glucose monomers.

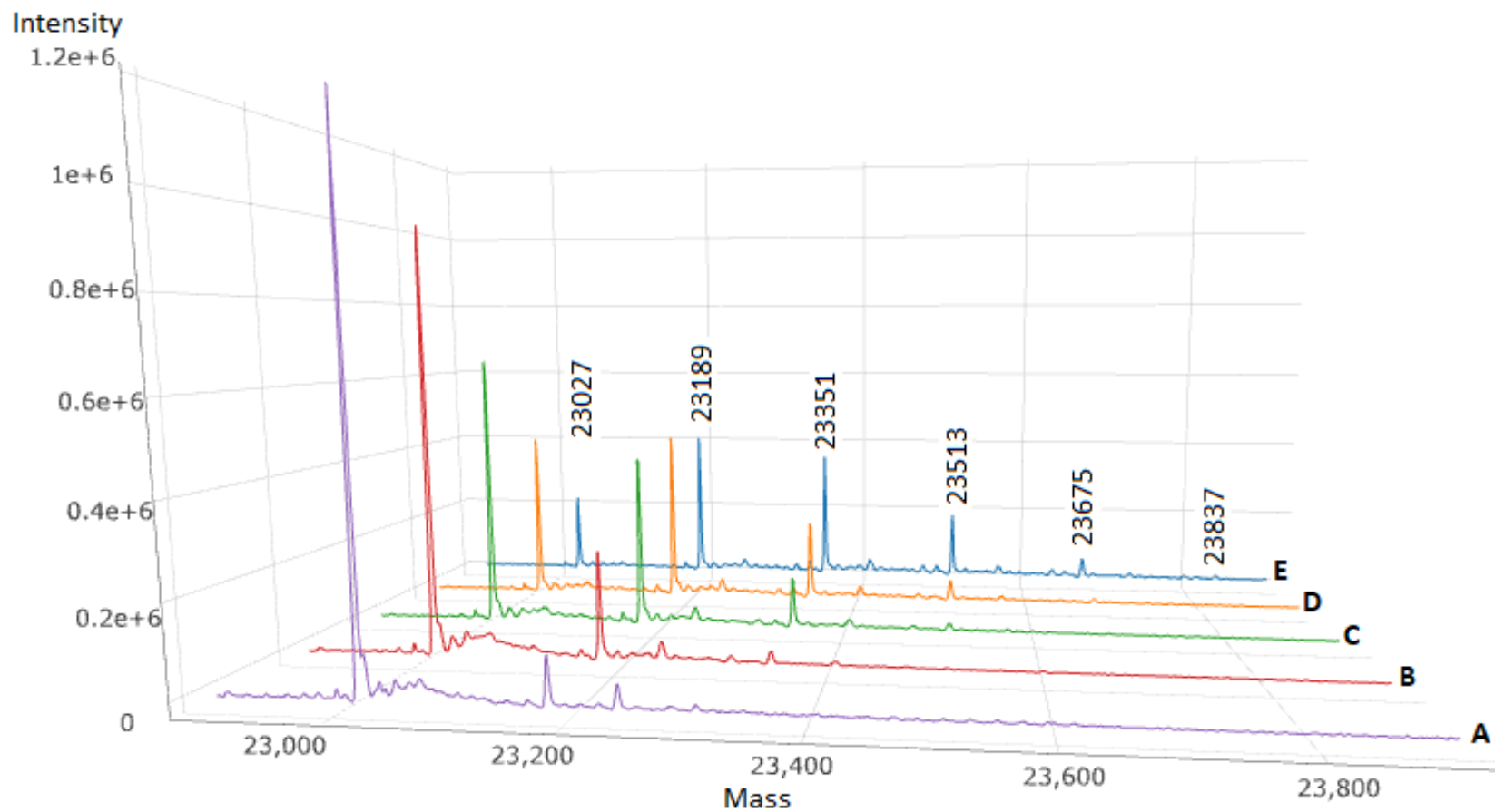


Figure 2-8 The deconvoluted mass spectra of CRP samples incubated with different concentrations of glucose for 3 days. A; with 30 mM glucose, B; with 100 mM glucose, C; with 300 mM glucose, D; with 500 mM glucose, and E; with 1000 mM glucose. The sample incubations were carried out at 37°C in 0.2 M phosphate buffer, pH 7.4 for 3 day

### *2.3.2.3 LC-MS analysis of samples incubated for 6 days*

Following six days incubation, at 30 mM concentration of glucose (Figure 2-9, A), the peak that represents the unmodified form of the protein remains base peak in resultant spectra, but the distribution between un- mono- and higher glycosylated forms shifts with increasing glucose concentration, with a marked alteration above 300 mM glucose (Figure 2-9, C). The intensity of the unmodified peak is barely detectable above background at a glucose concentration of 1 M (Figure 2-9, E).

Reduction in the intensity of the unmodified protein corresponded with the appearance of new glycoforms and an increase in the intensity of previously observed glycosylation peaks. The peak representing di-glycosylated CRP was barely detectable at 30 mM glucose level (Figure 2-9, A), but increased six-fold with 100 mM glucose level, alongside the emergence of tri-glycosylated protein signal (Figure 2-9, B). At 300 mM glucose (Figure 2-9, C), the mono-glycosylated peak is the prominent peak in the spectrum. However, the intensity of this peak gradually decreased in proportion as the glucose concentration further increased; in parallel, there was an increase in the intensities of multi-glycosylated forms with higher glucose levels.

The peaks of the di- and tri-glycosylated CRP were the major peaks in the deconvoluted spectrum of the sample incubated with 500 mM glucose (Figure 2-9, D), the intensity of these two peaks exceeded fourfold the intensity of the peak corresponding to non-glycosylated CRP. Furthermore, the intensity of tetra-glycosylated CRP is almost identical to that produced by the mono-glycosylated CRP, and both are more than two fold higher than that provided by the unmodified protein (Figure 2-9, D). At this level of glucose (500 mM) signals at 23,837, and 23,999 Da were quite evident and represented the addition of five and six glucose

molecules to the protein, respectively, beside a trace signal at 24,162 Da that could represent the addition of seven glucose molecules (Figure 2-9, D).

At the highest glucose level (1 M), most of the protein is glycosylated, and a series of glycosylated forms of the protein are seen (Figure 2-9, E). The peaks representing the tri and tetra-glycosylated form of the protein dominate the spectrum while the di- and penta-glycosylated forms are the second dominant forms; both are more than two fold higher than the mono-glycosylated form. The peak at 24,162 Da is quite evident demonstrating the addition of seven glucose molecules to the protein. Additionally, a very low signal at 24,323 Da can be recognised at this glucose level to report the condensation of eight glucose molecule to the protein (Figure 2-9, D).

Notably, there is a minor peak with an average mass of 23,080 Da (+53 Da) was a common feature in the mass spectra of most of the CRP preparations including the non-glycosylated sample as shown in figure 2-6 B. The above peak is found to follow the pattern of CRP glycosylation; it is thought to be an experimental artefact due to the ingredient impurities or a result of a combination modification. The odd mass difference (+53 Da) between the minor peak at 23,080 Da and the main peak at 23,027 suggest presence of 1 or 3 ammonium adducts.

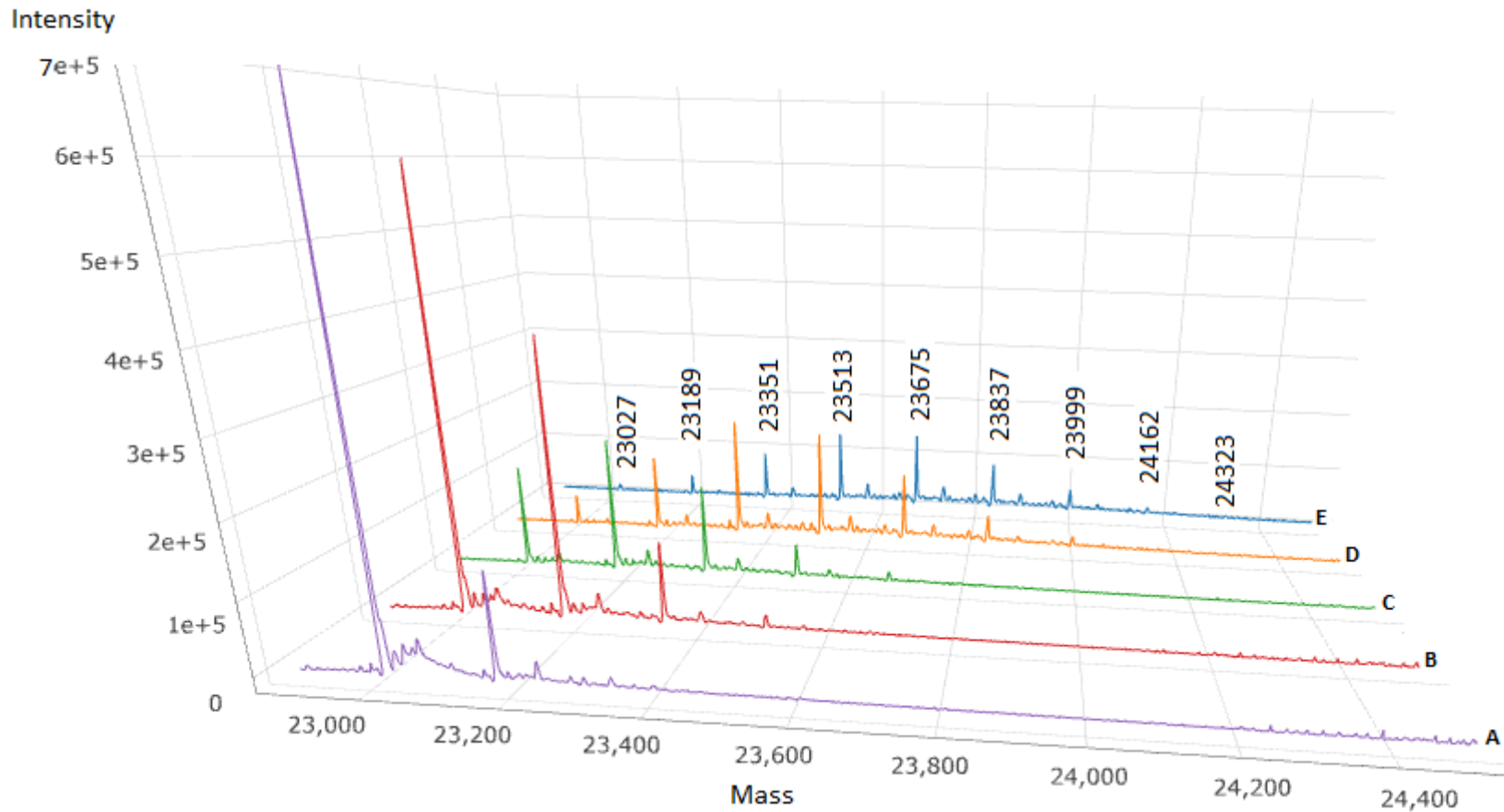


Figure 2-9 The deconvoluted mass spectra of CRP samples incubated with different concentrations of glucose for 6 days. A; with 30 mM glucose, B; with 100 mM glucose, C; with 300 mM glucose, D; with 500 mM glucose, and E; with 1000 mM glucose. The sample incubations were carried out at 37°C in 0.2 M phosphate buffer, pH 7.4 for 6 days.

### 2.3.3 The percentage distribution of glycoforms and the extent of glycation

The LC-MS analysis of each sample was carried out in triplicate. The peak intensity for a given species averaged and used to calculate the extent of glycation and the proportion of each glycoform for each sample, the calculations were based on the assumption that signal intensity represents the relative amount of the protein in the sample, the results of these calculations are presented in Figure 2-10.

As shown in Figure 2-10, A, the extent of glycation of CRP incubated for 19 hours increased with increasing glucose concentration, from 3.7% of the protein signal at 30 mM glucose to 48.8% at 1000 mM glucose. At this incubation time, the mono-glycated form is the predominant glycosylated signal, at all glucose concentrations; its percentage increases in parallel with increasing glucose concentration from 3.7% at 30 mM of glucose to 35.7% at 1000 mM. The di-glycated form represented only 1.9% of the protein signal at 300 mM glucose, increasing to 3.6% at 500 mM glucose and further increased to 11.1% at 1000 mM glucose. The tri-glycated form was seen only at 1000 mM glucose, where it represented 2% of the protein signal.

Increasing the incubation time to 3 days doubled the extent of glycation of CRP compared to 19 hours incubation (Figure 2-10, B). Again, the extent of glycation here increased in parallel with the increase in glucose concentration. The mono-glycated form is also the main glycosylated form during this incubation period, However, the percentage of the mono-glycated form started to fall after 500 mM (from 38.4% at 500 mM of glucose to 32.9 at 1000 mM of glucose), but it remained higher than the second highest glycosylated form, the di-glycated form, which counts 29% of the protein signal. The di-glycated form increased from 2.5% at 100 mM glucose and in parallel with the increasing glucose to reach 29% at

1000 mM glucose level. The tri-glycosylated form started at 1.8% at 300 mM and reached 14.7% with 1 M glucose. The tetra-glycosylated form started at 1.2% at 500 mM but increased to 4.9% at 1000 mM, condensation of five glucose molecules was seen only at 1000 mM and represent 1.2% of the protein.

After 6 days of the incubation (Figure 2-10, C), heterogeneous mixtures of glycosylated forms were detected, and almost complete removal of the un-glycosylated CRP peak was observed at the highest dose of glucose. Again, the observed glycosylation occupancy increased with increasing glucose concentration. However, the increment is only slight after 500 mM; from 93% at 500 mM to 97.6% at 1000 mM. The mono-glycosylated form is the predominant form in the samples incubated with 30, 100 and 300 mM of glucose (17.6%, 34.6% and 36%, respectively), while the di-glycosylated form represents the second most prevalent form. At 500 mM glucose, the di-glycosylated form becomes the dominant form (26.8%) and followed by the tri-glycosylated form (24.3%). At 1000 mM glucose, the tri and tetra-glycosylated forms were the primary forms (23.7% and 23.6%) while the di-glycosylated and the tetra-glycosylated forms are the second dominant (16.2% and 15.4%), the mono-glycosylated form percentage fall to 6.8%.



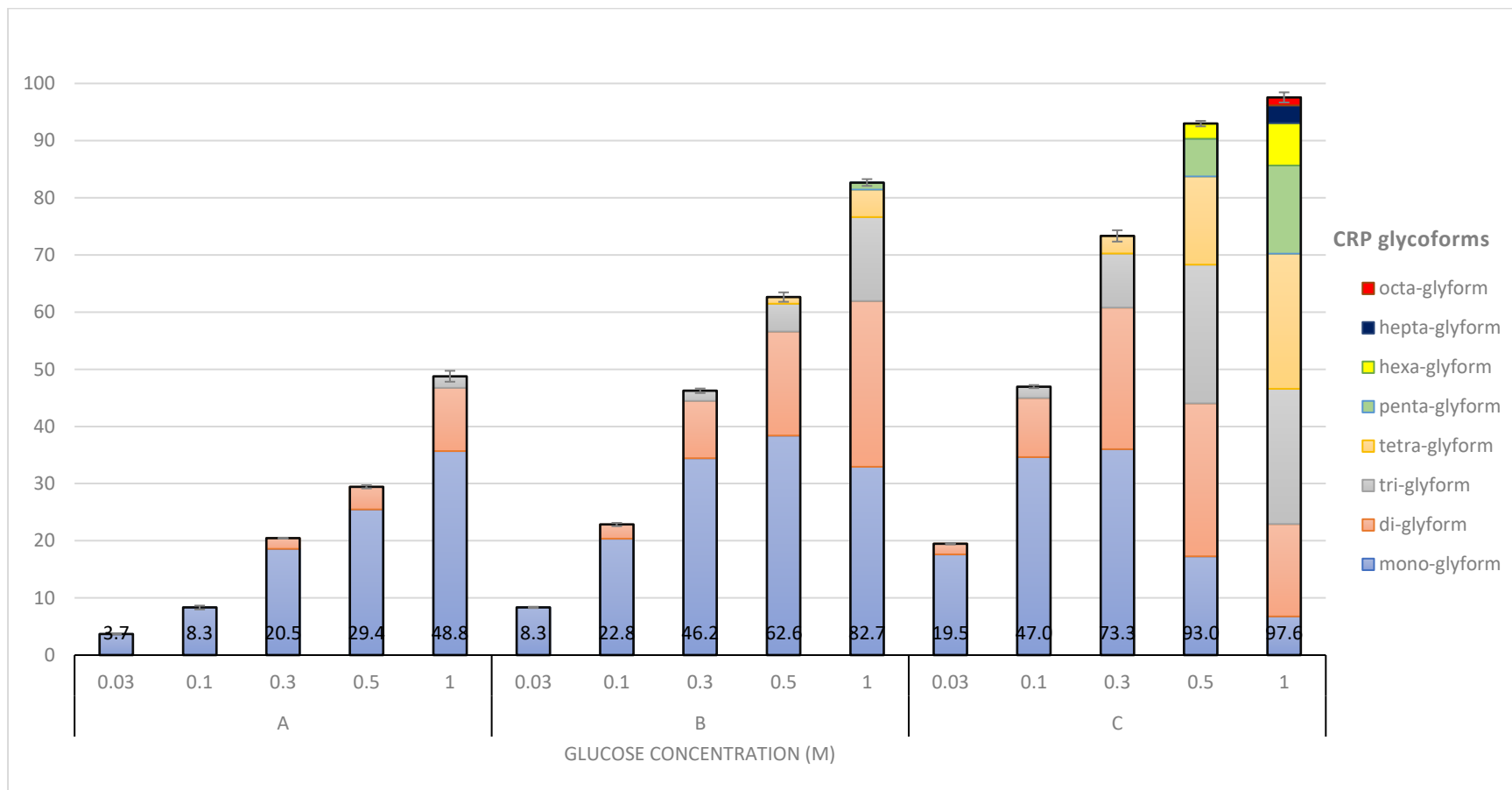


Figure 2-10 The extent of glycation and the percentage of glycoform signal distributions of *in vitro* glycated samples. CRP samples were incubated with different concentration of glucose for A: 19 hours, B: 3 days, C: 6 days. The columns represent the extent of glycation for each sample, the values placed at the base of the column while the error bars represent the standard deviation of the extends of glycation. Different coloured divisions represent individual glycoform percentages in each sample.

### 2.3.4 Glucose Index (G-idx)

The G-idx, which is the average number of glucose residues per molecule of CRP, was calculated as detailed in section 2.2.6 and plotted as a function of time, the results are presented in the Figure 2-11.

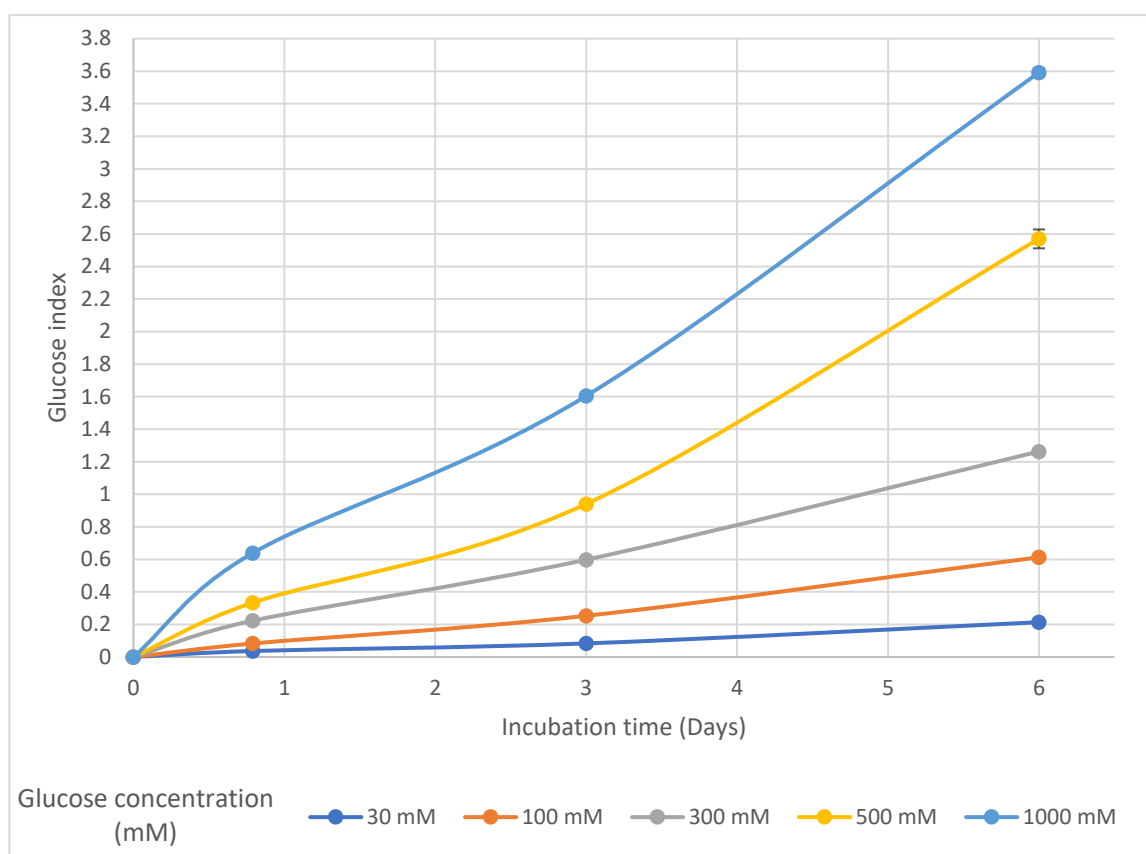


Figure 2-11 The G-idx of the *in vitro* glycated samples versus the incubation time. Different coloured lines represent the glucose concentrations used to produce glycated isoforms, the error bars (too low for the majority of the values) represent the standard deviation of G-idx

For all the *in vitro* glycated samples, the G-idx shows a direct relationship with the further increase in the incubation time; the longer the incubation time, the larger numbers of glucose attached per CRP molecule.

G-idx were less than one for the majority of the samples; sample incubated with 1000 mM for 19 hours, samples incubated with 500 mM or less for 3 days, and samples incubated with 100 mM or less for up to 6 days. More than 1 values were seen only when the samples exposed to 1000 mM of glucose for 3 days, or 300 to 1000 of glucose for 6 days. The sample incubated with 1000 mM of glucose for 6 days has 3.6 glucose molecules per each CRP molecules.

## 2.4 Discussion

Despite extensive studies having been previously carried out on CRP, no information in the literature is available regarding the existence of glycosylated versions of CRP either *in vivo* nor *in vitro*. To understand the glycosylation process in more detail and to be able to interpret the mass differences that could be found *in vivo*, *in vitro* glycosylation of CRP has been carried out in defined experimental conditions with the aim to extrapolate the generated information to the *in vivo* situation.

The glycosylation process relies on several factors including the half-life of the protein (Schleicher & Wieland, 1986), type of glycosylation agent and its concentration (Eble *et al.*, 1983, McPherson *et al.*, 1988), pH, type of buffer and the temperature (Kańska & Boratyński, 2002, Smith & Thornalley, 1992). Two factors, whose influence were taken into consideration, were the glucose concentration and the incubation time, while other factors were kept constant. Different glucose concentrations were used (from 30 mM to 1 M), and the incubation time ranged from 19 hours, which is the half-life of the protein in serum, to six days, to increase the opportunity of observation of reaction products. The protein-glucose mixture was prepared using phosphate buffer (pH 7.4), since other buffers or different pH may increase protein glycosylation as a result of denaturation and revealing the unexposed potential glycosylation sites (Lee, *et al.*, 2011). All glycosylation experiments were carried out at 37 °C, similar to the physiological temperature as higher temperature results in an increased rate of glycosylation (Gadgil, *et al.*, 2007).

#### 2.4.1 Preliminary characterising of *in vitro* glycated CRP with SDS-PAGE and western blotting

SDS-PAGE is one of the simplest and most widely used technique in proteomics, and has been used in the analysis of protein glycation for detecting or for separating the glycation product for further analysis (Galván-Moroyoqui *et al.*, 2016, Bai *et al.*, 2012, Ledesma-Osuna *et al.*, 2008, Chen *et al.*, 2012). In SDS-PAGE proteins are separated according to their molecular weight, depending on their differential rates of migration through a sieving matrix (a gel) under the influence of an applied electric field (Rabilloud *et al.*, 2009). In this work, SDS-PAGE has been used to see to assess if there is an apparent increase in the molecular weight of the protein after treating it with glucose.

CRP incubated with different concentrations of glucose for a period equal to the half-life of the protein in human serum revealed a single band with no apparent difference in their mobility as compared to the control (Figure 2-4, A). When the incubation time was increased to three days, broader bands with less distinct edges were associated with the samples treated with 500 and 1000 mM of glucose (Figure 2-4, B). These broader bands are suggestive of the behaviour of individual molecules as a distribution of molecular masses (Ledesma-Osuna *et al.*, 2008, Chen *et al.*, 2012).

Further increasing incubation time to 6 days resulted in a slight decrease in the electrophoretic mobility of the samples treated with the higher glucose concentrations (300, 500, and 1000 mM of glucose) as compared to the control (Figure 2-4, C), this decrease in electrophoretic mobility is strongly suggestive of an increase in CRP molecular weight and, accordingly, the occurrence of glycation.

SDS-PAGE alone has limited resolution, meaning unless the samples are profoundly different in molecular weights and low in complexity it will fail to provide sufficient

separation capacity to be used as the sole indicator of glycation. Even in the case of samples with low complexity such as our samples, proteins with close molecular weight due to the attachment of few glucose molecules will co-migrate, even with the excellent optimisation of the separation (Rabilloud *et al.*, 2009). Additional SDS-PAGE experiment was performed by loading fewer amounts of the glucose treated samples (data not shown) as an attempt to resolve the glycated bands; we were, however, unable to see separate/multiple bands.

Some workers modified the SDS-PAGE technique (named Flu-PAGE) to detect and identify *in vitro* glycated proteins. In their method the glycated proteins were incubated first with a fluorophore-appended boronic acid, then the labelled glycated proteins are subjected to electrophoresis and visualised with visible blue light or UV (Morais *et al.*, 2013). This method would be a useful alternative to conventional SDS-PAGE if we did not intend to use mass spectrometry.

SDS-PAGE is a useful technique to identify cross-linked AGEs because the significant difference in mass between the cross-linked protein and their non-cross-linked counterpart results in a noticeable change in the electrophoretic mobility (Galván-Moroyoqui *et al.*, 2016, Sheng *et al.*, 2018). However, it is not expected to identify a cross-linked structure of CRP as the incubation time was limited to six days in pseudo-physiological condition (Nawale *et al.*, 2006, Anguizola *et al.*, 2013).

Western blotting was performed to assess whether the effect of glycation could be visualised using the immunoreactivity of CRP with its anti-CRP antibody. Alteration in the antigen-antibody binding because of glycation reaction has been reported with several proteins (Ma *et al.*, 2013, Seo *et al.*, 2013). The changes in immunoreactivity of targeted protein could be attributed to blockage of the lysine residue or the structural changes as a

result of protein glycation, leading to either epitope modification or steric hindrance thereby affecting the antigen-antibody recognition (Ma *et al.*, 2013). The western blotting results show that the anti-CRP antibody have comparable reactivity toward all the glycosylated samples regardless of the incubation time and the glucose concentration used (see Figure 2-5), and the protein at the higher glucose concentrations/times show similar bands intensity compared to the non-glycosylated CRP control, this would indicate that CRP-antibody interaction is not affected by glycation.

The anti-CRP antibody used in western blotting was a monoclonal anti-CRP antibody (clone-8) which recognise an epitope located on the denatured and reduced CRP monomer. The manufacturer does not reveal which epitope is involved; however, it is reported that monoclonal anti-CRP antibody (clone-8) cross-react with an epitope on Hsp60. The epitope occupies amino acids sequence from 218-232 on Hsp60 and displays 26.6% amino acid identity to CRP amino acid region 77 to 90 (Udvarnoki *et al.*, 2007). This sequence (77-90) have only one possible glycation site, K87. If this sequence is where the antibody bind to the protein, then either K87 is not glycosylated or its glycation has no effect on the antibody-CRP interaction.

#### 2.4.2 Using LC-MS for the intact analysis of glycosylated proteins

Among the multiple analytical methods available to detect proteins glycation (Priego Capote & Sanchez, 2009, Lacinová *et al.*, 2010), the mass spectrometric-based techniques provide the most direct and reliable approach for detection and characterisation of glycosylated protein. The usefulness of mass spectrometry for both *in vivo* and *in vitro* monitoring of non-enzymatic glycation of protein has been well documented (Bai *et al.*, 2012, Farah *et al.*, 2005, Lapolla, *et al.*, 2013, Thornalley, *et al.*, 2000).

In this work, a Q-TOF-MS was coupled to nESI ion source, the nESI ion source shows a better overall performance for posttranslational modification detection than MALDI-TOF mass spectrometer (ter Haar *et al.*, 2011, Nadler *et al.*, 2017). Moreover, the nESI can easily be combined with HPLC as the coupling is fast, simple and provides an extraordinarily powerful analytical tool (Hart & Gaskell, 2005). The Q-TOF mass spectrometer has been used in MS mode, which allows the measurement of the accurate mass of the intact protein and the relative abundance of its isoforms (Stastna & Van Eyk, 2012).

During glycation, the covalent coupling of a glucose moiety to the protein via the Maillard reaction, accompanied by loss of a water molecule, lead to a mass increase of 162 Da per glucose molecule incorporated to the protein, this characteristic mass shift incurred to protein by the Amadori product ( $\Delta m$  162 Da) can be measured by mass spectrometric analysis of the intact protein. Utilising the LC-MS to determine the mass of intact protein molecules will detect the glycosylated protein and monitor the reaction as a function of the net gain in the molecular weight of the protein due to its binding to glucose. Determination of net increase in the mass of intact protein was the basis of measurement of the degree of glycation of many previous studies dealing with the analysis of protein glycation by MS (Farah *et al.*, 2005, Thornalley *et al.*, 2000, Lapolla *et al.*, 2013, Leblanc *et al.*, 2016, Gadgil *et al.*, 2007).

The acquired multiply charged  $m/z$  spectra of the Q-TOF-MS were deconvoluted using maximum entropy function (MaxEnt1) to give the actual molecular mass spectra. MaxEnt1 preserves the quantitative information of the multiply charged electrospray spectra. The intensities of the peaks in the deconvoluted spectra were used to calculate the extent of glycation, percentage of each glycoform, and the average number of glucose molecules per



molecule of CRP (G-idx). These calculations based on the assumption that limited glycation has no to little effect on the ionisation state of proteins and the intensities of the peaks are representative of their corresponding species (Yeboah & Yaylayan, 2001). Quantifying and understanding the degree of glycation based on peaks intensities has been used in several studies and found to be reasonably accurate (Gadgil *et al.*, 2007, Miller *et al.*, 2011b, ter Haar *et al.*, 2011, Leblanc *et al.*, 2016).

#### 2.4.3 LC-MS analysis of the *in vitro* glycosylated CRP samples

The typical spectrum of a non-glycosylated CRP standard, as acquired by nESI Q-TOF-MS, shows a broad distribution of multiply-charged monomer peaks (see Figure 2-6, A). This is one of the advantages of utilising the ESI-MS instruments, where the protein can acquire more than one charge (usually multiple charges) during the ionisation process, and each different charge state will give rise to a separate peak for the same compound in the spectrum. The set of peaks will reduce errors during the calculation of the mass and hence give higher mass accuracy.

After software-assisted deconvolution of the acquired  $m/z$  spectrum (Figure 2-6, B), the main peak at 23027 Da corresponds to the expected mass of CRP unit was obtained (Kiernan *et al.*, 2004). CRP is a pentameric protein at physiological pH, composed of five identical noncovalently associated subunits (Shrive *et al.*, 1996). However, the finding of a single subunit of CRP suggests that the non-covalent interactions between CRP units have been disrupted, as might be expected when using a typical ESI buffer system. The acid and organic solvent in ESI buffer degrade the quaternary and tertiary structure, unfold the protein and expose all their basic side chains to the solvent. However, the non-covalent interactions between proteins are usually destroyed in their denatured state (Hilton &

Benesch, 2012, Loo, 1997). Furthermore, it was observed that incubating CRP with a glucose-free phosphate buffer (standard CRP sample) for a period up to six days did not result in any apparent modification in the LC-MS spectrum (data not shown).

Investigation of glucose-treated CRP by LC-MS at intact protein level demonstrated for the first time, that human CRP can be glycosylated. This can occur to a different extent depending on both the incubation time and the glucose concentration used. Overall, the LC-MS spectra of the *in vitro* glycosylated samples shows charge state distributions similar to that produced by control samples, but additional peak(s) were associated with each charge state. These new peaks were of varying intensity depending on the incubation time or glucose concentration used. Deconvolution of these spectra indicates that these additional peaks represent new forms of the protein due to its glycosylation.

Incubating CRP with different concentrations of glucose for a period equal to the half-life of the protein *in vivo* reveal additional peaks in the spectrum besides the original peak that represents the non-glycosylated CRP. CRP with 1 to 3 glucose molecules attached was found after 19 hours of incubation with different glucose concentration (see Figure 2-7).

At 30 mM glucose level (Figure 2-7, A), which is the lowest glucose concentration used, the deconvoluted spectrum of CRP shows a low-intensity peak at 23,189 Da beside the major peak at 23027 Da that represents the unmodified CRP. The difference in mass of 162 Da between the new and the non-glycosylated CRP peak clearly indicate that the former peak represents the mono-glycosylated CRP. This finding suggests that, despite its short half-life of 19 hours (Vigushin *et al.*, 1993), CRP could be glycosylated in a patient with a high glucose level as in poorly controlled diabetes. Proteins with a half-life comparable to CRP were also

found to be glycosylated in diabetic patients (Spiller *et al.*, 2017, Sensi *et al.*, 1990, Hammer *et al.*, 1989).

During this incubation period, increasing glucose concentration resulted in a parallel increase in the intensity of the mono-glycosylated form of CRP; this adds further confidence in the identity of this peak. Moreover, the samples incubated with higher levels of glucose revealed additional peaks representing the progressive addition of glucose residues; the peak at 23,351, which seen at 300 mM glucose level and thereafter, has a mass increment of 324 Da over the unmodified peak suggests that the peak is the representative of CRP conjugated with two glucose molecules, while the signal at 23,513 Da, which was observed at 1000 mM glucose, is shifted by 486 Da from the original peak in the spectrum and hence represent a condensation of three glucose molecule (Figure 2-7).

Using the intensities of the peaks to estimate the glycosylation level, clearly demonstrate the increase in the glycosylation of the protein from 3.7% at 30 mM glucose to 48.8% at 1 M glucose level (Figure 2-10, A). The mono-glycosylated form was the dominant form during this incubation period. Because the glycosylation condition used were identical for all the samples including the incubation time, then the increase in the extent of glycosylation can be solely attributed to the increase in glucose concentration. Several *in vitro* glycosylation experiments on other proteins showed the relationship between increased glycosylation level and glucose concentration. (Farah *et al.*, 2005, Lapolla *et al.*, 1993, Goodarzi *et al.*, 2004). Farah *et al.*, incubated insulin preparation with 55, 110, 165, and 220 mM glucose for 24 h at 37 °C, pH 7.4 with subsequent mass analysis of the resultant products. In their work, mono-glycosylated insulin was detected with the lowest glucose level with increasing intensity with further

increasing glucose concentration, while the di-glycated form was noticed only when the insulin exposed to 220 mM glucose.

The glycation process of CRP was not only affected by glucose concentration, but also by the incubation time. If we use the CRP sample incubated with the lowest glucose concentration, 30 mM, as an example to explain the effect of incubation time we found that the extent of glycation of this sample was increased from 3.7% at 19 hour incubation to 8.3% after 3 days, and increased further to nearly 20% with further increase the incubation time to 6 days (Figure 2-10). The same trend was seen with all other samples. It can also be observed that the extent of glycation of CRP exposed 30 mM of glucose (which is lowest glucose concentration used in this work but still higher than that those in diabetic patients) for a period equal to its half-life *in vivo* is relatively low and represented by only a mono-glycated form. This degree of glycation could be too low to influence the biological functionality of CRP mainly because of its short half-life *in vivo* (Vigushin *et al.*, 1993). In a poorly-controlled diabetic, longer half-life proteins such as albumin are glycated by condensation of 4-15 glucose molecule while IgG was found to be glycated with 7-28 glucose (Lapolla *et al.*, 2013), the glycation of both proteins known to affect their functionality (Anguizola *et al.*, 2013, Goetze *et al.*, 2011).

It is observed that samples incubated with glucose for three days favoured the formation of mono-glycated CRP even when the protein exposed to 1000 mM of glucose, while glycoforms with the higher level of glycation tetra-, penta-glycated CRP represented the lowest proportion (see Figure 2-8 and 2-10, B). The mono-glycated CRP is the only detected glycated form in the samples incubated with 30 and 100 mM for 19 hours and with 30 mM

for 3 days, suggesting that if CRP found glycosylated *in vivo* it probably will be a mono-glycosylated form.

By the sixth day of incubation (see Figure 2-9), there was a shift in the distribution of CRP glycoforms toward the highly-glycosylated forms with higher glucose concentration. The predominant glycoform was the mono-glycosylated CRP for the samples incubated with up to 300 mM glucose. At 500 mM glucose, the di- and tri-glycosylated CRP become the most dominant (26.8% and 24.3%, respectively) while the mono-glycosylated CRP decreased to the third lowest observed glycoform (17.3%). At 1000 mM of glucose, the tri- and tetra was the most observed glycoform, representing 23.7% and 23.6% of total glycoforms, respectively.

It was observed that the same level of glycosylation and glycoforms distribution that was achieved by longer incubation times can be achieved with higher glucose concentration in shorter times. For instance, the samples incubated with 30 and 100 mM of glucose for six days produced a mass profile comparable to the samples incubated with 300 and 1000 mM but for 19 hours, respectively. This suggests that the degree of CRP glycosylation achieved after six days of exposure to the glucose can be reached after 19 hours if the glucose concentration increased to tenfold.

The G-idx elucidate the dynamic of the glycosylation process against the time. For constant sugar level, the degree of glycosylation of CRP was found to be augmented by prolonging the time of exposure to that level of glucose. The relationships between G-idx of each sample and the incubation time show a linear ascending trend, suggesting that the further increase in incubation time could result in more glycosylation. At 1 M glucose level almost all the protein is glycosylated (97.6%) with an average of 3.6 glucose molecule per each CRP molecule (Figure

2-11). Increasing the incubation beyond this period will be aimless; however, it could result in a further increase in G-index value as it reflects the heterogeneity of the glycation product.

Utilising LC-MS to analyse glycosylated protein at intact level offers several advantages including minimal sample pre-treatment, fast detection of glycosylated forms, provides means to estimate the relative quantity of the glycoforms from their intensities in the spectrum (Zhang *et al.*, 2008). One drawback of this approach is that it does not result in any information regarding the modification site within the amino acid sequence of the protein.

#### 2.4.4 Conclusion

Investigation of glucose-treated CRP by LC-MS at intact protein level demonstrates, for the first time, that human CRP can be glycosylated to a different extent depending on both concentration and time of exposure to glucose. The LC-MS analysis of intact CRP and the intensities of the peaks of the glycosylated forms can be effectively used in the detection and determination of the extent of CRP glycation. The extent of CRP glycation seems to be low with the formation of only a mono-glycosylated form when the protein incubated with 30 mM of glucose for a period equal to its half-life *in vivo*. Samples incubated with glucose for 19 hours and three days favoured the formation of mono-glycosylated CRP even when the protein exposed to 1000 mM of glucose, while glycoforms with the higher level of glycation were obtained upon the glycation of CRP with higher glucose levels (0.5 and 1 M of glucose) and longer incubation time. Western blotting shows that the glycation of CRP does not affect its reaction with the monoclonal anti-CRP antibody.

## Chapter 3: Revealing the glycation sites of *in vitro* glycated C-reactive protein

### 3.1 Introduction

#### 3.1.1 Overview

In chapter two, we demonstrated the usefulness of the intact mass measurement by electrospray ionisation-mass spectrometry in the identification of the glycated isoforms and estimation of the glycation level of the *in vitro* glycated CRP samples treated with different concentrations of glucose for different time intervals. In this chapter, further studies were carried out involving enzymatic digestion of the glycated CRP samples with trypsin and endoproteinase Glu-C, followed by product ion tandem mass spectrometric analysis of the digests, the aim of these experiments were to identify the precise glycation sites in CRP. The Mascot search engine was used to identify the glycated peptides and the possible glycation site; the product ion spectra of glycated CRP peptides were then further interpreted manually, and the neutral losses from the glucose-containing parent and fragment ions were employed to sequence the peptides and to confirm the Mascot result or identify the modified residues when Mascot fail.

### 3.1.2 Proteomic approaches to identify modification sites by mass spectrometry

Analysis of protein post-translational modifications (PTM) such as glycation, using mass spectrometry can provide information about the exact location of the modified residues (Mann & Jensen, 2003). Intact protein analysis by MS alone does not result in any information regarding the precise modification site within the amino acid sequence of the protein (Zhang *et al.*, 2008). Tandem mass spectrometry (MS/MS) based experiments which involve fragmentation of either the intact, undigested protein (top-down approach) (Zhang & Ge, 2011), or of peptides derived from proteolytic digestion of intact proteins (bottom-up approach) are typically required for the confident assignment of the modification sites (Larsen *et al.*, 2006, Silva *et al.*, 2013).

Top-down and bottom-up approaches are the two major classes of modern MS-based proteomics (Zhang & Ge, 2011, Verrastro *et al.*, 2015, Zhang *et al.*, 2008). The top-down approach remains nascent and requires a high mass resolution/accuracy instrument such as Fourier-transform ion cyclotron resonance (FTICR) MS (Zhang & Ge, 2011). By contrast, the well-established bottom-up approach has been commonly employed in the identification of glycated proteins and their modified residues (Gadgil *et al.*, 2007, Jahouh *et al.*, 2013, Zhang *et al.*, 2008, Wang *et al.*, 2014, Swamy-Mruthinti & Schey, 1997). The top-down approach has never been applied in the analysis of non-enzymatically glycated proteins (Zhang *et al.*, 2008). This introduction will, therefore, focus on bottom-up approaches and the technologies and strategies that have been developed or applied to characterise glycated proteins and determine their modification sites.

The bottom-up approach, which is also named 'shotgun' analysis when performed on a mixture of proteins (Zhang *et al.*, 2013), involves the following steps: protein digestion by



single or multiple enzymes, separation of the generated peptides typically by liquid chromatography (LC) and then peptide analysis using by MS and MS/MS (Zhang *et al.*, 2013, Rathore *et al.*, 2018). Identification of the protein and mapping of the modification sites relies on the resulting peptide  $m/z$  values, and their product ion spectra with the assistance of database search bioinformatic tools (Tsiatsiani & Heck, 2015). Figure 3-1 shows the steps involved in analysing protein through a bottom-up approach.

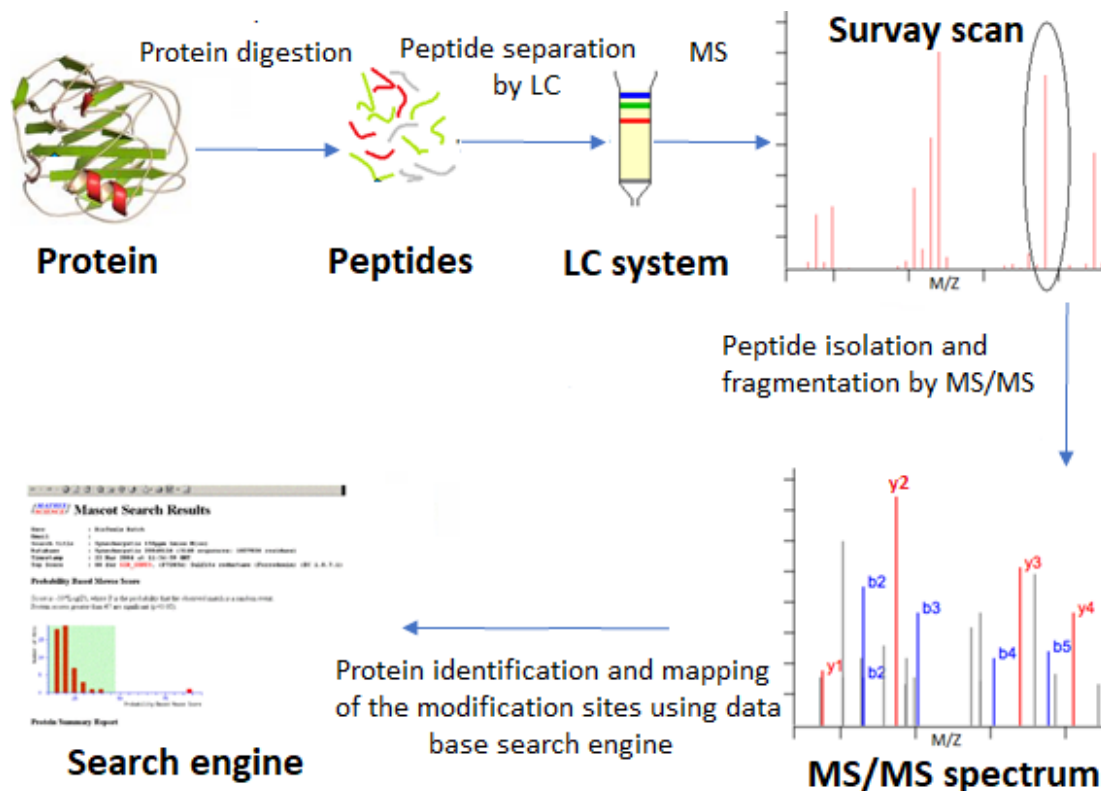


Figure 3-1 Schematic presentation for the general steps in bottom-up proteomics. In bottom-up approach the proteins are first digested into peptides prior to separation and MS analysis, the resulting peptides may be further fragmented before another step of MS analysis, identification of the protein is achieved by the assistance of database search bioinformatic tools.

### 3.1.3 Peptide production via enzyme digestion

The important step during sample preparation for bottom-up experiments is the conversion of proteins into appropriately-sized peptides suitable for mass spectrometric analysis (Hustoft *et al.*, 2012). Different proteolytic enzymes with varying degrees of specificity are currently available for this task (Tsiatsiani & Heck, 2015), and selecting the appropriate enzyme depends largely on experimental goals (Verrastro *et al.*, 2015).

Trypsin is the most commonly protease used in proteomic studies (Tsiatsiani & Heck, 2015, Hustoft *et al.*, 2012). It cleaves specifically and efficiently at the C-terminal side of lysine and arginine residues except if they are followed by a proline (Hustoft *et al.*, 2012). Because of its cleavage specificity and the relative abundance of lysine and arginine; trypsin-generated peptides are relatively short peptides generally with a basic residue at the C-terminus, meaning that tryptic peptides carry at least two positive charges in acidic medium; one on the N-terminal amine and another on the basic (K or R) residue (Manadas *et al.*, 2010). Tryptic peptides are well-separated using conventional chromatographic fractionation, and generally produce information-rich and readily-interpretable tandem mass spectra that help in peptide identification (Tsiatsiani & Heck, 2015).

When the comprehensive mapping of PTMs of proteins is the goal of the study, generating as close to complete protein sequence coverage as possible is required. However, one drawback of trypsin is that in some cases the generated peptides may be too short to be sequenced by MS (Swaney *et al.*, 2010, Tsiatsiani & Heck, 2015), resulting in limited protein sequence coverage. Therefore, it is often impossible to identify all the post-translational modifications decorating a given protein using trypsin alone (Tsiatsiani & Heck, 2015). Moreover, the presence of PTMs is known to affect the cleavage efficiency of proteases

(Larsen *et al.*, 2006). Protein glycation alters the digestion pattern of trypsin (or other enzymes cleaving at K or R residues) and results in the generation of longer peptides as a result of miscleavage, since some K and R side chains are modified in the glycation reaction with carbohydrates (Larsen *et al.*, 2006, Priego-Capote *et al.*, 2014). Nevertheless, several studies concerned with the mapping of glycation sites have utilised trypsin as a sole proteolytic enzyme (Brady *et al.*, 2007, Zhang *et al.*, 2008, Silva *et al.*, 2018, Frolov *et al.*, 2006). Using other proteases that have different cleavage specificities and/or using multi-protease protein digestion alongside trypsin can increase the protein sequence coverage and improve the identification of PTMs (Tsiatsiani & Heck, 2015, Rathore *et al.*, 2018).

Glu-C specifically cleaves peptide bonds C-terminally to glutamate or aspartate (at a lower rate) and Asp-N, cutting N-terminally to asparagine and cysteine, are both suitable alternative to trypsin when trypsin cleavage sites are unavailable. Glu-C and Asp-N are less reactive than trypsin (Hustoft *et al.*, 2012). However, because of the limited cleavage sites for these enzymes, Glu-C and Asp-N generate longer peptides than does trypsin, therefore, allowing increased sequence information per peptide (Tsiatsiani & Heck, 2015). Glu-C and Asp-N have been used alone or in combination with trypsin to digest glycated proteins for identifying the glycation sites by tandem mass spectrometry (Jahouh *et al.*, 2013, Priego-Capote *et al.*, 2014, Meltretter *et al.*, 2008, Sun *et al.*, 2005). Pepsin, a nonspecific protease, was also reported to improve sequence coverage of glycated proteins (Wang *et al.*, 2013).

### 3.1.4 Peptide separation

Following protein digestion, the resulting peptide mixture is typically separated by HPLC before mass spectrometry (Xie *et al.*, 2012). Reversed-phase high-performance liquid chromatography (RP-HPLC) is commonly used for separation of the peptides (Aguilar, 2004, Manadas *et al.*, 2010, Simpson *et al.*, 2006). RP separation is based on the binding of the peptides in aqueous, moderately polar mobile phase to hydrophobic ligands immobilised on the stationary phase (commonly n-octadecyl C18); (Manadas *et al.*, 2010). Polar peptides elute first, while the hydrophobic peptides interact more strongly with the hydrophobic groups in the stationary phase, causing their retention, i.e. the retention time of a given peptide is directly proportional to its hydrophobicity (Aguilar, 2004). Elution of the peptides from hydrophobic stationary phases is generally achieved by means of applying a gradient of increasing organic content (such as acetonitrile or methanol, with an ionic modifier such as trifluoroacetic acid or formic acid); (Aguilar & Hearn, 1996). The nature of the buffer system used in RP-HPLC allows direct coupling to the MS system, making it the preferred separation method in bottom-up proteomics (Tsiatsiani & Heck, 2015). RP-HPLC not only separates peptides based on their hydrophobicity but also affords a means for removal of the polar salts which are deleterious to MS analysis (Zhang & Ge, 2011).

Besides RP-HPLC, other separation techniques have been used online or offline, including ion-exchange chromatography (most commonly strong cation exchange), hydrophilic interaction LC, and IEF. These methods can be used individually or in combination (usually with RP-HPLC) to resolve peptides from more complex proteomes in two-dimensional liquid chromatography (Manadas *et al.*, 2010, Simpson *et al.*, 2006).

### 3.1.5 Peptide mass fingerprinting

Peptide mass fingerprinting (PMF) is a method based on MS analysis of the peptides generated from proteolytic digestion of a known or unknown protein followed by matching of the obtained masses with a list of theoretical peptide masses generated from *in silico* digestion of the protein sequence (Thiede *et al.*, 2005). PMF can be used for protein identification and investigating the presence of PTMs. The matched peptide masses identify the protein sequence, while the unmatched values could arise from several factors including the presence of PTMs (Mann & Jensen, 2003, Holmes & Giddings, 2004). Evidence for the occurrence of PTMs upon a specific peptide can be inferred from the resultant mass shift between the measured value and theoretical mass (Zhang *et al.*, 2003). Examples of some mass shifts ( $\Delta M$ ) resulted from the formation of early and advanced glycation end product are listed in Table 3-1. PMF has been applied in the identification of glycated peptides and their corresponding proteins using LC-MS (Ames, 2005, Brock *et al.*, 2003), or MALDI-TOF MS (Farah *et al.*, 2005, Wa *et al.*, 2007). In some cases, it was even possible to predict the modified residue when there is only one lysine or arginine residue in the given glycated peptide (Wa *et al.*, 2007). Using LC-MS-based peptide mass fingerprinting, Brock *et al.* determined the major sites of glycation of RNase, incubated with glucose in phosphate buffer, are K41, K7, K1, and K37 (Brock *et al.*, 2003).

Table 3-1 Summary of possible early and advanced glycation adducts, according to literature. A; the possible modification that involves lysine residues, B; arginine residues, C; or both lysine and arginine residues due to glycation processes (Vlassara, 1994, Lyons & Jenkins, 1997, Thornalley, P. J., 1999).

	Glycation adducts	Mass (Da)
<b>A</b>	Fructosyl-lysine	162.0528
	Fructosyl-lysine -1H <sub>2</sub> O	144.0423
	Fructosyllsine -2H <sub>2</sub> O	126.0317
	Nε-(carboxymethyl hydroxy)-lysine (CML)	72.0211
	Nε-(carboxymethyl)-lysine (CEL)	58.0055
	Pyrraline	108.0211
<b>B</b>	Nδ-[5-hydro-5-(2,3,4-trihydroxybutyl)-4-imidazol-2-yl]ornithine	144.0423
	Tetrahydropyrimidine (THP)	144.0423
	Imidazolone B	142.0266
	Argpyrimidine	80.0262
	Nδ-(5-hydro-5-methyl-4-imidazol-2-yl)-ornithine (MG-H1)	54.0106
	Nδ-(5-hydro-4-imidazol-2-yl)-ornithine (G-H1)	39.9949
<b>C</b>	1-alkyl-2-formyl-3,4-glycosyl-pyrrole (AFGP)	270.074

While PMF is a fast and simple method, as it does not involve amino acid sequencing of the peptide with tandem mass spectrometry (MS/MS). PMF frequently does not provide adequate information to determine the glycation site, particularly in cases where the peptide has more than one possible modification site. Further, some of the  $\Delta M$  may be poorly diagnostic, preventing confident assignment of glycation from another kind of PTM, for instance, the  $\Delta M$  of 162 could result from nonenzymatic glycation with hexose or enzymatic O- and N-glycosylation with hexose (Mann & Jensen, 2003).

### 3.1.6 Tandem mass spectrometry

Tandem mass spectrometry experiments involve two or more stages of mass analysis, often with analyte dissociation steps between subsequent stages. Generally, during the first stage, precursor ions of a specific  $m/z$  range are isolated, ion activation is performed to promote the production of product ions, the  $m/z$  values of which are then recorded (Glish & Vachet, 2003, El-Aneed *et al.*, 2009).

Several dissociation methods have been applied to promote precursor ion fragmentation (Silva *et al.*, 2013). The most widely available and hence applied fragmentation method is collision-induced dissociation (CID); (Zhang *et al.*, 2008, Silva *et al.*, 2013). CID is based on the collision between selected precursor ions with inert collision gas molecules (such as helium, nitrogen or argon). Dissipation of the resultant kinetic energy takes place via vibrational translation. For unimolecular precursor ions, this ultimately results in fragmentation of covalent bonds; in the case of peptides the most abundant cleavage sites are amide bonds (Zhang *et al.*, 2013, Larsen *et al.*, 2006). CID of peptides, therefore, results in general in the predominant formation of two types of product ions; the product ions derived from the N-terminal side of the peptide are referred as b-ions, and product ions derived from the C-terminus are referred to as y-ions (El-Aneed *et al.*, 2009). The different type of peptide backbone fragment ions according to the nomenclature of Roepstorff and Biemann are shown in Figure 3-2 (Biemann, 1992).

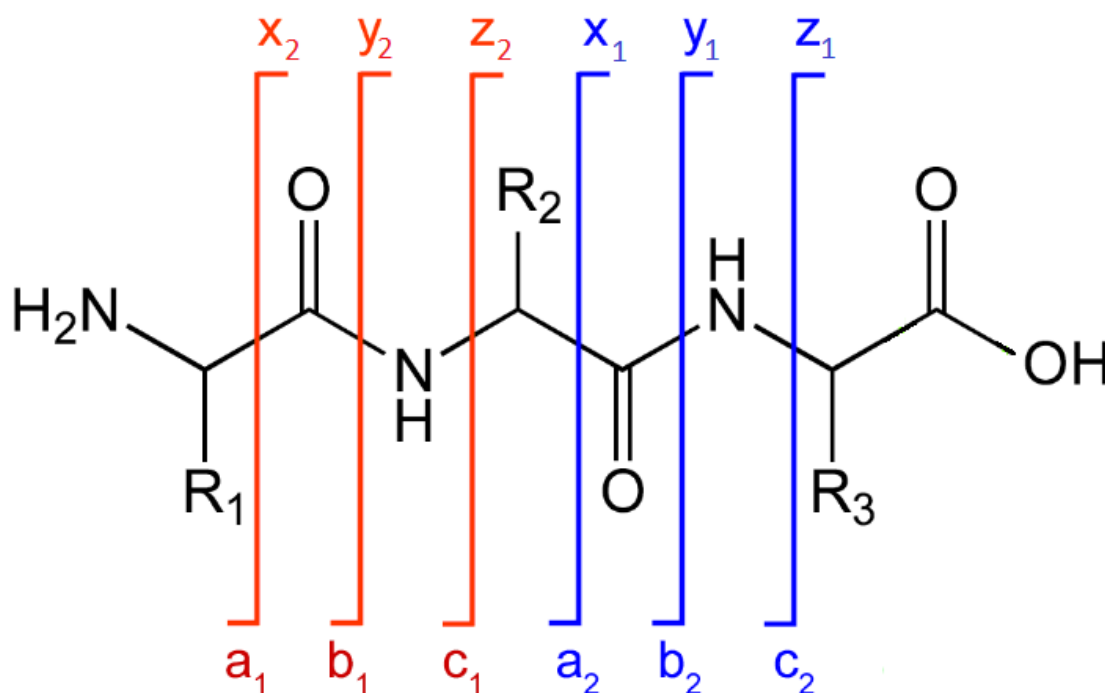


Figure 3-2 Peptide ion fragmentation in tandem mass spectrometry. Six types of backbone fragment ions can be generated, a-, b- and c- are the N-terminal fragment ions while x-, y-, and z-ions are the C-terminal fragment ions. Reproduced from (Biemann, 1992).

For the Q-TOF mass spectrometer, which was used in this study, the peptide precursor ion is selected within the first quadrupole (Q1) mass filter, fragmented by CID within the collision cell (an RF-only quadrupole), and the resulting fragment ions (principally b- and y-ions in addition to the residual precursor) are measured by the TOF analyser (Chernushevich *et al.*, 2001).

Alternative fragmentation techniques for MS/MS-based experiments are available (Quan & Liu, 2013). Higher energy collisional dissociation (HCD) is a subset of CID whereby the precursor ion is accelerated to higher kinetic energy (Quan & Liu, 2013). HCD is principally used with Orbitrap analysers (Olsen *et al.*, 2007). It has the advantage of increased ion fragments, able to detect and generate immonium ions, and has no low mass cut-off (Olsen



*et al.*, 2007, McAlister *et al.*, 2011). Therefore, HCD can provide more informative MS/MS spectra for quadrupole ion trap-based instruments (McAlister *et al.*, 2011). HCD generates b- and y-ions with a high prevalence of y-ions; b-ions can be further fragmented to a-ions or smaller species (Olsen *et al.*, 2007). Other methods to induce fragmentation include: electron-capture dissociation (ECD), used mainly with FTICR (Sleno & Volmer, 2004), and electron-transfer dissociation (ETD), used with ion trap mass spectrometers and Q-TOF instruments (Kim & Pandey, 2012). Both these methods involve electron-driven radical species generation from a multiply protonated peptide, followed by N-C $\alpha$  bond dissociation, resulting in c- and z-type products (Kim & Pandey, 2012). ECD and ETD allow identification of PTMs that are labile under CID. Uptake of these methods has been limited due to dedicated instrumentation requirements (Silva *et al.*, 2013).

Several types of scan mode can be performed by tandem mass spectrometers (Niessen, 1999). The basic and the most commonly used scan mode is the product ion scan, where a specific precursor ion is selected during the first stage of the mass analysis, and the product ions generated from its fragmentation are detected during the second stage of mass analysis (Niessen, 1999). Product ion scans are used to infer peptide sequences and hence identify the proteins in a biological sample (Glish & Vachet, 2003). However, if a group of compounds share a common and specific product ion, then precursor ion scans can be used for their identification. Precursor ion scanning involves specifying a target product ion to be monitored in the second mass analyser while the first mass analyser scans all the precursor ions that decompose to produce the specified product ion (Glish & Vachet, 2003, Niessen, 1999). Neutral loss scanning is another scan mode that can be performed by some instruments, here the instrument is set to transmit only ions where a specific neutral mass loss is observed in product ion populations. The  $m/z$  offset corresponds to the mass of

neutral loss. For example; 18 Da loss to identify precursor ions losing a water molecule (Glish & Vachet, 2003, Niessen, 1999). Precursor ion and neutral loss scans are sensitive and specific techniques for the identification of modified peptides (Salzano & Crescenzi, 2005, Larsen *et al.*, 2006). Finally, selected reaction monitoring scans are typically used for quantification. Both the precursor and product ion  $m/z$  values are specified here, and the detection takes place only when the precursor ion selected by the first mass analyser fragments to produce a given product ion (Glish & Vachet, 2003, Niessen, 1999). Figure 3-3 summarise the different MS/MS scan mode.

Q-TOF instruments can be operated with product ion scan, precursor ion scan (however, with much lower sensitivity than to triple quadrupoles); (Chernushevich *et al.*, 2001), and selected reaction monitoring-like experiments by utilising software to algorithmically extract the targeted signals from full or partial product ion scans (Bereman *et al.*, 2012).

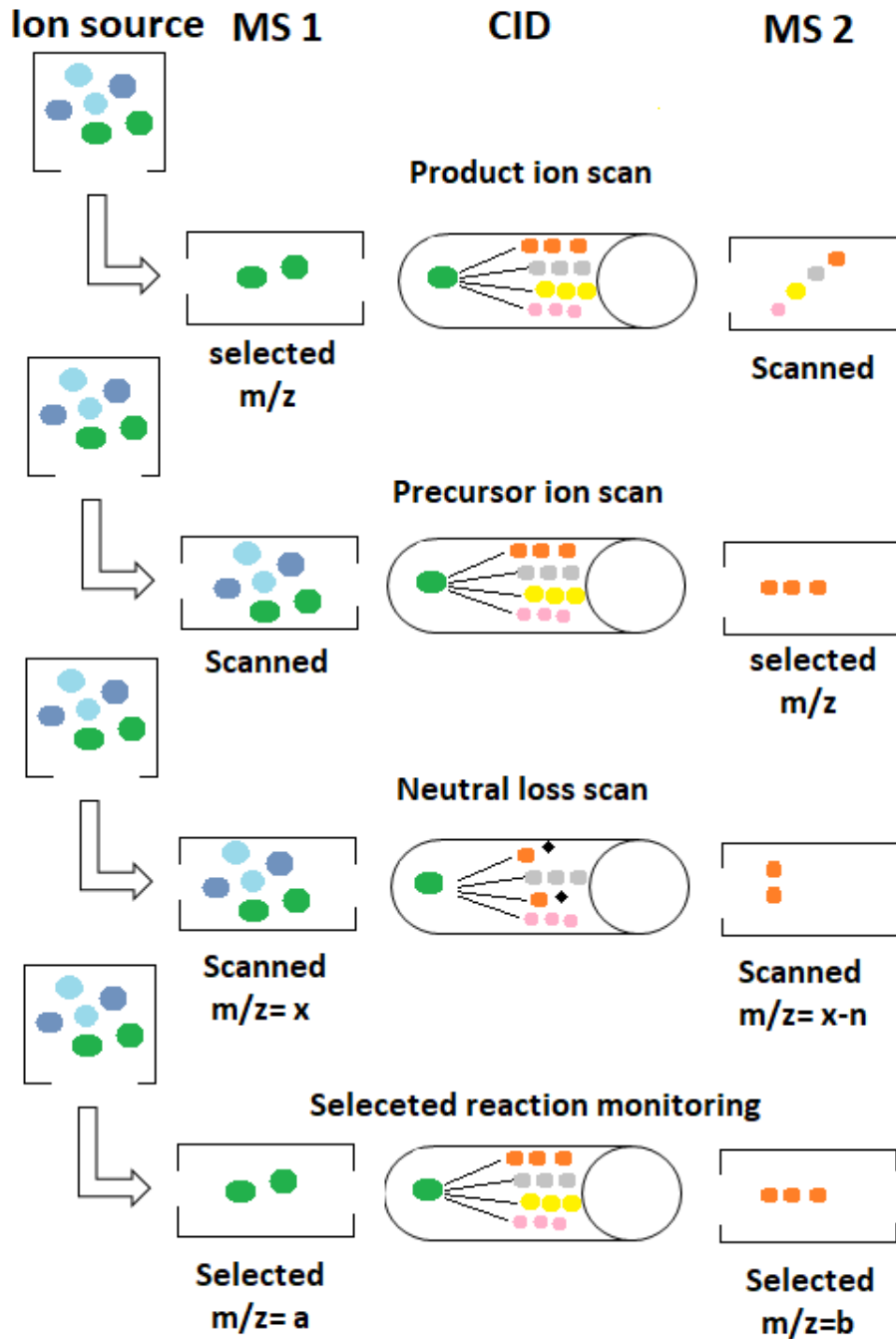


Figure 3-3 Tandem mass spectrometry scan types. Product ion scan: MS1 select precursor ion and MS2 scan production. Precursor ion scan: MS1 scan precursor ion and MS2 monitor specific product ion. Neutral loss scan: MS1 scan precursor ion and select ions with a neutral loss in MS2. Selected reaction monitoring: MS1 select precursor ion and monitor specific product ions in MS2. MS and CID stand for mass spectrometer and collision-induced dissociation, respectively

### 3.1.7 Identification of peptides

Once acquired, the MS/MS dataset (containing product ion spectra and precursor ion  $m/z$  values) have to be analysed. The peptide sequence can be determined from the fragmentation spectrum by calculating the mass shift between adjacent fragment peaks and determining any match between mass increments and amino acid masses (Aebersold & Goodlett, 2001). If a specific PTM modifies the analysed peptide ion, the type of modification and the location of the modified residue may be similarly determined from the mass increment (Mann & Jensen, 2003). However, this is only possible if the PTM remains intact during the fragmentation process with adequate fragment ions covering the peptide sequence (Mann & Jensen, 2003). A number of known PTMs show high lability under CID conditions, including phosphorylation and glycosylation (Kim & Pandey, 2012).

The large, complex datasets generated during tandem mass spectrometry experiment are highly amenable to bioinformatics analysis, as peptide fragmentation processes are predictable; the presence of modifications (post-translational, chemicals or accidental) further increases the complexity of sequence determination. Various algorithms have been developed for the automatic interpretation of tandem mass spectra. These algorithms fall into two types (Nesvizhskii & Aebersold, 2004): *de novo* sequencing algorithms; these identify peptide sequences directly from the tandem MS data based on amino acid masses e.g. PEAKS algorithm (Zhang *et al.*, 2012), and database searching algorithms, which compare MS/MS spectra in the experimental dataset against theoretical fragmentation spectra of peptides generated from *in silico* digestions of protein sequence database entries (Helsens *et al.*, 2007). The output of such comparisons is the identity of the peptides with a score or expectation value to quantify the degree of similarity between the compared spectra (Nesvizhskii & Aebersold, 2004). Examples of these algorithms are

Mascot (Perkins *et al.*, 1999), Sequest (Eng *et al.*, 1994), and X!Tandem (Craig & Beavis, 2004).

Mascot, which is developed by Matrix Science, is a probability-based scoring system which supports PMF, sequence query, and MS/MS ions search (Perkins *et al.*, 1999). Mascot has been extended to include a modified peptides search, where PTMs such as glycation can be included as variable modifications during the search (Creasy & Cottrell, 2002). It should be noted that identifications of modified peptides are often less precise than unmodified peptides, mainly when more than one modification is present, because of the increase in the search space involved in variable searching with modifications (Mann & Jensen, 2003). MS/MS spectra from peptides with labile modifications are often less informative due to neutral loss being the predominant fragmentation. Nevertheless, these difficulties can be overcome by validation of the tandem mass spectrum against possible modified peptides predicted from the putative sequence (Mann & Jensen, 2003).

### 3.1.8 Fragmentation behaviour of glycated peptides under CID condition

Glycated peptides have characteristic fragmentation patterns distinct from those of normal non-modified peptides when using CID or HCD (Rabbani *et al.*, 2016, Frolov *et al.*, 2006). When the peptide is modified by glycation, the intramolecular vibrational energy resulted from the collision redistributes to include the peptide side chain modification, causing dissociation of weakest bonds on side chain modifications before the amide bond cleavage (Zhang *et al.*, 2008). The labile Amadori product tends to lose up to three molecules of water to generate pyrylium ions and loses three water molecules with additional loss of formaldehyde to form furylium and immonium ions (Jerić *et al.*, 2002). Consecutive loss of four water molecules has also been reported, attributed to the five-membered furanose

ring form, which can lose up to four water molecules (Frolov *et al.*, 2006). The fragmentation of glycosylated peptides under CID or HCD conditions and the proposed structures of the fragment ions mentioned above are illustrated in figure 3-4. The signals of the successive neutral losses dominate the MS/MS spectra of glycosylated peptides; glycosylated b- and y-ions signals are usually very poor or even undetectable (Frolov *et al.*, 2006). The characteristic product ions can be used to demonstrate the presence of a glycosylated peptide (Wang *et al.*, 2014, Leblanc *et al.*, 2016, Frolov *et al.*, 2006).

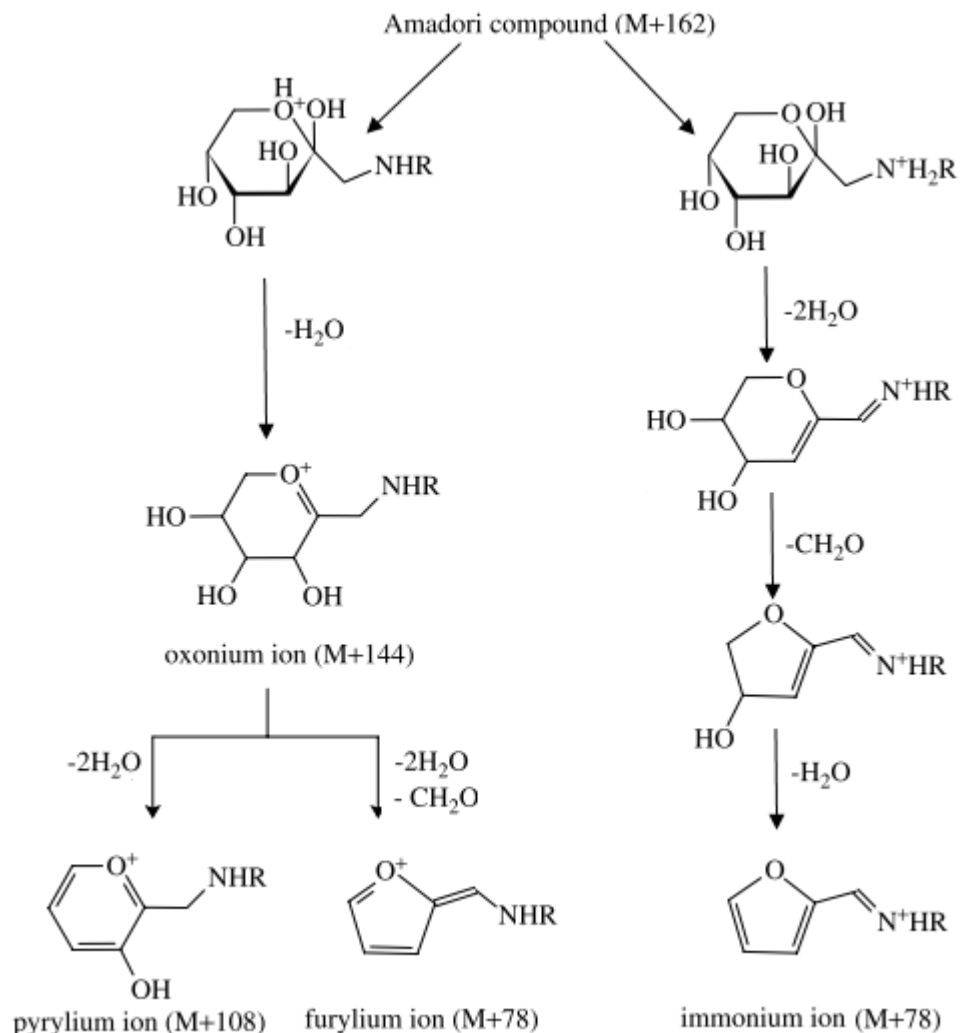


Figure 3-4 Fragmentation of Amadori compounds under CID conditions. Fragmentation of Amadori compounds by CID results in the formation of oxonium, pyrylium, furylium and immonium ions. Reproduced from Rabbani *et al.*, 2016.

### 3.1.9 Other technologies and strategies to characterise glycated peptides

Glycated peptides have been demonstrated to generate a number of low mass ions related to modified lysine immonium ion species including: glucose-derived lysine immonium ion, pyrylium-derived lysine immonium ion and furylium-derived lysine immonium ion (246.2, 192.1, and 162.1 Da respectively when considering the mass of 84.1 Da for the unmodified lysine immonium ion signal); (Frolov *et al.*, 2006, Priego-Capote *et al.*, 2014). The presence of such ions in product ion spectra is a good indicator that a given peptide is modified by glycation. However, among the modified lysine immonium ions, the pyrylium-derived lysine immonium ion generates the most intense signal (Frolov *et al.*, 2006). Precursor ion scanning experiments for pyrylium-derived lysine immonium ion at  $m/z$  192.1 have been used to identify the glycation peptides of the glycated bovine serum albumin using Q-TOF MS; the peptide sequence was determined based on the abundant signals from the pyrylium and furylium glycated fragments ions (Frolov *et al.*, 2006).

Under CID conditions, neutral loss of an entire glucose molecule (-162) is also observed in addition to the neutral loss of multiple water molecules in the product ion spectra of glycated peptides. Neutral loss-triggered product ion scans based on characteristic neutral loss of -162 Da have been used to determine the glycation sites of human serum albumin using Q-TOF MS. The neutral loss scan experiment was composed of two segments; in the first segment the glycated peptides are identified based on their neutral loss of 162 Da, and in the second segment, the identified glycated peptides were fragmented at higher collision energy to obtain the amino acid sequences and the sites of glycation (Gadgil *et al.*, 2007).

Data-dependent neutral loss triggered MS3 (NLMS3), and multi-stage activation tandem mass analysis (MSA) are similar advanced CID techniques that can also be applied to

identify glycation sites (Zhang *et al.*, 2008). In NLMS3 experiments, product ions produced on CID via neutral loss of 3 H<sub>2</sub>O or 3 H<sub>2</sub>O + HCHO are selected for MS3 analysis. In MSA, the neutral loss products are subjected to MS3 without isolation step, and in the presence of the other fragments of the precursor ion (Zhang *et al.*, 2008).

The neutral loss phenomena associated with glycated peptides under CID condition can be avoided by using alternative fragmentation methods such as ETD (Zhang *et al.*, 2007). ETD provides more extensive sequence fragmentation and retains the Amadori compound intact (Zhang *et al.*, 2007). ETD has been applied successfully for sequencing of glycated peptides in several studies, where it outperformed CID by producing abundant and almost complete series of c- and z-type fragment ions enabling direct and confident identification of modification sites (Zhang *et al.*, 2007, Zhang *et al.*, 2008). However, ETD based MS/MS also has limitations; it is less effective for ions with *m/z* over 850 and doubly charged ions (Good *et al.*, 2007), and ETD capability is less widely available (Zhang *et al.*, 2008).

#### 3.1.10 Experimental Aims

The present study aimed to identify the potential glycation sites of CRP, exposed to 1 M glucose concentration over a time course of six days, by enzymatic digestion and tandem mass analysis with the Q-TOF-MS. The second aim was to identify the favourable glycation site of the protein when exposed to the lower glucose levels. A computational-based method to calculate the distance from the potential glycation sites to the nearby acidic and basic residues in the 3D structure of the protein, as well as the exposed surface area of the potential glycation sites and their pKa values, were used to understand the microenvironment of the potential glycation sites. The information gained from this study were utilised to investigate the effect of diabetic glucose levels on CRP *in vivo*.



## 3.2 Materials & Methods

### 3.2.1 Glycated CRP samples used in this study

From the *in vitro* glycated samples prepared in section 2.2.2, samples incubated for 19 h, 3 days, and 6 days with 1 M glucose solution were selected for the tandem mass spectrometry analysis. The rationale for selecting the samples incubated with the highest glucose concentration was to maximise the observation of glycation sites.

Additional samples were also prepared to determine the primary amino acid site modified by glycation. This group included CRP samples incubated with decreasing concentration of glucose; 200, 100, 50 and 30 mM, for 19 hours. The incubation time of the CRP sample exposed to 30 mM was also extended to 3 days. The sample preparation was according to the method described in section 2.2.2. In addition to the *in vitro* glycated samples, a CRP sample prepared under the same condition but without the addition of glucose was used as a control. The final concentration of all CRP samples was 1  $\mu\text{g}/\mu\text{L}$  (43.4 pmol/ $\mu\text{L}$ ) in 50 mM ammonium bicarbonate, pH 7.4. The samples were stored at 4 °C until analysis.

### 3.2.2 Sample pre-treatment and protein digestion

The *in vitro* glycated CRP samples and the non-glycated CRP control (see section 3.2.1) were all treated and digested using the same reagents, protocol, and the same equipment to minimise variations in digestion process due to alterations in the experimental conditions. HPLC grade water was used in the preparation of ammonium bicarbonate buffer (50 mM, pH 8), dithiothreitol (100 mM) and iodoacetamide (100 mM).

CRP samples were first denatured by addition of, MS-compatible detergent, Rapigest (Waters) to enhance the enzymatic digestion of the protein. Rapigest (lyophilised powder, 1 mg) was reconstituted in 100  $\mu\text{L}$  of 50 mM ammonium bicarbonate buffer to give a stock

solution of 10 g/L; this solution is stable for up to 1 week at 4°C. To a 50 µL aliquot of each sample (containing 50 µg of the protein), Rapigest was added to a final concentration of 2 g/L then incubated at room temperature for 30 minutes.

Protein disulfide bonds were reduced by the addition of dithiothreitol (DTT, Sigma-Aldrich) to a final concentration of 10 mM. The solutions were heated to 99.5 °C for 5 min, then kept at 60 °C for 30 min. Upon cooling and gentle vortexing, iodoacetamide (IAA, Sigma-Aldrich) was added to a final concentration of 15 mM to prevent disulfide bonds from reforming. The samples were then incubated in the dark for 1 hour. The alkylation reaction with IAA was stopped by further addition of DTT to reach 25 mM. A sufficient quantity of 50 mM ammonium bicarbonate was added to reduce the Rapigest concentration to 1 g/L in a 125 µL reaction volume. Finally, each CRP solution was divided into two equal portions (each contains ≈25 µg of the protein), the first portion to be used for trypsin digestion, while the second portion to be digested with endoproteinase Glu-C.

Both trypsin (proteomics Grade, Sigma-Aldrich) and Glu-C (protease V8 Sequencing Grade, Roche) were freshly reconstituted. For trypsin, the lyophilised product (20 µg) was dissolved in 20 µL of 1 mM HCl, while Glu-C (50 µg) was dissolved in 50 µL HPLC grade water. Trypsin and Glu-C stock solutions can be stored at 2–8 °C for 2 weeks and 1-2 days, respectively. Aliquots of the stock enzyme solutions were diluted with 50 mM ammonium bicarbonate to 0.1 µg/µL. The working solution of trypsin (0.1 µg/µL) was then added to the first portion of each CRP sample to give substrate to enzyme ratios of 30 to 1, an equal volume of Glu-C working solution (8.3 µL of 0.1 µg/µL) was added but to the second portion to give the same substrate to enzyme ratio. The solutions were incubated for 24 hours at 37 °C with gentle agitation. A second application of each enzyme was added to the

corresponding aliquot after 8 hours of incubation to enhance protein digestion. The digestion was stopped by lowering the pH to less than two by adding TFA to a final concentration of 0.5%. The acidified solutions were then incubated at 37 °C for 30-45 min before centrifugation and removal of the supernatant for analysis. The digestion products (80 µL samples) were stored frozen (-20 °C) until analysis.

### 3.2.3 LC-ESI-MS/MS Analysis of Protein Digests

Before MS/MS analysis, the CRP peptide digests were desalted using C18 Zip-tips (Millipore). The zip-tip was wetted ten times with a 10 µL portions of 100% ACN, and then equilibrated ten times with a 10 µL portions 0.1% F.A solution. A 10 µL aliquot of CRP digest (containing ≈ 3 µg of the protein) was adsorbed onto the reversed-phase medium by 10 repeated aspirate-dispense cycles of sample loading. The tip was washed immediately with 100 µL of 0.1% F.A. Finally, A 10 µL portion of 80% ACN in 0.1% F.A solution was placed into a microcentrifuge tube, and peptide elution was performed by passing this solution through the Zip-tips with three aspirate-dispense cycles. The eluted peptides (≈10 pmole/µL assuming 80% recovery from the Zip-tip) were diluted with HPLC grade water to a final concentration of 1 pmol/µL. Samples were centrifuged for 3-5 minutes to remove precipitate, then transferred to autosampler vial (Waters Corporation). The vials were loaded onto autosampler, and an aliquot of solvent A was also loaded to be used as a blank. The sample run was performed after three blank runs to ensure that the LC-MS system is stable before samples analysis. Moreover, two blank runs were added between samples to minimise observation of system contamination and carryover. The same LC-MS equipment described in 2.2.5 was used for peptide analysis but employing C18 trapping column (Acclaim PepMap100, ID 0.3 mm, length 5 mm, particle size 5 µm, pore size 100 Å, Dionex) and C18 analytical column (Acclaim PepMap100, ID 75 µm, length 15 cm, particle size 3 µm,

pore size 100 Å, Dionex). The HPLC solvents consisted of 0.1% F.A, 5% ACN, as solvent A, and 0.1% F.A, 95% ACN for solvent B.

For each HPLC separation, digested protein (2 pmol) were loaded onto C18 trap column throughout 5 minutes using 95% solvent A, 5% solvent B at 30 µL/min to desalt and concentrate the peptides. After the trapping period, the peptides were separated on a C18 analytical column at a flow rate of 0.35 µL/min. The HPLC gradients consisted of 2 ramping steps; a 35 minutes ramp from 5% to 60% solvent B, followed by a 10 minutes ramp to 95% solvent B then preserving on 95% B for additional 5 minutes before returning to the initial starting conditions at the end of the run.

After separation, peptides eluting from the HPLC column were delivered to a Q-TOF mass spectrometer and introduced into the gas phase through the nanoESI source by applying a DC potential of 1-2 kV to the sample solution within the nano-spray emitter capillary. Survey spectra were acquired in positive ion mode over the  $m/z$  400–1600, with data-dependent product ion analysis of multiply charged precursors where intensity >50 counts per scan,  $m/z$  50 to 1600. The collision offsets used during the CID-MS/MS analyses were preset according to precursor  $m/z$ . Data were collected using MassLynx version 4.1 (Waters Ltd., Manchester, UK).

#### 3.2.4 Collection and analysis of mass spectrometry data

Raw peptide mass values from each whole mass spectrometry file were extracted as text formatted mgf files using Mascot Distiller (Matrix Science, London). The mgf files were subjected to database searching using Mascot database search engine (Matrix Science, London, UK). A custom database, comprising the CRP sequence and common laboratory-observed contaminants was used for these searches with the following parameters:

enzyme: trypsin or V8-E, missed cleavages:  $\leq$ two, fixed modifications: carbamidomethyl (C), variable modifications: oxidation (M), Gln->pyro-Glu (N-term Q), glycation (K, R and N), mass values: monoisotopic peptide, mass tolerance=  $\pm$ 100 ppm, MS/MS tolerance=  $\pm$ 0.6Da.

MS/MS spectra of the precursors whose  $m/z$  values were assigned to glycated CRP peptides through Mascot were further analysed manually to interpret unassigned peaks within the spectrum and identify the glycation site. The  $m/z$  values that did not correspond to any peptide sequence in the selected database during the Mascot search were considered as possible CRP glycated peptides and further analysed as follows: the experimental masses of these  $m/z$  values were compared with a table model containing hypothetical molecular weight values of all possible glycated CRP peptides bearing up to three glucose molecules. The data table was generated by performing a theoretical digestion of CRP using ProteinProspector tool MS-Digest (<http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest>) with the following parameters: enzyme: trypsin or V8-E, missed cleavages:  $\leq$ two, fixed modifications: carbamidomethyl (C), variable modifications: oxidation (M), Gln->pyro-Glu (N-term Q), peptide mass values: 500-5000 Da, minimum peptide length = 5. Subsequently, to the generated CRP peptide masses the mass increment due to glycation (by up to 3 glucose molecules) was added. When there was a match, within 100 ppm mass threshold, then the product ion spectrum of the possible glycated peptide was manually processed.

To facilitate the interpretation of the fragment ions in the spectrum of the matched  $m/z$  value (the possible glycated peptide), an *in silico* fragmentation of the possible glycated peptide were performed using the ProteinProspector tool MS-Product

(<http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct>). Then the values are compared, and only the matches within 0.6 Da threshold were considered.

### 3.2.5 Physicochemical properties of lysine and arginine residues in CRP

Factors with the potential to impact the glycation process of individual lysine and arginine residues of CRP were analysed, these factors include the local pKa values of the lysine and arginine residues, the accessibility of the side chain amino group in these residues, and the nearby acidic (Glu and Asp) and basic (Lys, Arg and His) residues.

The crystal structure of CRP in Protein Data Bank, PDB ID: 1B09, was used in this analysis. Local pKa values were estimated by using a protein pKa predictor (PROPKA 3.0), available on [http://nbc-222.ucsd.edu/pdb2pqr\\_2.0.0/](http://nbc-222.ucsd.edu/pdb2pqr_2.0.0/) (Li, Hui *et al.*, 2005). The fractional solvent accessible surface areas (FAS) of the side chain of lysine and arginine were calculated by using VADAR (Volume, Area, Dihedral Angle Reporter) algorithm, <http://vadar.wishartlab.com> (Willard *et al.*, 2003). The location of lysine and arginine residues within the protein structure and the distance to the nearby acidic and basic residues were computed using PyMOL (<http://www.pymol.org>). The distance from lysine or arginine residue to acidic residues (Glu, and Asp) was calculated from the  $\epsilon$ -NH<sup>+</sup> group of lysine or the guanidino group of arginine to the free acid group of the side chain of the acid residues ( $\gamma$ -COOH group of glutamate or  $\beta$ -COOH of aspartate). Similarly, the distance of lysine or arginine to basic residues was calculated from the  $\epsilon$ -NH<sup>+</sup> or guanidino group to the basic groups of the side chain of the basic residues ( $\epsilon$ -NH<sup>+</sup> of lysine, guanidino group of arginine, or the imidazole group of histidine). Only residues within 10 Å distance in tertiary structure or within three residues distance in the primary structure was included in this calculation.

### 3.3 Results

The non-glycated and the *in vitro* glycated CRP samples prepared according to section 3.2.1 were enzymatically digested by trypsin and Glu-C separately, and the resulting digests were analysed by LC-nESI-QqTOF-MS/MS. The obtained data were submitted to Mascot, which allowed identification of CRP peptides. The product ion spectra of putative glycated peptides were further inspected manually to narrow down the likely glycation site(s).

#### 3.3.1 The sequence coverage of the non-glycated CRP control

The sequence coverage obtained from the tryptic and Glu-C digests of the untreated non-glycated CRP are summarised in Figure 3-5. Sequence coverage of 53% was achieved with trypsin, and 91% from Glu-C digest, combining the mapping results of both enzymes generated 100% sequence coverage thus covering all the probable glycation sites.

```

1      QTDMSRKAFV FPKESDTSYV SLKAPLTKPL KAFTVCLHFY TELSSTRGYS
51     IFSYATKRQD NEILIFWSKD IGYSFTVGGG EILFEVPEVT VAPVHICTSW
101    ESASGIVEFW VDGKPRVRKS LKKGYTVGAE ASIILGQEQD SFGGNFEFSQ
151    SLVGDIGNVN MWDFVLSPDE INTIYLGPF SPNVLNWRAL KYEVQGEVFT
201    KPQLWP

```

Figure 3-5, Sequence coverage of the control CRP sample obtained from the tryptic and Glu-C digest. Residues in red represent coverage obtained from Glu-C digests, while the bold, underlined residues represent coverage obtained from trypsin digestion.

## 3.3.2 Tryptic digest of CRP preparations pre-exposed to 1 M glucose for 19 h, 3 and 6 days

The data acquired from LC-MS/MS analysis of the tryptic digest of CRP samples exposed to 1 M glucose for 18 hours, 3 days and 6 days were submitted to Mascot. The matched CRP peptides are summarised in Table 3-2.

Table 3-2 Peptides identified in the tryptic digest of CRP samples incubated with 1 M glucose in the time course of the experiment. The duration of incubation were A: 19 hours, B: 3 days and C: 6 days. Incubated samples were analysed by LC-nESI-QqTOF-MS/MS, and the acquired data were searched against Mascot using database comprising CRP sequence and common laboratory-observed contaminants. ✓ indicates the observation of a peptide in a given sample. Glycated peptides are underlined, and \* represents the number of attached glucose moieties. All the identified peptides were within a mass tolerance of ±100 ppm.

Sequence	Residues	Experimental mass (Da)	Charge state	A	B	C
KAFVFPK	7-13	835.49	2		✓	
ESDTSYVSLK	14-23	1127.52	2		✓	
ESDTSYVSLKAPLTKPLK	14-31	1976.03	3	✓		
GYSIFSYATK	48-57	1135.53	2	✓	✓	✓
GYSIFSYATKR	48-58	1291.63	2, 3	✓	✓	✓
QDNEILIFWSK <small>Gln-&gt;pyro-Glu (N-term Q)</small>	59-69	1374.67	2	✓	✓	✓
QDNEILIFWSK	59-69	1391.68	2	✓	✓	✓
RQDNEILIFWSK	58-69	1547.79	2, 3	✓	✓	✓
YEVQGEVFTKPQLWP	192-206	1819.89	2	✓	✓	✓
ALKYEVQGEVFTKPQLWP	189-206	2132.1	2, 3	✓	✓	✓
DIGYSFTVGGSEILFEVPEVTVAPVH ICTSWESASGIVEFWVDGKPR	70-116	5166.44	4	✓	✓	✓
<u>KAFVFPKESDTSYVSLK</u> *	7-23	2107.05	3, 4	✓	✓	✓
<u>AFVFPKESDTSYVSLK</u> *	8-23	1978.95	2, 3	✓	✓	✓
<u>GYSIFSYATKR</u> *	48-58	1453.70	2	✓	✓	✓
<u>ALKYEVQGEVFTKPQLWP</u> *	189-206	2294.15	3		✓	✓
<u>YEVQGEVFTKPQLWP</u> *	192-206	1981.96	2			✓



The sequence coverage obtained for samples incubated for 19 h, 3 and 6 days were found to be 54.4%, 54.4% and 50.5%, respectively. This sequence coverage included 10 out of 13 lysine residues (K119, K122 and K123 were not covered) for the sample incubated for 19 h and 3 days; and 8 out of 13 lysine residues (K28, K31, K119, K121 and K122 were not covered) for the sample incubated for 6 days. 4 out of 6 arginine residues (R6, R47, R118, R188) were not covered by the sequence coverage of all the samples digested by trypsin. However, three of these arginines (R6, R47, and R188) were observed cleavage sites for trypsin.

Peptides 7-23, 8-23 and 48-58 were found to be glycated with a single glucose moiety, these glycated peptides were identified in the tryptic digest of CRP samples exposed to 1 M glucose for 19 h, 3 and 6 days (see table 3-2). If the trypsin cleavage site is discounted as a potential glycation site, then K13 is the only potential glycation site in the glycated peptides 7-23 and 8-23 while K57 is the potential glycation site of the peptide 48-58.

The peptide comprising residues 189-206 includes two possible glycation sites, K191 and K201; this was modified with single glucose in the samples incubated for 3 and 6 days. The peptide comprising residues 192-206 has a single potential glycation site. K201 was modified by one glucose after all incubation time of 6 days.

## 3.3.3 Glu-C digest of CRP preparations pre-exposed to 1 M glucose for 19 h, 3 and 6 days

The samples exposed to 1 M glucose for 18 hours, 3 days and 6 days were also digested by Glu-C endoproteinase, the resulting digests were analysed by LC-MS/MS, and the acquired data were submitted to Mascot. The results are summarised in table 3-3.

Table 3-3 Peptides identified in the Glu-C digest of CRP samples incubated with 1M glucose in the time course of the experiment. The incubation times were A: 19 hours, B: 3 days and C: 6 days. Samples were analysed by LC-nESI-QqTOF-MS/MS, and the acquired data were searched against Mascot database using comprising CRP sequence and common laboratory-observed contaminants. ✓ indicates the observation of a peptide in the samples. The glycated peptides are underlined, and \* represents the number of attached glucose. All the identified peptides were within a mass tolerance of  $\pm 100$  ppm.

Sequence	Residues	charge state	Experimental mass (Da)	A	B	C
QTDMSRKAFVFPKE Gln->pyro-Glu (N-term Q)	1-14	3	1665.8	✓	✓	✓
QTDMSRKAFVFPKESDTSYVSLKAPLT KPLKAFTVCLHFYTE Gln->pyro-Glu (N-term Q)	1-42	6	4863.38	✓		
LSSTRGYSIFSATKRQDNE	43-62	3, 4	2322.05	✓	✓	✓
ILIFWSKDIGYSFTVGGSE	63-81	2	2118.05	✓	✓	✓
VTVAPVHICTSWE	89-101	2	1497.73	✓	✓	✓
FWVDGKPRVRKSLKKGTVGAE	109-130	3, 4	2520.36	✓	✓	✓
ASIILGQEQDSFGGNFE	131-147	2	1810.83	✓	✓	✓
GSQSLVGDIGNVNMWDFVLSPDE	148-170	2	2478.12	✓	✓	✓
GSQSLVGDIGNVNMWDFVLS DEINTIYLGGPFSPNVLNWRALKYE	148-193	4	5124.48	✓	✓	✓
INTIYLGGPFSPNVLNWRALKYE	171-193	2, 3	2664.34	✓	✓	✓
VFTKPQLWP	198-206	2	1114.61	✓	✓	✓
<u>ILIFWSKDIGYSFTVGGSE</u> *	63-81	2	2280.11	✓	✓	✓
<u>FWVDGKPRVRKSLKKGTVGAE</u> *	109-130	3, 4, 5	2682.40	✓	✓	✓
<u>LSSTRGYSIFSATKRQDNE</u> *	43-62	3	2484.12			✓
<u>INTIYLGGPFSPNVLNWRALKYE</u> *	171-193	3	2826.42			✓
<u>FWVDGKPRVRKSLKKGTVGAE</u> **	109-130	5	2844.45			✓

Sequence coverage of 91.3% was obtained for the sample incubated for 19 h; this coverage included all the lysine and arginine residues in the primary sequence of CRP. Lower coverage (77.8%) was obtained from samples incubated for 3 and 6 days, all lysine and arginine residues except K23, K28 and K31, were included in this coverage.

When the mapping results of both trypsin and Glu-C are considered, the total coverage increases to 100%, 94.7%, and 90.8% for the samples incubated for 19 h, 3 and 6 days respectively, this coverage included all the probable glycation sites, except the sample incubated for 6 days where K28 and K31 were not covered.

Glu-C digestion products from samples incubated with 1 M glucose for 19 hours shows two peptides modified by one glucose molecule; residues 63-81 and residues 109-130. The peptide 63-81 have only a single lysine (K69), while the peptide 109-130 has 4 lysine and 2 arginine residues. The same modified peptides were also detected after three days of incubation. However, after six days of exposure to 1 M glucose additional modified peptides were detected; peptides 43-62 and 171-193 both were modified by one glucose while the peptide 109-130 was seen modified by two glucose molecules. Residues 43-62 cover K57 and R58, while residues 171-193 cover R188 and K191.

### 3.3.4 Interpretation of MS/MS spectra of the modified peptides

Where multiple potential glycation sites were present within a single peptide, unambiguous assignment of the modified amino acid residue was inconsistently reported by Mascot. For this reason, the CID product ion of each glycated peptide was reviewed manually, looking for the unassigned peaks due to neutral losses, as well as identifying the glycation sites in each modified peptide.

#### 3.3.4.1 Glycation sites in CRP after 19 hours exposure to the 1 M glucose solution

Based on the trypsin digest (table 3-1) and endoproteinase Glu-C digest (table 3-2), three tryptic and two GluC peptides were found modified by one glucose molecule after 19 h of incubation.

Figure 3-6 displays the product ion spectrum of the glycated peptide KAFVFPKESDTSYVSLK at  $m/z$  703.347<sup>3+</sup>. Several intense ions detected in this spectrum arise from neutral losses of water molecules from the fructosamine-containing product ions, mainly loss of 3 molecules of water (-3 H<sub>2</sub>O), 4 molecules of water (-4 H<sub>2</sub>O) and loss of 84 Da (-3 H<sub>2</sub>O+HCHO), in addition to the loss of entire glucose molecule (-C<sub>6</sub>H<sub>10</sub>O<sub>5</sub>). The presence of these neutral loss signals from the product ions  $y_{12}$  to  $y_{16}$ , and  $b_{15}$ ,  $b_{10}$ , and  $b_8$  confirmed not only the glycation status of the peptide but also allowed positive identification of the site of modification as K13. A similar finding was observed for the glycated peptide AFVFPKESDTSYVSLK at  $m/z$  660.65<sup>+3</sup> (Figure 3-7), which has the same sequence as the former peptide but without the N-terminus K7. The signals of fructosamine-containing product ions (such as  $y_{14}$ ,  $y_{12}$ ,  $y_{11}$ ,  $b_{11}$ ,  $b_9$ , and  $b_7$ ) and the intense signals of the neutral loss from the product  $y$ -ions ( $y_{11}$  to  $y_{14}$ ) confirm that this peptide was also glycated at K13.

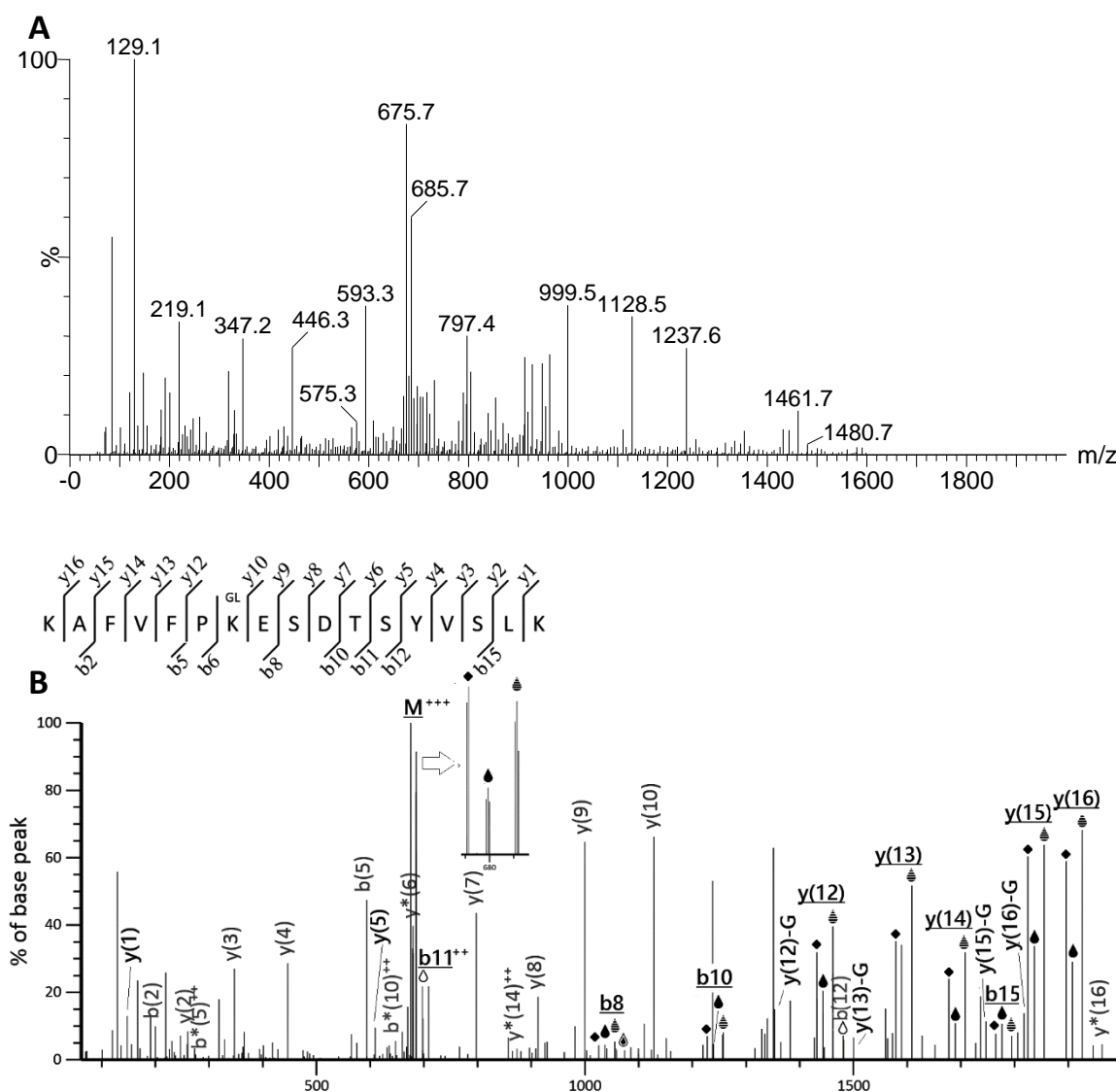


Figure 3-6 Assignment of product ion data for peptide sequence KAFVFPKESDTSYVSLK. A) nESI product ion mass spectrum of the precursor at  $703.34^{+3}$   $m/z$ , combined spectra were smoothed/centroided using MassLynx 4.1. B) Mascot identification of precursor ion at  $703.34^{+3}$   $m/z$ , assigned to peptide sequence KAFVFPKESDTSYVSLK. The glycation site was identified by manual interpretation of the spectrum. K13 is the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragment ion series, respectively. + is the number of charges, \* indicates deamination. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\diamond$  (loss of 1  $H_2O$ ),  $\blacklozenge$  (loss of 2  $H_2O$ ),  $\blacktriangle$  (loss of 3  $H_2O$ ),  $\blacklozenge$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O$ +HCHO), and G (loss of  $C_6H_{10}O_5$ ).

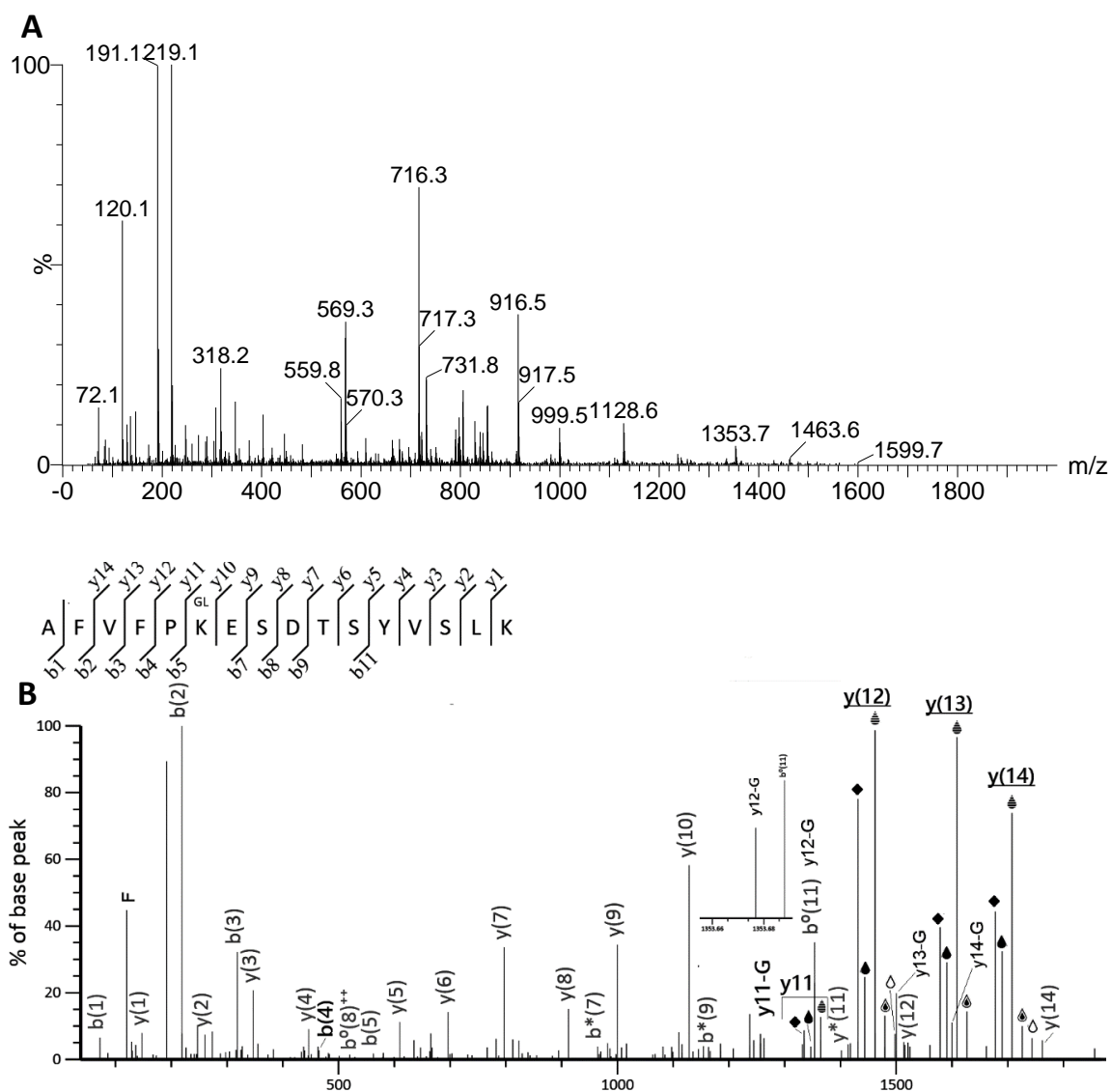


Figure 3-7 Assignment of product ion data for peptide sequence AFVFPKESDTSYVSLK. A) nESI product ion mass spectrum of the precursor at  $660.65^{+3}$   $m/z$ , combined spectra were smoothed/centroided using MassLynx 4.1. B) Mascot identification of precursor ion at  $660.65^{+3}$   $m/z$ , assigned to peptide sequence AFVFPKESDTSYVSLK. The glycation site was identified by manual interpretation of the spectrum. K13 is the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragment ion series, respectively. + is the number of charges, \* indicates deamination. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\triangle$  (loss of 1  $H_2O$ ),  $\blacktriangle$  (loss of 2  $H_2O$ ),  $\bullet$  (loss of 3  $H_2O$ ),  $\blacklozenge$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O$ +HCHO), and G (loss of  $C_6H_{10}O_5$ ).



The MS/MS spectrum of the glycated peptide GYSIFSYATKR at  $m/z$  727.84<sup>+2</sup> shows abundant signals corresponding to mass losses of 3 H<sub>2</sub>O, 4 H<sub>2</sub>O and 3 H<sub>2</sub>O+HCHO from the parent ion confirming the glycation status of the peptide, while the fructosamine-containing product ions and their correlated neutral losses identify the glycation site as K57.

Additional glycated peptides were identified in the Glu-C digests of CRP samples incubated with 1M glucose for 19h and more extended periods: ILIFWSKDIGYSFTVGGSE, which has a single modification site (K69), and FWVDGKPRVRKSLKKGTVGAE, with six possible modification sites.

Mascot assigned the doubly charged peptide at  $m/z$  1141.0617<sup>+2</sup>, ILIFWSKDIGYSFTVGGSE, to the MS/MS spectrum in Figure 3-9.  $\gamma$ -ion and b-ion series corresponding to the entire peptide chain was observed, with a robust b-ion series showing mass shifts of +162 Da corresponding to glycated peptide fragments and assign the glycation to K69. The neutral loss patterns in MS/MS observed, but the major losses observed was -18 Da corresponding to the loss of 1 H<sub>2</sub>O. Other kinds of losses were also detected but with lesser intensity.



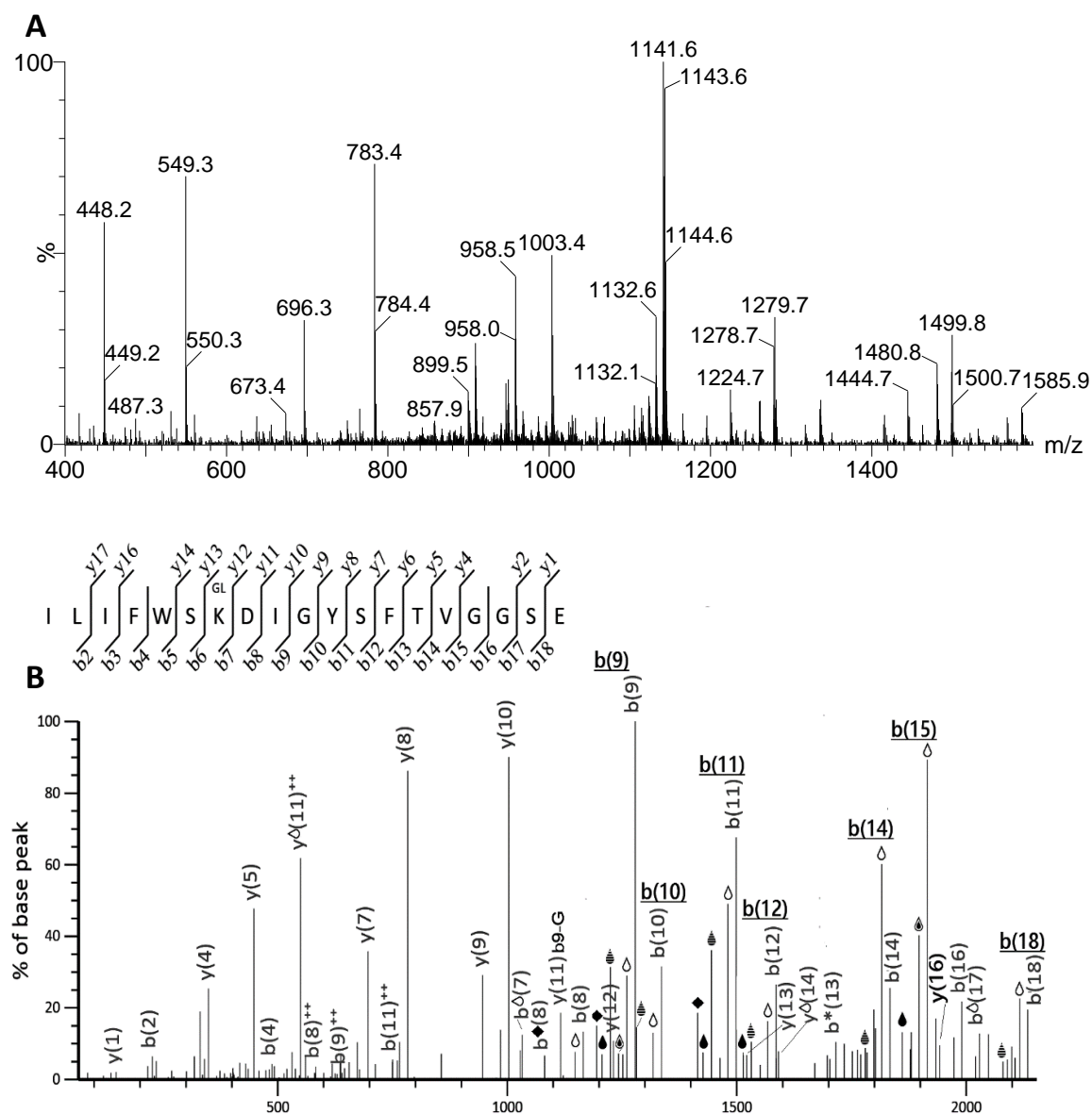


Figure 3-9 Assignment of product ion data for peptide sequence ILIFWSKDIGYSFTVGGSE. A) nESI product ion mass spectrum of the precursor at  $1141.06^{+2}$   $m/z$ , combined spectra were smoothed/centroided using MassLynx 4.1. B) Mascot identification of precursor ion at  $1141.06^{+2}$   $m/z$ , assigned to peptide sequence ILIFWSKDIGYSFTVGGSE. The glycation site was identified by manual interpretation of the spectrum. K69 is the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragment ion series, respectively. + is the number of charges, \* indicates deamination. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1  $H_2O$ ),  $\circ$  (loss of 2  $H_2O$ ),  $\ominus$  (loss of 3  $H_2O$ ),  $\blacklozenge$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O+HCHO$ ), and G (loss of  $C_6H_{10}O_5$ ).

The MS/MS spectra of triply, quadruply or the quintuply charged glycated peptide FWVDGKPRVRKSLKKG YTVGAE shows only limited sequence information that does not help to assign the modification to any of the 6 possible modification residues in the sequence.

Figure 3-10 shows the fragmentation spectrum of the quadruply charged peptide at  $m/z$  671.61<sup>+4</sup>.

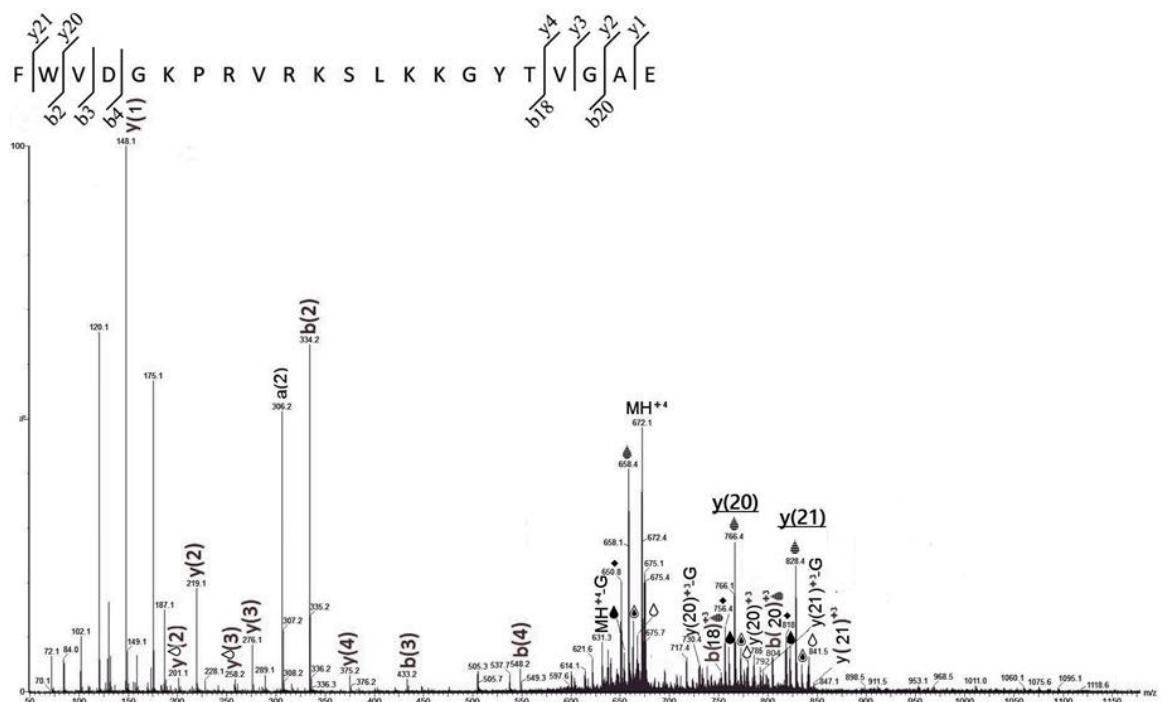


Figure 3-10 Assignment of product ion data for peptide sequence FWVDGKPRVRKSLKKG YTVGAE. Combined spectra of the precursor at 671.61<sup>+4</sup>  $m/z$  were smoothed/centroided using MassLynx 4.1. The product ion data were manually assigned to the peptide sequence FWVDGKPRVRKSLKKG YTVGAE with one glucose molecule. The fragmentation spectrum did not provide enough information to identify the glycation site. b<sub>n</sub> and y<sub>n</sub> are N- and C-terminal fragment ion series, respectively. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1 H<sub>2</sub>O),  $\blacklozenge$  (loss of 2 H<sub>2</sub>O),  $\blacktriangle$  (loss of 3 H<sub>2</sub>O),  $\blacklozenge$  (loss of 4 H<sub>2</sub>O),  $\blacklozenge$  (loss of -3 H<sub>2</sub>O+HCHO), and G (loss of C<sub>6</sub>H<sub>10</sub>O<sub>5</sub>).

The parent ion is observed as abundant signal shifted by +162 and associated with peaks of varied intensity corresponding to neutral losses of 1 to 4 molecules of water, loss of 84 Da (3 H<sub>2</sub>O+HCHO) and 162 Da (C<sub>6</sub>H<sub>10</sub>O<sub>5</sub>). The glycated product ions  $\gamma_{20}$  and  $\gamma_{21}$  were identified and associated with signals corresponding to all described neutral losses ( H<sub>2</sub>O, 2 H<sub>2</sub>O, 3 H<sub>2</sub>O, 4 H<sub>2</sub>O, 3 H<sub>2</sub>O + HCHO, and C<sub>6</sub>H<sub>10</sub>O<sub>5</sub>); however, the ions corresponding to the loss of 3 molecules of H<sub>2</sub>O was most abundant. Moreover, signals corresponding to neutral losses of 3 H<sub>2</sub>O from fructosamine-containing product ions  $b_{18}$  and  $b_{20}$  were observed.

Despite this information confirming the presence of glycation but any attempt to identify the glycation site with such information-weak fragmentation spectra are likely to result in false positives.

#### *3.3.4.2 Glycation sites in CRP after 3 days exposure to 1 M glucose solution*

Increasing the incubation time to 3 days allowed the glycation on K191 besides on the previously recognised glycation sites. The MS/MS spectrum of the tryptic glycated peptide ALKYEYVQGEVFTKPQLWP- at  $m/z$  765.724<sup>+3</sup> (Figure 3-11) afford a series of product ions corresponding to most of the peptide chain. Although this peptide has two potential glycation sites, K191 and K201, the presence of unglycated  $\gamma$ -ions from  $\gamma_6$  to  $\gamma_{15}$  and glycated  $b$ -ions  $b_4$  and from  $b_6$  to  $b_{12}$  suggest the modification to be predominantly associated with K191 rather than K201. Consecutive losses of water and formaldehyde from the glycated  $b$ -ions species (mainly from  $b_9$  to  $b_{12}$ ) further suggest glycation of K191. The signals corresponding to the loss of three water molecules were the strongest among the other neutral loss signals.

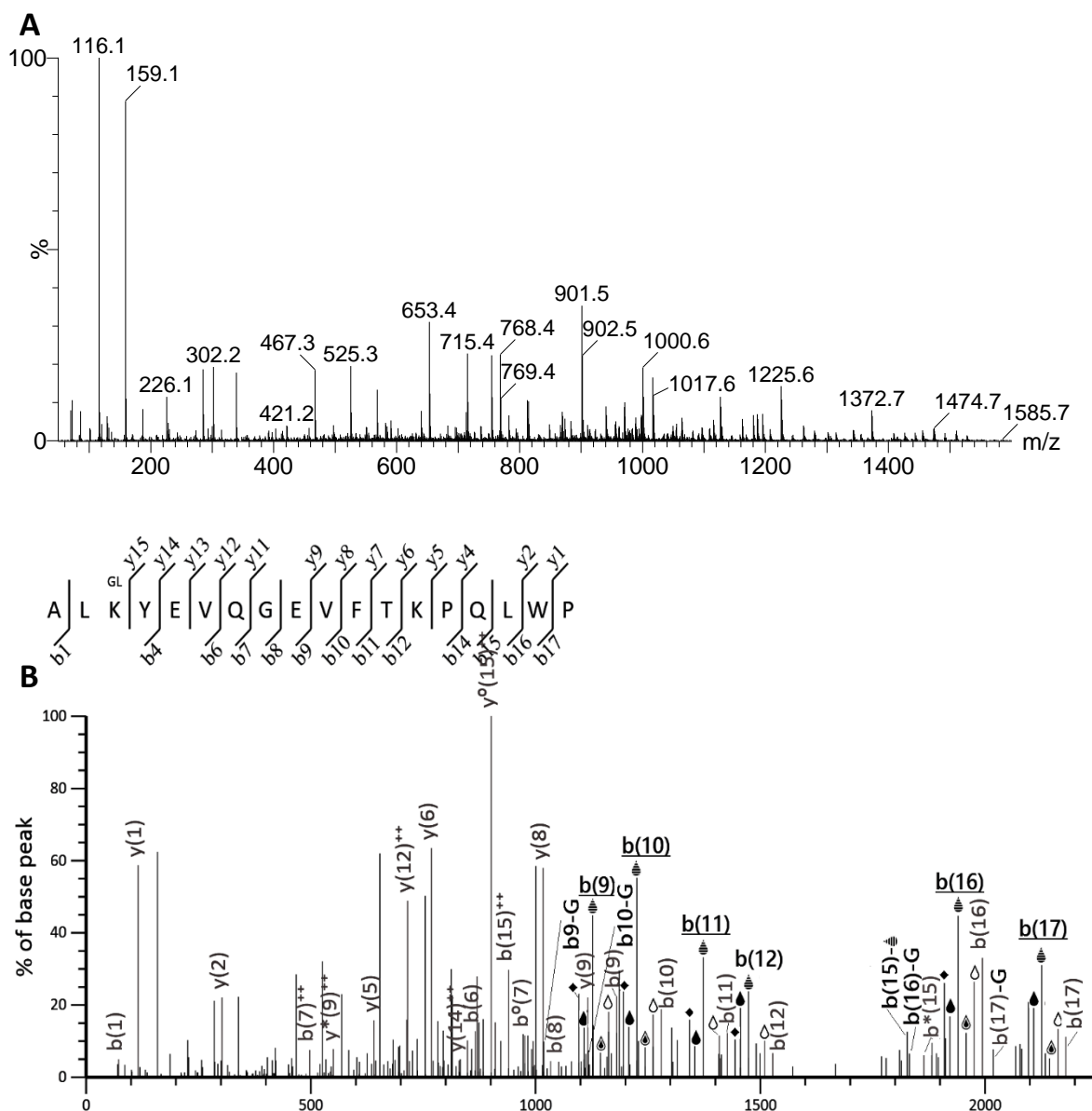


Figure 3-11 Assignment of product ion data for peptide sequence ALKYEYVQGEVFTKPKQLWP. A) nESI product ion mass spectrum of the precursor at  $765.72^{+3}$   $m/z$ , combined spectra were smoothed/centroided using MassLynx 4.1. B) Mascot identification of precursor ion at  $765.72^{+3}$   $m/z$ , assigned to peptide sequence ALKYEYVQGEVFTKPKQLWP. The glycation site was identified by manual interpretation of the spectrum. K191 is the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragments, respectively. + is the number of charges, \* indicates deamination. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1  $H_2O$ ),  $\Delta\Delta$  (loss of 2  $H_2O$ ),  $\Delta\Delta\Delta$  (loss of 3  $H_2O$ ),  $\Delta\Delta\Delta\Delta$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O$ +HCHO), and G (loss of  $C_6H_{10}O_5$ ).

### 3.3.4.3 Glycation sites in CRP after 6 days exposure to 1 M glucose solution

K201 seems to be glycated only after increasing incubation to 6 days. The MS/MS spectrum of the doubly charged tryptic peptide at  $m/z$  991.9854<sup>+2</sup>, YEVQGEVFTKPQLWP.-, confirmed the glycation on this residue, Figure 3-12. The presence of the Amadori  $\gamma$ - and  $\beta$ - product ions shifted by +162 Da together with the unglycated ions, cover the majority of the glycated peptide sequence, and allowed the identification of the modification site. Signals corresponding to the loss 1 H<sub>2</sub>O, 3 H<sub>2</sub>O and -C<sub>6</sub>H<sub>10</sub>O<sub>5</sub> were the most recognised losses from the fructosamine-containing product ions.

During this incubation period, the peptide FWVDGKPRVRKSLKKGTVGAE was observed with a  $m/z$  value consistent with glycation by two glucose molecules. The product ion spectrum showed very few product ions sufficient to confirm the presence of the two glucose molecules, however, these fragments are inadequate to confirm or even to exclude any of the potential modification sites.

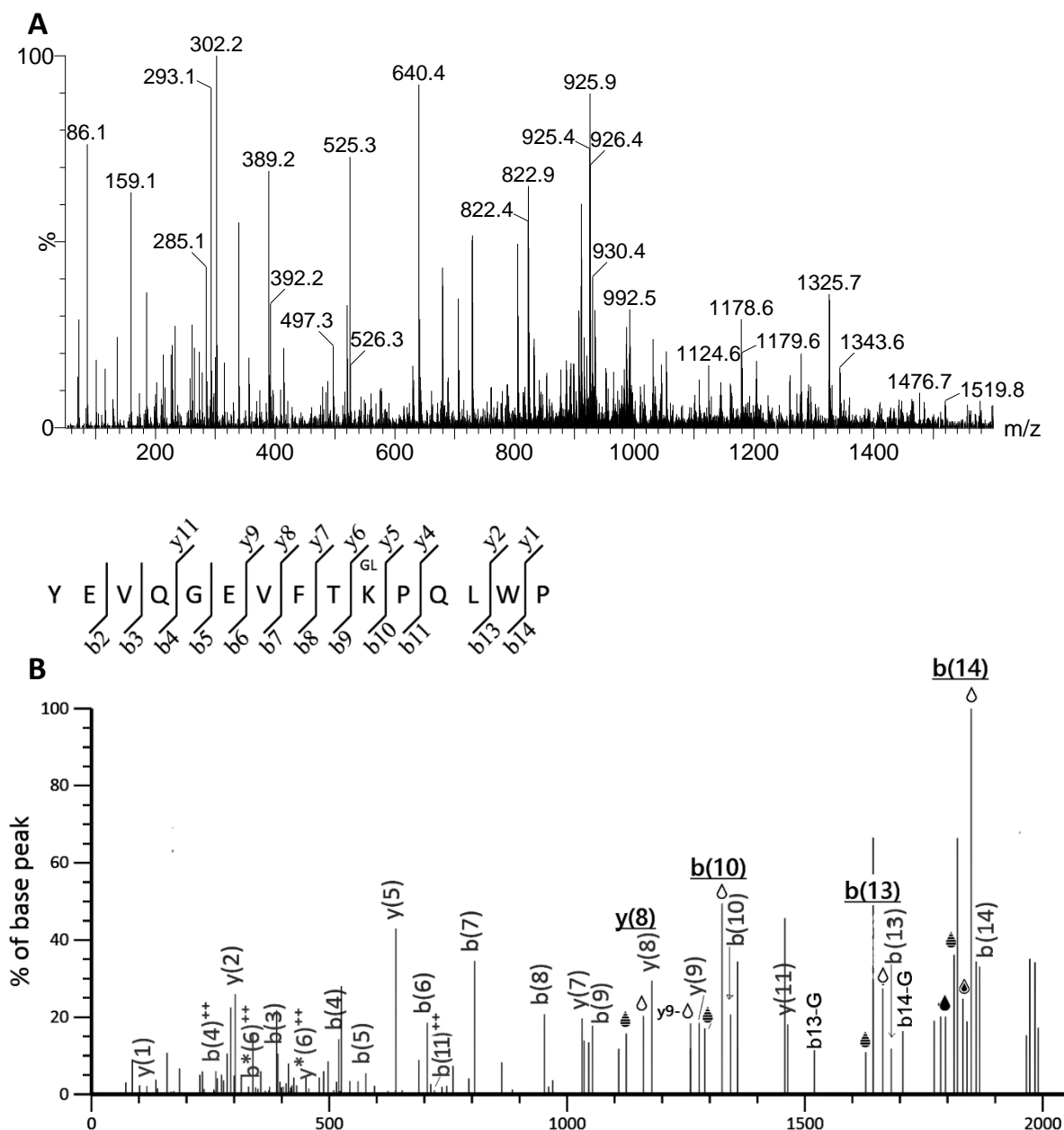


Figure 3-12 Assignment of product ion data for peptide sequence YEVQGEVFTKPLWP. A) nESI product ion mass spectrum of the precursor at  $991.98^{+2}$   $m/z$ , combined spectra were smoothed/centroided using MassLynx 4.1. B) Mascot identification of precursor ion at  $991.98^{+2}$   $m/z$ , assigned to peptide sequence YEVQGEVFTKPLWP. The glycation site was identified by manual interpretation of the spectrum. K201 is the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragments, respectively. + is the number of charges, \* indicates deamination. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1  $H_2O$ ),  $\blacklozenge$  (loss of 2  $H_2O$ ),  $\blacktriangle$  (loss of 3  $H_2O$ ),  $\blacklozenge$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O+HCHO$ ), and G (loss of  $C_6H_{10}O_5$ ).

Some of the glycated residues were identified in the digest of both enzymes. Figure 3-13, shows the fragmentation spectrum of the glycated peptide LSSTRGYSIFS<sup>GL</sup>YATKRQDNE at  $m/z$  829.04<sup>+3</sup> from the Glu-C digest, the modification was assigned to the residue K57. This glycated residue (K57) was already identified in the glycated peptide GYSIFS<sup>GL</sup>YATKR from the tryptic digest of CRP exposed to glucose solutions for 19 h and more extended periods.

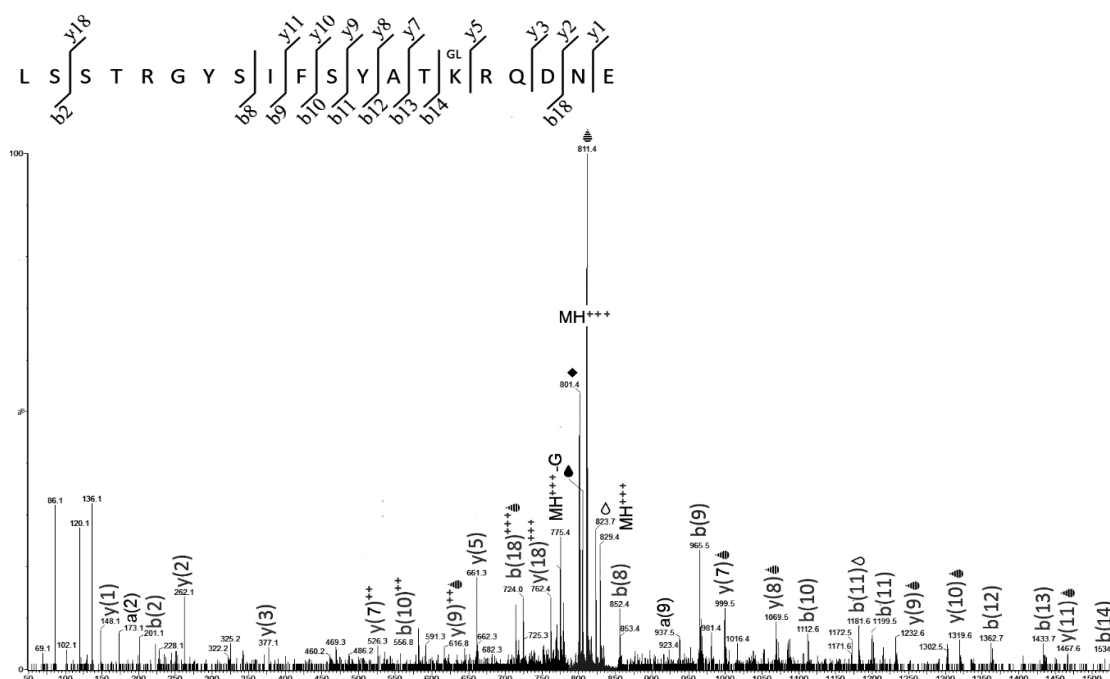


Figure 3-13 Assignment of product ion data for peptide sequence LSSTRGYSIFS<sup>GL</sup>YATKRQDNE. Combined spectra of the precursor at 829.04<sup>+3</sup>  $m/z$  were smoothed/centroided using MassLynx 4.1. The product ion data were manually assigned to the peptide sequence LSSTRGYSIFS<sup>GL</sup>YATKRQDNE with one glucose molecule. K57 is identified as the glycation site and indicated in the above sequence by (GL). bn and yn are N- and C-terminal fragment ion series, respectively. + is the number of charges. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1 H<sub>2</sub>O),  $\circ$  (loss of 2 H<sub>2</sub>O),  $\ominus$  (loss of 3 H<sub>2</sub>O),  $\blacklozenge$  (loss of 4 H<sub>2</sub>O),  $\blacklozenge$  (loss of -3 H<sub>2</sub>O+HCHO), and G (loss of C<sub>6</sub>H<sub>10</sub>O<sub>5</sub>).

Furthermore, K191 is identified as glycation site in the glycated Glu-C peptide INTIYLGPFSPNVLNWRALKYE; the fragmentation spectrum is shown in Figure 3-14. The same residue was identified in the tryptic glycated peptide ALKYEVQGEVFTKPQLWP.

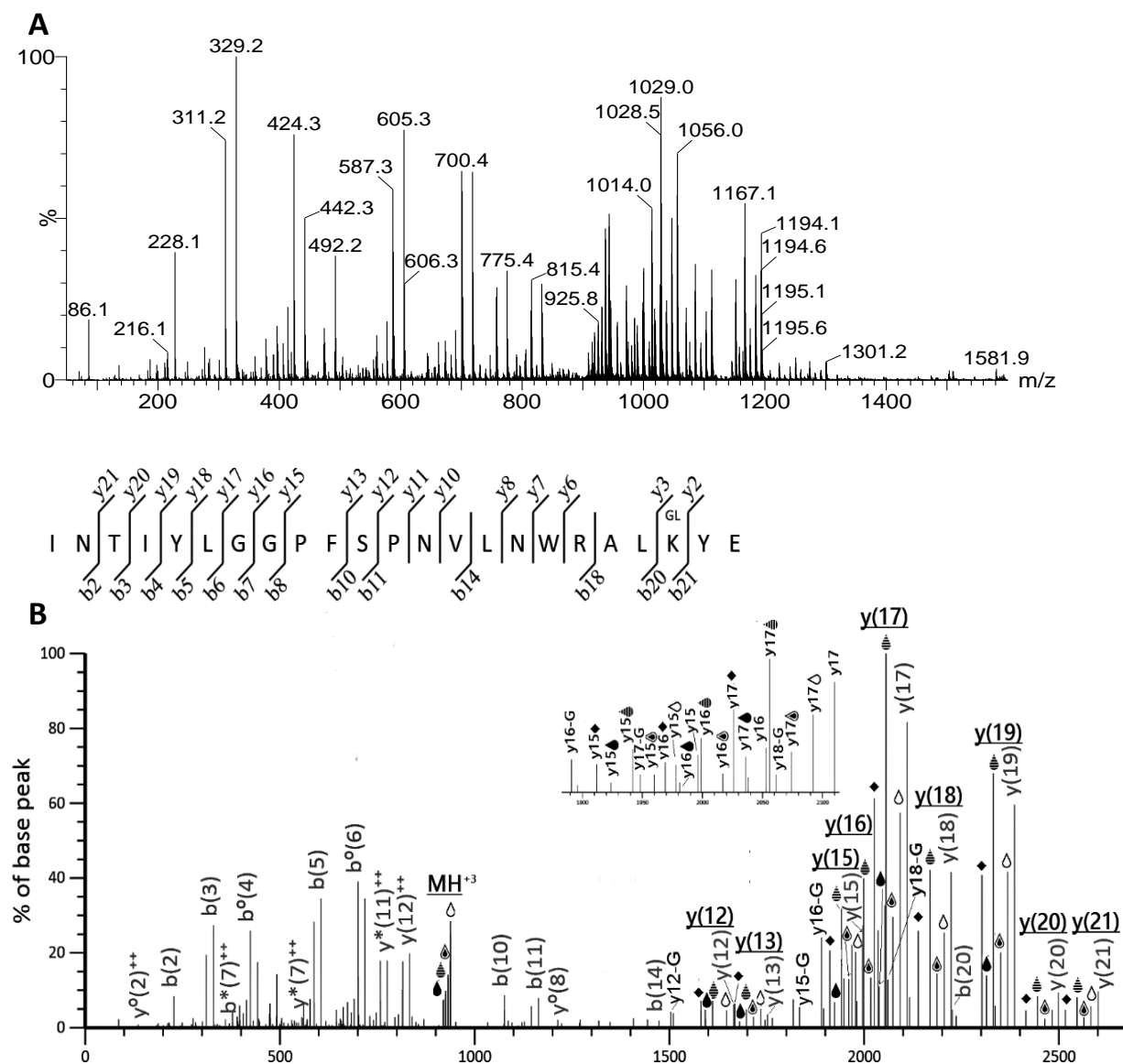


Figure 3-14 Assignment of product ion data for peptide INTIYLGPFSPNVLNWRALKYE. A) nESI product ion mass spectrum of the precursor at  $943.14^{+3}$   $m/z$ , combined spectra were smoothed/centroided using MassLynx 4.1. B) Mascot identification of precursor ion at  $943.14^{+3}$   $m/z$ , assigned to peptide INTIYLGPFSPNVLNWRALKYE. The glycation site was identified by manual interpretation of the spectrum. K191 is the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragments, respectively. + is the number of charges, \* indicates deamination. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1  $H_2O$ ),  $\blacklozenge$  (loss of 2  $H_2O$ ),  $\blacktriangle$  (loss of 3  $H_2O$ ),  $\blacklozenge$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O$ +HCHO), and G (loss of  $C_6H_{10}O_5$ ).



#### 3.3.4.4 Unmatched *m/z* values after Mascot search

Due to the fragmentation behaviour of the glycated peptides, *m/z* values corresponding to glycated CRP peptides may not be assigned to CRP during the Mascot search. For this reason, all the unmatched *m/z* values during the Mascot search were considered as possible CRP glycated peptides and further explored as detailed in section 3.2.4.

After analysis, two *m/z* values were found to have a match with a glycated CRP peptide within 100 ppm threshold. The species at *m/z* 829.04<sup>+3</sup>, which present in Glu-C digest of CRP sample incubated for 3 days, match the glycated peptide LSSTRGYSIFSYATKRQDNE. The fragmentation spectrum of this peptide assigns the modification to K57. This modification site is already identified in the tryptic digest of CRP incubated for 19 h (Figure 3-8) and longer periods, and in the Glu-C digest but only in the sample incubated for 6 days, see figure 3-13.

The second matched peptide, which presents in the tryptic digest of CRP sample incubated for 3 days, has *m/z* 713.7105<sup>+3</sup> and corresponds to glycated peptide ESDTSYVSLKAPLTKPLK, Figure 3-15. The fragmentation spectrum of this triply charged peptide shows a high-quality series of fragment ions shifted by +108 Da and +78 Da corresponding to the formation of pyrylium and furylium ions from the fructosamine-containing product ions  $y_9$  to  $y_{17}$ , while the conventional glycated  $y$  fragments (M+162) were missing. Pyrylium and furylium ions fragments confirmed the presence of glycation and provided enough information to cover most of the peptide sequence. The strong signals of the unglycated fragments  $y_8$ ,  $y_7$  and a detectable signal of  $y_5$  ions localised the glycation site on K23 and excluding K28.

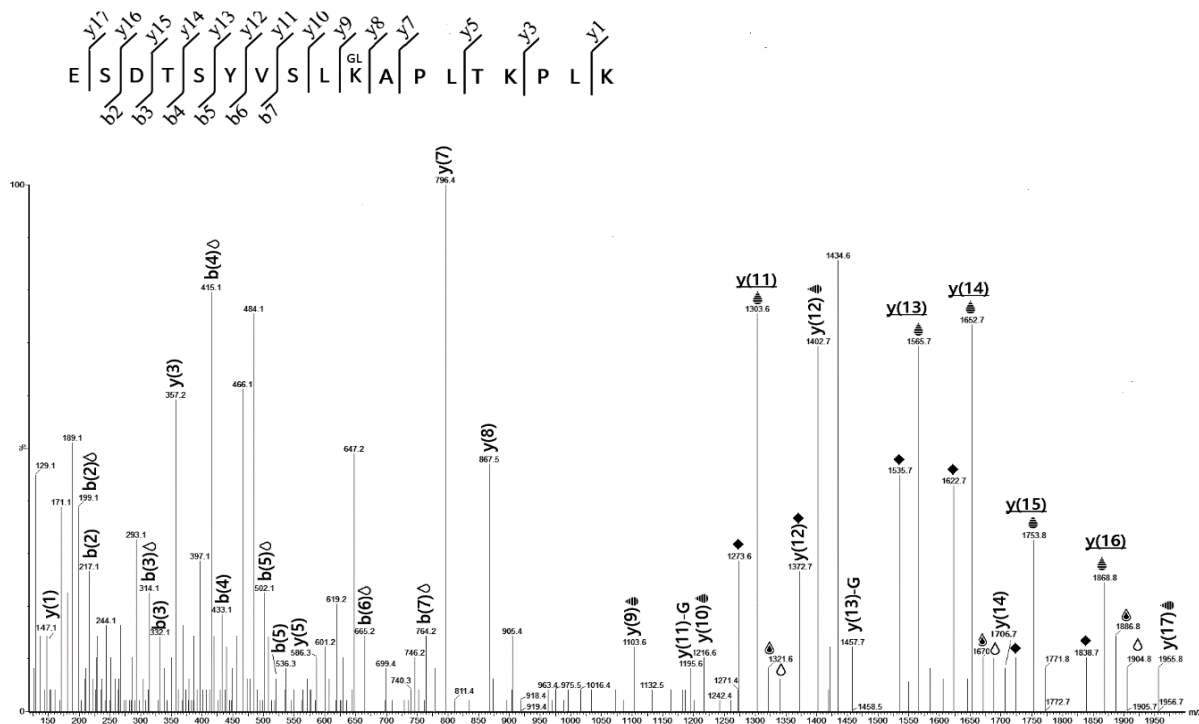


Figure 3-15 Assignment of product ion data for peptide sequence ESDTSYVSLKAPLTKPLK. Combined spectra of the precursor at  $713.71^{+3}$   $m/z$  were smoothed/centroided using MassLynx 4.1. The product ion data were manually assigned to the peptide sequence ESDTSYVSLKAPLTKPLK with one glucose molecule. K23 is identified as the glycation site and indicated in the above sequence by (GL).  $b_n$  and  $y_n$  are N- and C-terminal fragment ion series, respectively. Neutral losses from the fructosamine-containing parent and fragment ions are indicated by  $\Delta$  (loss of 1  $H_2O$ ),  $\Delta\Delta$  (loss of 2  $H_2O$ ),  $\Delta\Delta\Delta$  (loss of 3  $H_2O$ ),  $\Delta\Delta\Delta\Delta$  (loss of 4  $H_2O$ ),  $\blacklozenge$  (loss of -3  $H_2O$ +HCHO), and G (loss of  $C_6H_{10}O_5$ ).

### 3.3.5 Determination the most reactive lysine residue

CRP samples were incubated with decreasing concentration of glucose to determine the most reactive lysine residue. The samples were enzymatically digested by trypsin and Glu-C separately, and the resulting digest was analysed by LC-nESI-QqTOF-MS/MS. The data obtained were submitted to the Mascot search engine which allowed identification of glycated CRP peptides. The  $m/z$  values that did not match any peptide sequence during the Mascot search were considered as possible CRP glycated peptides and further analysed as explained in section 3.2.4. The results are summarised in Table 3-4.

Table 3-4 The Glycated peptides identified in the tryptic and Glu-C digests of the CRP samples incubated for 19 hours with a varying concentration of glucose. The incubation time of the CRP sample exposed to 30 mM of glucose was also extended to 3 days. The samples were analysed by LC-nESI-QqTOF-MS/MS, and the acquired data were searched against the Mascot using a database comprising the CRP sequence and common laboratory-observed contaminants. The peptide sequence was confirmed by manual interpretation of its fragmentation spectrum.

Glucose concentration/ incubation time	Peptide sequence	Observed <i>m/z</i> value	Experimental mass	Mass error ppm
200 mM/ 19 h	KAFVFPKESDTSYVSLK	703.3563	2107.0471	-12.05
100 mM/ 19 h	KAFVFPKESDTSYVSLK	703.3604	2107.0595	-6.17
50 mM/ 19 h	-	-	-	-
30 mM/ 19 h	-	-	-	-
30 mM/ 3 days	KAFVFPKESDTSYVSLK	703.3459	2107.0157	-26.95

CRP exposed to 200 mM of glucose for 19 h shows one glycated peptide, KAFVFPKESDTSYVSLK, at  $m/z$  703.36<sup>+3</sup> with the glycation site in a mid-chain position corresponding to K13 of CRP as confirmed by manual interpretation of the product ion spectrum. This glycated peptide was identified in the tryptic digest of the protein while no modified peptide was observed in the Glu-C digest. The same result was obtained after reducing the glucose to 100 mM for the same incubation time. However, no glycated peptide can be identified (in both tryptic or Glu-C digest) in the CRP samples exposed to glucose concentration less than 100 mM for 19 h, while the sample exposed to 30 mM of glucose but for 3 days reveal the same glycation site on K13.

### 3.3.6 Characteristics of glycated lysine residues

The proximity of acidic and basic groups to the lysine and arginine residues in the primary or tertiary structure of CRP, the local pKa value of arginines and lysines in CRP sequence, and the fractional accessible surface (FAS) of the side chain of these residues were calculated and presented in Table 3-5. This information was utilised to inform our understanding of observed differences in lysine residue glycation, and also to predict the

most likely modification sites within a glycated peptide bearing more than one potential modification site.

When comparing the glycated residues detected in the present study (group A in Table 3-4) with the unglycated lysine and arginine residues (group C in Table 3-4), it was noted that glycated lysines are located in closer proximity to acidic residues, with the acidic residue is mainly situated C-terminally to the glycation site.

All glycated lysines except K191 are located adjacent to carboxylate groups within distances  $<4 \text{ \AA}$ . The glycated lysine K191, observed only after three days of incubation, is  $9.7 \text{ \AA}$  distant from Glu 193. However, this lysine has the highest FAS (0.692) compared to other glycated residues. On the other hand, K201 has the lowest FAS (0.232) compared to the other glycated lysines and was found modified only after 6 days of incubation.

The modified lysines that have been detected after 19 h incubation times (K13, K57 and K69) share the property of having an acidic residue in the C-terminal side, and shorter distance between the  $\epsilon\text{-NH}^+$  group of the glycated residue to the carboxylate group of the acidic residues; the distance did not exceed  $3.6 \text{ \AA}$  to the nearest free acid group.

Group B in the Table 3-4 summarise the properties of the lysine and arginine residues of the FWVDGKPRVRKSLKKGTYVGAE peptide. This peptide was modified with up to two glucose molecules. However, the fragmentation spectrum did not provide enough information to identify the glycation sites. The arginens in this peptide (R116 and R118) has the highest pKa values and lowest FAS compared to the lysines of the same region. In contrast, K114 has relatively low pKa value (10.38) and the greatest FAS (0.832) compared to the other residues. With the other three residues (K119, K122, K123), the pKa and FAS

were either both low or both high. For example, K199 which has the lowest pKa value (9.61) it has the smallest FAS (0.24) compared with the other lysines in the same sequence.

Table 3-5 Physicochemical characteristics of CRP lysine and arginine residues. Lysine and arginine residues were divided into three groups; group A included the glycated lysine detected in this study, B: the lysine and arginine residue of the glycated peptide FWVDGKPRVRKSLKKGTVGAE, C: unglycated lysine and arginine in this study. The distance to the nearby acidic and basic residues was computed using PyMOL, the local pKa values were estimated by using PROPKA 3.0, The FAS of the side chain of lysine and arginine were calculated by using VADAR. The \* represents residues on the adjacent protomer.

Group	Residue	Distance to basic residues in Å	Distance to acidic residues Å	Estimated pKa	Estimated FAS
A	K13	6.6 to K123*	3.3 to E197, 12.1 to E14, 14.8 to D16	10.684	0.556
	K57	9.1 to R58	3.1 to E130, 14.6 to D60	11.27	0.55
	K69		3.6 to D70	10.88	0.44
	k191		9.7 to E193	10.352	0.692
	K23	9.5 to K191	3.8 to E193	10.798	0.516
	k201	6.3 to R118	3.1 to E101*, 6.5 to E108*	10.466	0.232
B	K114	11.1 to R116	8.1 to D112	10.38	0.832
	K119	8.3 to R116	3.9 to E42*, 5.6 to E85	9.614	0.24
	K122	14.3 to K123	7.6 to E62	10.16	0.364
	K123	6 to K13*, 14.3 to K122	9.7 to E101, 2.8 to E197*	11.204	0.472
	R116	8.3 to K119, 11.1 to K114	4.6 to E42*	12.218	0.194
	R118	6.3 to K201, 15 to R116	9.3 to E101, 2.8 to D155*	13.296	0.008
C	K7	11.6 to R6	6.6 to D3	10.392	0.454
	K28	8.5 to K31	8.3 to D163	10.432	0.876
	K31	8.5 to K28	8 to E101	10.25	0.694
	R6	8.7 to H38, 11.6 to K7	5.4 to D169*, 10 to D3	12.82	0.686
	R47		2.9 to E14, 9.5 to D16	11.942	0.216
	R58	9.1 to K57	8.2 to E130, 10.2 to D60	12.49	0.928
	R188		5.1 to D3	12.456	0.576

### 3.4 Discussion

The goal of the present study was to characterise the putative glycation sites of CRP exposed to glucose, using tandem mass spectrometry and to determine the most reactive residues for glycation in order to extrapolate the generated information to the *in vivo* situation. The general strategy employed in the study may be summarised as follows: CRP digestion with trypsin and Glu-C, tandem mass spectrometry analysis of digests, Mascot searching of data against customised database comprising the CRP sequence and common laboratory-observed contaminants, manual validation of the MS/MS spectra that were assigned to glycated CRP peptides through Mascot and interpretation of unassigned product ions, analysis of the  $m/z$  values that did not match any peptide sequence in the selected database during the Mascot analysis.

#### 3.4.1 C-reactive protein and enzymatic digestion

Protein digestion is the most critical step in the sample preparation for tandem mass spectrometry analysis (Hustoft *et al.*, 2012). CRP preparations were digested with trypsin due to its advantage in preparing peptides of appropriate mass range for MS sequencing (Hustoft *et al.*, 2012). Since the cleavage sites of trypsin could be involved in the glycation reaction (Lacinová *et al.*, 2010), endoproteinase Glu-C was also used as an alternative single protease to enhance protein sequence coverage.

CRP has been found to be relatively resistant to enzymatic digestion. After several digestion trials with low-efficiency cleavage (data are not shown), it was found that additional denaturation steps, higher CRP: enzyme molar ratio and longer incubation times with proteases were required to improve its degradation. Taylor *et al.* demonstrated that CRP in its native pentameric form is highly-resistant to enzymatic digestion, both degradation

to its monomeric form and the use of high enzyme: CRP molar ratio is required to achieve digestion with trypsin (Taylor & Van Den Berg, 2007). Monomerisation of CRP can be achieved either by heat or by denaturation with a chelating agent (Taylor & Van Den Berg, 2007, Potempa *et al.*, 1987). It found that one hour of heating at 70 °C is required to achieve complete monomerisation of CRP in the presence of calcium (Taylor & Van Den Berg, 2007). However, in the absence of calcium abbreviated incubation times can achieve the same results (Potempa *et al.*, 1983), because removal of calcium decreases the stability of the CRP pentamer and increases susceptibility to proteolysis and denaturation (Shrive *et al.*, 1996). Using chelating agents to remove calcium from the solution is not recommended with trypsin digestion (Nord *et al.*, 1956).

In this work, CRP preparations were denatured by using the combination of a mild denaturant (Rapigest) to unfold the protein and reveal the inaccessible cleavage sites (Huang *et al.*, 2009), and by heating the samples near 100 °C for 5 min followed by 30 minutes heating at 60 °C. Furthermore, the digestion time was extended to 24 h, with two additions of the enzyme at 30:1 to promote complete protein digestion. Combining the mapping results of both trypsin and Glu-C digested peptides resulted in high sequence coverage sufficient for the analysis of the N-terminus, all arginines, and at least 11 out of 13 lysine residues on CRP in all of the samples that were studied.

#### 3.4.2 Fragmentation behaviour of glycated peptides

Glycated peptides lose water molecules from the glycated lysine residues during CID as a result of dissociation of the most labile bonds in the peptide side chain modification (Zhang *et al.*, 2008). The MS/MS spectra of glycated CRP peptides showed a high abundance of peaks reflecting these multiple neutral losses from the glucose-containing parent and

fragment ions at a characteristic intensity distribution. Similar findings have been previously observed with glycated peptides analysed by CID (Lapolla *et al.*, 2004, Frolov *et al.*, 2006, Mennella *et al.*, 2006). The characteristic neutral losses observed with CRP glycated peptides, analysed by Q-TOF instrument, include consecutive elimination of 1 H<sub>2</sub>O, 2 H<sub>2</sub>O, 3 H<sub>2</sub>O, 4 H<sub>2</sub>O, 3 H<sub>2</sub>O + HCHO, and neutral loss of the entire Amadori product (C<sub>6</sub>H<sub>10</sub>O<sub>5</sub>). The most intense neutral loss signal commonly arose from pyrylium ions (neutral loss of 3 H<sub>2</sub>O) while the second most intense neutral loss signal typically derives from the furylium ions (neutral loss of 3 H<sub>2</sub>O + HCHO). Interestingly, pyrylium ions were observed from all glycated peptides studied here, frequently followed by the furylium ion signal, indicating that observation of pyrylium and furylium ions is strongly suggestive of the presence of a glycated peptide. The neutral loss of entire Amadori product was also observed in glycated  $\gamma$ -ion series. The conventional Amadori b- and  $\gamma$ -ions were very weak or even undetectable, while the non-modified product ions do not show any neutral loss, but the typical  $\gamma$ - and b-ions.

Although this fragmentation pattern has been previously reported (Lapolla *et al.*, 2004, Frolov *et al.*, 2006, Mennella *et al.*, 2006), unambiguous assignment of the modified amino acid residue was not consistently reported by Mascot. This was especially the case where multiple potential glycation sites were present on a single peptide. For instance, the glycated tryptic peptides KAFVFPKESDTSYVSLK and AFVFPKESDTSYVSLK share the same sequence except for a missed cleavage site at the N-terminus of the first sequence. Mascot confirmed the glycated peptide K<sub>7</sub>AFVFPK<sub>13</sub>ESDTSYVSLK sequence but assigned a higher probability score for glucose condensed on K7 than K13; this is less plausible, as lysine glycation should obstruct trypsin cleavage (Wa *et al.*, 2007, Zhang *et al.*, 2003). Since the glycated b- and  $\gamma$ -ions were very weak or missing, the majority of product ions used by



Mascot in the assignment of the glycated residue are those not involved in Amadori product formation, while the signals from successive neutral losses, including the strong signals of pyrylium and furylium ions were neglected, resulting in confounding the automatic sequence assignment and glycation site identification by Mascot.

Manual assignment of the unmatched neutral losses provides the missing piece of information and reveals the ambiguity of the modification site. Although neutral losses of H<sub>2</sub>O and 2 H<sub>2</sub>O can also occur from other peptide functional groups, such as neutral loss of H<sub>2</sub>O from peptides containing hydroxyl or carboxylic acid residues (Zhang *et al.*, 2008), neutral losses of 3 H<sub>2</sub>O, 4 H<sub>2</sub>O, 3 H<sub>2</sub>O + HCHO and C<sub>6</sub>H<sub>10</sub>O<sub>5</sub> are significantly less common and characterise the Amadori-modified peptides (Zhang *et al.*, 2008, Zhang *et al.*, 2008b). Identification of these fragment ions confirmed the presence of glycation and allowed positive determination of the site of modification. Furthermore, besides the glycated product ions, the presence of surviving precursor ion of glycated molecular weight and neutral loss products from precursor also confirm the presence of a glycated peptide. When all potential neutral losses from the product and precursor ions were considered in the MS/MS spectra analysis, the majority of the backbone ions can be determined, and the modification site can be identified and localised (Frolov *et al.*, 2006). Considering the neutral loss signals from the glycated product ions confidently shows that both peptides, KAFVFPKESDTSYVSLK and AFVFPKESDTSYVSLK, were glycated at K13 (see Figure 3-6 and 3-7). Another example showing the utility of neutral loss peaks in assigning the modification site is the tryptic glycated peptide ALKYEYQGEVFTKQLWP- at  $m/z$  765.724<sup>+3</sup>, bearing two possible glycation sites; K191 and K201 (see Figure 3-11). Here, the strong pyrylium and furylium ion signals derived from glycated b-ion species allowed positive determination of the primary glycation site at K191 rather than K201. The precursor peptide ion at

713.7105<sup>+3</sup> did not find any match against CRP sequence via Mascot searching as no glycosylated b- and y- ions were observed. The MS/MS spectrum of this peptide shows a high-quality series of fragment ions with mass shifts of +108 amu and +78 amu vs their unmodified counterparts, corresponding to the formation of pyrylium and furylium ions from the glycosylated residue, confirming the presence of glycation. When these ions are utilised together with the unglycosylated fragments, the majority of the peptide sequence can be covered, and the glycation site can be localised to K23 (see Figure 3-15).

#### 3.4.3 Glycation sites of the *in vitro* glycated C-reactive protein

The potential glycation sites of CRP include 13 lysines, 6 arginines and the N-terminal amine. These sites are primarily clustered at the outer boundaries of each subunit in the tertiary structure of CRP as shown in Figure 3-16.

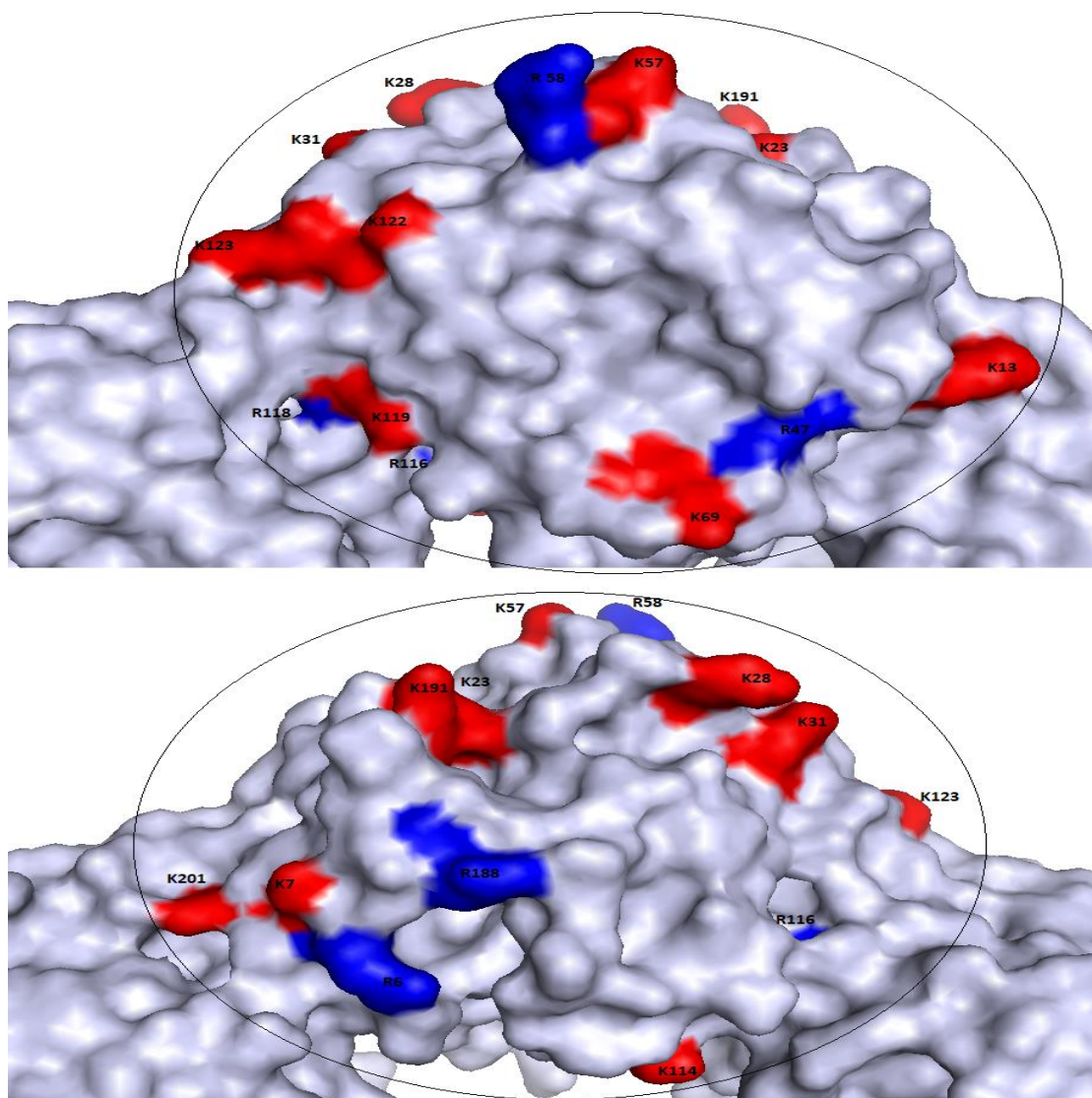


Figure 3-16 Crystal structure of CRP highlighting putative glycation sites. CRP structure PDB entry 1B09, showing one protomer (Top; ligand recognition face, bottom; the effector face). Lysine residues (K) are coloured red, arginine residues (R) are coloured blue

Based on trypsin and Glu-C experiments (Tables 3-2 and 3-3) and after manual validation of the MS/MS spectra for each glycated peptide, three glycation sites were identified in CRP preparations exposed to 1M of glucose for 19 hours; these are K13, K57 and K69. Occupation of the same glycation sites was also observed after increasing the incubation time to 3 and 6 days, suggesting that these residues are favourable glycation sites at this glucose concentration. After 3 days of incubation, K191 (which is also seen after 6 days)

and K23 were modified, while K201 was observed as being modified only after 6 days of incubation. This finding suggests that glycation is not randomly occurring, and the free amino groups of CRP differ in their rates of glycation (Zhang *et al.*, 2008). Besides the glycation sites mentioned above, the peptide FWVDGKPRVRKSLKKGTVGAE was found to be glycated by one glucose molecule after 19 h, and by two glucose after six days of exposure to 1M glucose concentration. The MS/MS spectrum of this peptide shows limited sequence information, sufficient to confirm the presence of the glucose moiety, but insufficient to localise the precise modification sites. In addition to the weak fragmentation pattern of glycated peptides under CID condition (Zhang *et al.*, 2008), the presence of 6 basic amino acids residues in the middle of this 22 residues peptide resulted in localised high positive charge density, which might be why simple peptide fragmentation is prevented (Xue *et al.*, 2015). It is worthy of mention that several residues of FWVDGKPRVRKSLKKGTVGAE peptide are involved in inter-subunit interactions with the neighbouring protomer in the CRP pentamer including the following amino acid pairs; K119 of one protomer with E42 of the next protomer, K123 with E197 of the next protomer, R118 with D155 and P202 from the next protomer (Thompson *et al.*, 1999, Shrive *et al.*, 1996). K201, which seen after 6 days of incubation, is also known to form a salt bridge with E101 from the next protomer (Shrive *et al.*, 1996).

Positively, the results of the tandem mass analysis are comparable with the result of intact mass analysis for the same samples. The number of condensed glucose molecules as determined intact mass analysis were 3, 5 and 8 molecules for samples incubated with 1 M of glucose for 19 h, 3 days and 6 days respectively (see Figures 2-6, 2-7 and 2-8). This corresponds closely to the total number of glycation sites identified in these samples; where 3 glycation sites plus one further site in 109-130 peptide were identified in sample

treated with 1M glucose for 19 h, 5 glycation sites plus one further site in 109-130 peptide, and 5 glycation sites plus two further sites in 109-130 peptide in the sample exposed to 6 days.

To identify the most reactive lysine residues towards glycation, CRP were exposed to lower concentrations of glucose for 19h. We observed that only the glycation site on K13 could be identified at a glucose level of 200 and 100 mM suggesting that this residue is more susceptible to glycation than other lysines. CRP incubated glucose levels less than 100mM for 19h incubation showed no evidence of glycation, suggesting that the extent of glycation was insufficient to be detected under our experimental conditions. However, increasing the incubation time, even at 30 mM glucose, allows observation of K13, as the glycation is directly proportional to the exposure time.

#### 3.4.4 Factors affecting glycation site specificity

Factors known to influence the reactivity of lysines or arginines towards the glycation reaction were investigated to understand why some of these residues are more prone to glycation or glycated at a faster rate than the others. The investigated factors were the position of the modified residues and their proximity to the nearby charged residues within the primary or tertiary structure, FAS, and pKa values of the side chain of the of the potential glycation sites (Shapiro *et al.*, 1980, Garlick & Mazer, 1983, Rabbani *et al.*, 2016).

All the identified glycation sites were lysine residues located at the outer surface of the protein, excepting K201, a residue which is implicated in intersubunit interaction with adjacent protomer (Shrive *et al.*, 1996). Because the formation of Schiff bases requires the free amino groups to be unprotonated to carry out the nucleophilic addition or substitution, lysine or arginine residues with lower pKa values would be expected to be

more susceptible to glycation reactions (Watkins *et al.*, 1985, Johansen *et al.*, 2006). However, the pKa values of the glycated lysine residues alone cannot directly explain the observed specificity (Shapiro *et al.*, 1980, Garlick & Mazer, 1983, Silva *et al.*, 2018). For instance, K57 has the highest pKa value among the lysine residues of CRP but is found to be glycated even during the shortest incubation period.

FAS and the properties of nearby amino acids been proposed to play a significant role in determining glycation site specificity (Rabbani *et al.*, 2016, Johansen *et al.*, 2006). In CRP, the potential glycation sites were all located close to carboxylate groups within the primary and/or tertiary structure, principally situated C-terminally to the putative glycation site. For the most reactive lysine residues that have been detected after 19h of incubation (i.e. K13, K57 and K69), the distance between the  $\epsilon$ -NH<sup>+</sup> group of the glycated residue to the nearest carboxylate group was below 3.6 Å. These residues, K13, K57 and K69, all share the property of having an acidic residue C-terminally namely; E14, D60 and D70, respectively. A similar finding has been previously observed with other glycated proteins (Shapiro *et al.*, 1980, Casey *et al.*, 1995, Watkins *et al.*, 1985, Johansen *et al.*, 2006). The proximity of carboxylate group to lysine in either the primary or tertiary structure catalyses the Amadori rearrangement (Casey *et al.*, 1995, Garlick & Mazer, 1983), the rate-limiting step in glycation, by providing a conjugate base to promote the removal of a proton from the Schiff base and hence accelerates the rate of glycation at these residues (Rabbani *et al.*, 2016). K201 is found glycated only after 6 days of incubation. The slower reactivity of K201, despite its proximity to E101, can be explained by the small FAS (0.232) of the side chain of this residue meaning less access to free glucose compared to the other glycated residues located on the surface of the protein.

The microenvironment of the potential glycation sites, pKa and the FAS have also been utilised to identify the possible glycation sites in the FWVDGKPRVRKSLKKGTVGAE peptide. This peptide includes 6 possible glycation sites (2 arginines and 4 lysines) adjacent to each other. The presence of arginine or another lysine near the lysine either in the primary or the tertiary structure has also been proposed to catalyse the glycation of the lysine (Chen *et al.*, 2012, Johansen *et al.*, 2006), suggesting that these glycation sites could catalyse each other during the glycation process (Wang *et al.*, 2013). Up to 2 glucose molecules could be condensed on this peptide after 6 days of incubation with glucose. The arginines (R116 and R118) of this peptide are the less probable glycation sites, because they have the highest pKa values and lowest FAS compared to the lysines on the same peptide (FWVDGKPRVRKSLKKGTVGAE). Among the 4 lysines residues, K114 has the highest probability of being modified, based on a favourable combination of pKa and FAS (10.38 and 0.832 respectively) compared to other lysines. D112 which is situated proximal to K114 in the primary sequence, can accept the proton from the Schiff base and accelerate the rate-limiting Amadori rearrangement step (Rabbani *et al.*, 2016). K114 is part of the putative C1q-binding site of CRP and plays a role in modulating the reactivity of CRP with C1q (Agrawal *et al.*, 2001, Agrawal & Volanakis, 1994). Thus, K114 glycation could have an impact on the CRP function. For the remaining lysines residues, K119, K122 and K123, the combination of pKa and FAS were either both low or both high. For example, K119 has the lowest pKa value (9.61) but the smallest FAS (0.24) compared with the other lysines in the same sequence, meaning it is less exposed to glucose despite its high reactivity. Moreover, K119 and K123 are implicated in the intersubunit interaction (Thompson *et al.*, 1999, Shrive *et al.*, 1996) which could partially shield the glycation of these lysines (Johansen *et al.*, 2006, Ito *et al.*, 2011).

### 3.4.5 Conclusion

MS/MS spectra of glycated CRP digest recorded by the Q-TOF shows multiple neutral losses from the glucose-containing parent and fragment ions at a characteristic intensity distribution; this characteristic fragmentation pattern was independent of the peptide sequence and could be applied to confirm the occurrence of glycation as well as to localise the glycation site. Six glycation sites have been identified in the CRP digest treated with glucose for a period up to 6 days. These are K13, K57, K69, K191, K23 and K201, in addition to two ambiguous glycation sites among six possible residues within one peptide (FWVDGKPRVRKSLKKGTVGAE). Based on the combined contribution of pKa and FAS to the lysine microenvironment, K114 could be an additional glycation site. Among the identified residues, K13 is the more reactive lysine. The common feature between the glycated lysines are their proximity to an acidic residue which, principally situated C-terminally to the glycated lysine. The established protocols/procedures and the results obtained from this study are a good starting point to investigate the presence of these glycation sites in human plasma from diabetic patients



## Chapter 4: Analysis of human serum samples from diabetic patients

### 4.1 Introduction

#### 4.1.1 Overview

The high levels of circulating blood sugar found in diabetic patients can cause protein glycation. While we have established that CRP can be glycated *in vitro* as illustrated in chapter two, and the putative glycation sites were identified (chapter three), the effect of the glycation reaction on CRP, *in vivo*, remains unexplored. Glycation of proteins impacts on their structure and consequently may alter the physical and functional properties of the protein which may impact physiologically. In addition, recent studies showed that monomeric isoform of CRP (mCRP), formed as a result of a dissociation process of CRP, can play a harmful pro-inflammatory role in the progression of diseases (Wang *et al.*, 2015). There is however a consensus that this isoform is preferentially expressed in tissues (Eisenhardt *et al.*, 2009, Ji *et al.*, 2009). The aim of this study is first seeking, then quantifying these modified forms of CRP (glycated and/or monomeric) in the serum of diabetes mellitus patients with elevated CRP levels. The serum samples utilised in our study were obtained from 24 diabetic patients with either type I (n=4) or type II (n=20) diabetes, and markedly raised levels of CRP (>100 mg/L). CRP was isolated from human serum by affinity chromatography. The glycated form and the site(s) of glycation were identified using our established protocols/procedures by mass spectrometry. The monomeric form of the protein was purified from serum then characterised based on size exclusion chromatography, ELISA and mass spectrometry. Any emergent relationship between the clinical condition and modified and clinically assayed CRP levels was investigated.

#### 4.1.2 Diabetes mellitus

Diabetes mellitus (DM) is a metabolic disorder characterised by chronic hyperglycaemia with disturbance of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action or both (American Diabetes Association, 2014). DM is one of the most serious and chronic diseases in the world, currently affecting ~8.5% of the adult population (World Health Organization, 2016). The vast majority of diabetes falls into two main types; type 1 DM, is usually caused by autoimmune destruction of the pancreatic islet  $\beta$ -cells, rendering the pancreas unable to synthesise and secrete insulin (Castano & Eisenbarth, 1990), while type 2 DM, is caused by a combination of insulin resistance and inadequate insulin secretion and is the most common form, accounting for 90% of diabetes cases (Sacks & McDonald, 1996). In both types of diabetes, hyperglycaemia can lead to severe microvascular complications such as retinopathy, neuropathy, and diabetic nephropathy. However, the major complication in diabetes is macrovascular complications, mostly atherosclerosis, that are associated with increased risk of developing CVD such as coronary artery disease, stroke, and peripheral artery disease (Fowler, 2008, Messier *et al.*, 2004). In fact, people with DM are at two-to three-fold increased risk of developing CVD, and this represents the major cause of mortality in patients with diabetes (Jousilahti *et al.*, 2011, Sarwar *et al.*, 2010). Protein glycation and formation of AGEs, as a result of persistent hyperglycaemia, plays an important role in the development of diabetic complications (Singh *et al.*, 2014). The levels of glycated proteins in diabetic patients are much higher than those of non-diabetic individuals (Rondeau & Bourdon, 2011). Although both types of DM could enhance protein glycation, the rate of increase may be different. Yoshiuchi *et al.* (2008) demonstrated that the higher level of glycaemic fluctuation and deviations seen in type 1 DM could be responsible for the increase in protein glycation relative to type 2 DM

(Yoshiuchi *et al.*, 2008). On the other hand, anti-type 2 diabetes drugs such as metformin and pioglitazone have been found to inhibit the non-enzymatic glycation of proteins, possibly by trapping highly reactive intermediates generated during glycation (Rahbar *et al.*, 2000). The consequences of glycation vary, depending upon the product formed and the molecular structures involved. In general, it could be said that glycation compromises normal molecular structure, function and/or half-life. Add to this the fact that glycated structures are also targets of specific clearance receptors whose ligation promotes inflammation, then one may quickly conclude that glycation can have considerable effects (Brownlee, 2000, Cloos & Christgau, 2002, Stern *et al.*, 2002). The effect of glycation has been described in many proteins such as haemoglobin, albumin, ferritin, apolipoprotein, collagen, insulin, histone, IgM, IgG, glucose-6-phosphate dehydrogenase, aldehyde reductase, glutathione reductase, Cu-Zn superoxide dismutase and the lens protein crystallin (Arasteh *et al.*, 2014, Ansari & Ali, 2011, Hunter *et al.*, 2003). To the best of our knowledge, there is no work described in the literature studying the effect of non-enzymatic glycation of CRP *in vitro* or *in vivo*. The results of our *in vitro* study show that CRP can be glycated to various extents, depending on both glucose concentration and exposure time; putative glycation sites were determined.

#### 4.1.3 Monomeric C-reactive protein within the human body

Although pentameric CRP is extremely stable under physiological conditions, there is increasing evidence to suggest a pro-inflammatory role for mCRP (Eisenhardt *et al.*, 2009, Ji *et al.*, 2009). *In vitro*, native pCRP can dissociate irreversibly upon treatment with acid, urea-chelation, or heating in the absence of calcium into mCRP (Potempa *et al.*, 1987). mCRP displays new epitopes (neoepitopes) which are normally concealed in pCRP, and has distinct physiochemical, antigenic and biological activities (Ablij & Meinders, 2002,

Potempa *et al.*, 1987). Recently, strong evidence has been presented that pCRP dissociation can take place under physiological conditions after binding to the cell membrane (Ji *et al.*, 2007, Eisenhardt *et al.*, 2009). Conversely, other studies report independent mCRP synthesis upon the cell surface (Salazar *et al.*, 2014). It has also been proposed that mCRP imposes pro-inflammatory effects throughout the human body, contributing to various auto-inflammatory diseases. The binding of mCRP to either native LDL or oxidised LDL (ox-LDL) is an effective method of activating the classical complement pathway, causing an intensified state of inflammation at localised sites throughout the body (Ji *et al.*, 2009). Endothelial cell activation and dysfunction have also been linked with mCRP. Incubation of human vascular endothelial cells with mCRP resulted in an increase in the secretion of monocyte chemoattractant protein (MCP-1), IL-8, and increased expression levels of the cell adhesion molecules ICAM-1, VCAM-1 and E-selectin (Khreiss *et al.*, 2004). Up-regulation of P-selectin by mCRP in a study carried out by Molins *et al.* 2008, demonstrated the ability to further increase and stabilise platelet-platelet and platelet-leukocyte interactions (Molins *et al.*, 2008). Coupled with increased tissue factor expression and alteration of fibrin formation and clotting properties, mCRP may be a contributing factor towards the prothrombic events routinely seen in cardiovascular disease (Molins *et al.*, 2008, Li *et al.*, 2012). In addition to the activation of endothelial cells, mCRP has also been linked to the activation of circulating platelets within atherosclerotic lesions (Eisenhardt *et al.*, 2009). Lipids present on the surface of activated platelets have been shown to mediate the binding and dissociation of pCRP into mCRP (Eisenhardt *et al.*, 2009). Proposed modes of action have been put forward, hypothesising a dissociation mechanism and explaining how mCRP may contribute to disease throughout the circulation, but there is still no clear biological role for a monomeric form of CRP (Eisenhardt *et al.*, 2009, Verma *et al.*, 2004).

Upon activation, lipid membranes undergo rapid changes in their membrane composition, resulting in the shedding of microparticles (MPs); (Owens & Mackman, 2011). Recent evidence has demonstrated that these MPs, following severe cardiovascular events, can bind and dissociate pCRP (Habersberger *et al.*, 2012). This evidence provides an *in vivo* explanation for the generation and transportation of mCRP, which can thus interact with other cells in the body and provide a pro-inflammatory stimulus (Habersberger *et al.*, 2012). Although it is not entirely clear what is the role of mCRP within the human body, experimental evidence provided in the literature indicates a potential pro-inflammatory role. Furthermore, this lack of understanding about mCRP has resulted in an incoherent generalisation throughout the literature of whether mCRP is primarily tissue bound or can be located within human serum. A recent pilot study carried out by Williams (2017) established that mCRP is located in human serum when circulating levels of CRP are high. In this study, the mCRP levels were observed to vary significantly, with some clear high and low outliers from the general distribution, between patients where the underlying aetiologies were unknown (Williams, 2017). In the current work, our aim is to investigate the presence of mCRP in diabetic patients who already have elevated CRP levels, not only to confirm or refute the previous finding of mCRP in patients with high clinically measured CRP levels, but also to establish where the mCRP levels for DM lie in relation to these. Another reason for selecting DM patients with elevated level of CRP is because the *in vitro* glycation trial of CRP shows that the stoichiometry of the glycation reaction is likely to be low, with difficulties associated with detecting a low abundance modification in the presence of other proteins in the serum.

#### 4.1.4 Experimental aims

While we have established that CRP can be glycosylated *in vitro*, the effect of the glycation reaction on CRP, *in vivo*, remains unexplored. Furthermore, recent evidence suggests that mCRP may be present in serum. The dissociation of pCRP to mCRP has been linked with proinflammatory processes and the progression of several diseases while the role and outcomes of CRP glycation remain unknown. The experimental aim of the research described in this chapter is to seek, then quantify, modified CRP in the serum of diabetes patients with elevated CRP levels and assess the relationship between pCRP, mCRP and the glycosylated isoforms. CRP was separated from human serum using affinity chromatography. The glycosylated form and the site(s) of glycation were detected using the established methodology that developed from our previous experimental research (chapter two and chapter three). The monomeric form of the protein was purified from serum then characterised based on chromatography, ELISA and western blot based on established procedures developed previously in the group (Williams, 2017).

## 4.2 Materials & Methods

### 4.2.1 Reagents and consumables

Reagents, solutions and proteins used in this chapter and their suppliers are summarised in Table 4-1.

Table 4-1 Reagents, solutions and proteins used in the study. All substances were purchased in the highest possible purity from the following supplier

Chemicals/proteins	Supplier
<b>3,3',5,5' tetramethylbenzidine (TMB)</b>	Sigma-Aldrich
<b>Acetonitrile (HPLC grade, 99.9%)</b>	VWR International
<b>Ammonium bicarbonate</b>	Sigma-Aldrich
<b>Ammonium persulfate (APS)</b>	VWR International
<b>Anti-mouse IgG (whole molecule) peroxidase Antibody</b>	Sigma-Aldrich
<b>bovine serum albumin (BSA)</b>	Sigma-Aldrich
<b>Bromophenol blue</b>	VWR International
<b>Calcium chloride</b>	Sigma-Aldrich
<b>C-reactive protein (CRP, &gt;99% pure)</b>	SCRIPPS Laboratories
<b>Dithiothreitol</b>	Sigma-Aldrich
<b>Ethylenediaminetetraacetic acid (EDTA)</b>	Sigma-Aldrich
<b>Formic acid (96%)</b>	Sigma-Aldrich
<b>Glu-C (Protease V8 Sequencing Grade)</b>	Roche
<b>Glycerol</b>	Sigma-Aldrich
<b>Glycine</b>	Sigma-Aldrich
<b>InstantBlue™ Protein Stain</b>	Expedeon
<b>Iodoacetamide</b>	Sigma-Aldrich
<b>Low range molecular weight markers</b>	Bio-Rad
<b>Monoclonal anti-C-reactive protein antibody (Clone-8)</b>	Sigma-Aldrich
<b>Potassium chloride</b>	Sigma-Aldrich
<b>Rapigest</b>	Waters
<b>Sodium azide</b>	Sigma-Aldrich
<b>Sodium chloride</b>	Sigma-Aldrich
<b>Sodium dodecyl sulfate SDS</b>	Sigma-Aldrich
<b>Sodium phosphate dibasic (99%)</b>	Sigma-Aldrich
<b>Sodium phosphate monobasic (99%)</b>	Sigma-Aldrich
<b>Tetramethylethylenediamine (TEMED)</b>	Sigma-Aldrich
<b>Trifluoroacetic acid (98%)</b>	Sigma-Aldrich
<b>Tris</b>	Fisher Scientific
<b>Trypsin (Proteomics Grade)</b>	Sigma-Aldrich
<b>Tween</b>	Sigma-Aldrich
<b>Water (HPLC grade)</b>	VWR International

#### 4.2.2 Ethical Application

Before starting the collection of human serum samples or any experimental procedure to take place, approvals from independent peer review and research ethics committee were required.

A peer review form was submitted to the Independent Peer Review Committee (IPRC) at Keele University, the application was approved in October 2016. Following IPRC approval, an application to the Integrated Research Application System (IRAS) was submitted. IRAS is an online system for applying the permissions and approvals to Research Ethics Committees (REC) and the NHS Research & Development offices. This application was formally approved in April 2017 by NRES Committee West Midlands – Edgbaston under REC reference 17/WM/0150; Protocol reference: RG-0091-16-FNS and IRAS reference 211529, see appendix 1.

#### 4.2.3 Recruitment of Participants

The identification of participants took place at the University Hospital of North Midlands (UHNM) considering the following inclusion criteria:

- I. Male or Female above the age of 18 years;
- II. They have type 1 or type 2 diabetes mellitus;
- III. They are a case with a raised CRP level of >100 mg/mL;
- IV. They are inpatients at University Hospitals of North Midlands;
- V. They are due to undergo routine blood testing as a part of their standard clinical care, and enough serum is left over for our study;
- VI. They are able to give informed consent.

No exclusion criteria were applied to the participants beyond the inclusion criteria.



All the recruited participants were inpatients of UHNM, being investigated for a range of clinical situations requiring CRP monitoring as part of their routine patient care. Eligible participants were identified by a research nurse or research team member in UHNM. Research staff approached eligible participants and gave a brief verbal explanation of the study. All eligible participants who showed an interest in our study were then provided with a patient information sheet (PIS, see Appendix 2). The research nurse guided the participants through PIS, outlined the aims, benefits, and risks of the study and answered any questions the participant may have. The participants were informed that they were due to undergo a routine blood test as a part of their standard clinical care, and the leftover sample, which is routinely discarded, would be utilised for our research. The research nurse further explained that no additional procedures, not part of the routine clinical care, were required and the taken samples were to be anonymised; therefore, none of them would be identifiable. The participants were given as much time as necessary to decide whether they wished to take part in our study. If the participant agreed to take part in our study, they were given a participant consent form (PCF, see appendix 3), which they were asked to read through carefully, sign and date. The participant consent forms were stored under the standard NHS procedure. Access to PIS and PCF was only provided to NHS members of the research team or the participants' clinical care team at the hospital. Participants were informed that they have the right to pull out of the study at any given time, without reason, with their sample and consent form being destroyed immediately. Once sample anonymisation and analysis had taken place this would no longer be possible as the patient identifiers were unknown to the researcher.

#### 4.2.4 Sample size calculation

The sample size calculation was based upon an estimated percentage of mCRP that we anticipate to be found in the serum of type 1 & 2 diabetic patients with raised levels of CRP (>100 mg/L). To our knowledge there is no published experimental data in either of these fields, meaning we are unable to perform a standard sample size calculation. Based on recent work by Williams, who detected mCRP in the serum of different patients with a raised level of CRP (Williams, 2017), we estimated that the proportion of diabetic patients with detectable levels of mCRP would be as high as 75%. From this, we were then able to calculate the margin of error (ME) and the confidence interval at 95% confidence level for a chosen sample size, these were calculated with the following equations:

Equation 4-1 The equations used to determine the margin of error ME (A) and the confidence interval (B) for a sample size.

<b>A)</b>	$ME = Z * \frac{\sqrt{\hat{p}(1-\hat{p})}}{n}$
<b>B)</b>	$\text{Confidence interval} = (\hat{p} - ME, \hat{p} + ME)$

where  $\hat{p}$  is the expected proportion of samples with a detectable level of mCRP,  $n$  is the sample size and  $Z$  is the critical value at the 95% confidence level which is equal 1.96.

It was calculated that with a sample size of 40 (20 of type 1 and 20 of type 2 diabetes) would be associated with a precision of  $\pm 13\%$ , meaning that between 62% to 88% (25-35 samples) of the 40 serum samples may contain mCRP at 95% confidence level.

#### 4.2.5 Acquisition of human serum samples

After obtaining a signed and dated PCF, a qualified phlebotomist at UHNM drew venous blood from the recruited participant according to the hospital care protocol. A standard phlebotomy procedure was followed. The blood was collected in red cap BD Vacutainer serum separator gel tubes without anticoagulant. The tubes were left for up to 30 minutes, to allow the blood to coagulate, and then centrifuged at 3,500 x g for 10 minutes at room temperature to separate the serum from other blood components. The serum was then collected in a tube and stored at 4°C in the hospital cold room until the routine analysis had been carried out. Once the analysis had been completed, the volume remaining of the serum (near 1 mL) was transferred into a 1.5 mL cryovial tube and anonymised via a unique sample ID number (e.g., CRP01). The anonymised samples were then returned to in the cold room at 4°C until collection. Upon collection, the CRP level and where possible, calcium levels, were provided, via an Excel spreadsheet, linked to the unique sample ID number. The samples were transferred to Keele University, Life Sciences, Huxley building in a chilled, sealed container. Upon arrival, 0.02% (v/v) of sodium azide was added to all samples, then stored at 4 °C in the cold room (HTA approved) of the Structural Biology Research Lab at Keele University until analysis. All data generated from the study were stored securely on a password protected computer, linked to the Keele University server and managed by the CI. A detailed record of the samples involving: the sample ID, collection date, sample volumes, CRP and calcium concentrations and dates of when samples were discarded, were kept throughout the study.

#### 4.2.6 Production of monomeric CRP

Monomeric CRP, mCRP, was prepared *in vitro* from commercially available CRP to be used as a standard control in our study. Human native CRP (purity>99%), purchased from SCRIPPS Laboratories (San Diego, California, USA), was used to prepare mCRP by urea and chelator treatment. Native CRP at final concentration of 0.63 mg/mL was incubated in dissociation buffer containing 3 M urea, 10 mM EDTA, 20 mM Tris and 280 mM NaCl. Sodium azide was added at a final concentration of 0.01% to prevent bacterial and fungal contamination during the dissociation process. The resulting solution was incubated in the cold room for 10 weeks. After 10 weeks had been passed, the dissociated protein was subjected to size exclusion chromatography (see section 4.2.10) to separate mCRP. The eluted size exclusion fractions corresponding to mCRP were concentrated using a Vivaspin™ Turbo 15 mL centrifugal concentrator (Fisher Scientific), 10 kDa molecular weight cut-off (MWCO). The centrifugal concentrator was washed with deionised water (DW) at 1,000 x g, 4°C for 10 minutes. After the washing step, the eluted mCRP was then concentrated at 1,000 x g, 4°C for 15 minutes or until the final volume of 0.5 mL had been achieved. The protein concentration was determined by UV absorbance at 280 nm (Nanodrop 1000, Thermo Scientific). An extinction coefficient of 1.7 was used to give the concentration values in mg/mL (De Beer & Pepys, 1982).

#### 4.2.7 Calibration of the size exclusion column

Size exclusion chromatography was used to separate the mCRP produced from dissociation of native pCRP *in vitro*, and to purify and separate both protein forms (mCRP and pCRP) from human serum samples. It was necessary to calibrate the size exclusion column before running the samples in order to accurately determine protein sizes via the reported elution volumes. HiLoad 16/60 Superdex 200pg column connected to the AKTA explorer 100 fast

protein liquid chromatography (FPLC) system (GE Healthcare) was used in this study. The system was controlled by UNICORN 5.11 software (GE Healthcare), and the elution volumes of the proteins were monitored by measuring the UV absorbance at 280 nm.

Before calibration, the tubing, pumps and all valves of the system were purged with 30 mL of DW at 4 mL/min flow rate and bypassing the column. The column was stored in 20% ethanol before use, three column volumes (120 mL per column volume) DW were used to purge the ethanol storage buffer from the column at 1.5 mL/min flow rate. The column was then equilibrated using three column volumes of elution buffer (50 mM Tris, 150 mM NaCl, 0.01% sodium azide). The elution buffer and DW used on the column were pre-filtered through a 0.2 µm cellulose acetate membrane filter (Sartorius Stedim Biotech, Goettingen, Germany) to remove particulate matter that may clog the system, and de-gassed to avoid introducing air into the column.

The calibration was performed using a Gel Filtration Low Molecular Weight (LMW) Calibration Kit (GEHealthcare). The LMW kit contains 5 proteins with molecular weights in the range 6.5 kDa to 75 kDa, and Blue Dextran 2000 to determine the void volume of the column (Table 4.2). To further expand the molecular weight range of the kit, CRP (99% pure) purchased from SCRIPPS Laboratories (San Diego, California, USA) was used with the calibration standards. The calibration standards were provided in a lyophilised form and reconstituted to a 20 mg/mL stock solution using buffers recommended by the supplier.

Two separate chromatography runs were performed, the first for the calibration standards including the CRP, and the second run for the Blue Dextran 2000 sample. Blue Dextran 2000 was run separately because the fraction of Blue Dextran is broad and may overlap the calibration standard peaks. To achieve the best resolution, the volume of the sample

loaded on the column needs to be in the range of 0.1 to 2% of the total column volume. Therefore, the final volume of the sample solution was 1mL. The proper volumes of calibration standards and CRP were mixed together to achieve the concentration of the proteins listed in table 4.2, and the final volume was adjusted to 1mL with the elution buffer. For the Blue Dextran 2000 sample, a fresh solution containing 1 mg/mL was prepared, and the final volume was adjusted to 1mL using the elution buffer.

Table 4-2 A table of the calibration standards and their respective molecular weights. The calibration standards were used to determine the calibration graph for the HiLoad 16/60 Superdex 200 pg size exclusion chromatography column. The final concentrations of the calibration standards that were used in the calibration process are also provided.

Protein Standard	Molecular Weight (Da)	Final concentration (mg/mL)
Aprotinin	6500	3
Ribonuclease A	13700	3
Carbonic anhydrase	29000	3
Ovalbumin	44000	4
Conalbumin	75000	3
CRP	115000	0.1

The prepared sample solution (calibration standards or Blue Dextran sample) was loaded onto the column via a 2 mL sample injection loop while the elution buffer passed through the column at low speed. A buffer with a pH of 6–8 and an ionic strength  $\geq 0.15$  M is recommended for elution. The proteins were eluted with one column volume of elution buffer (150 mM NaCl, 50 mM Tris, and 0.01% sodium azide, adjusted to pH 7.20) at a flow rate of 1 mL/min and elution volumes of the proteins were monitored by measuring the UV absorbance at 280 nm.

#### 4.2.8 Purification of the human serum samples by affinity chromatography

CRP was purified from human serum samples by using a phosphocholine-bound affinity column (5 mL immobilised p-aminophenyl phosphocholine agarose gel) connected to ÄKTAexplorer 100 (GE Healthcare). The software used to operate the system was UNICORN 5.11 (GE Healthcare). The buffers, solutions and DW used on the column were all filtered through a 0.2 µm cellulose acetate filter (Sartorius Stedim Biotech, Goettingen, Germany) to eliminate particulate matter that may clog the system, and de-gassed to avoid introducing air into the column.

Before running the samples, the tubing, pumps and all valves of the system were first purged with 30 mL of DW at flow rate 4 mL/min, bypassing the column. The column was then re-connected and purged with 4 column volume (5 mL/ column volume) of DW at a flow rate of 1.5 mL/min to remove the equilibration buffer that filled the column during storage. 10 column volumes of regeneration buffers 1 & 2 (see table 4-3) were passed through the column at 1.5 mL/min flow rate, to clean the column from protein remnants that may present from the previous run, and to enable CRP present in the serum to bind. After regenerating the column, the column was equilibrated with 20 column volumes of equilibration buffer (see table 4-3) at a flow rate of 1.5 mL/min. After 20 column volumes had been passed, 0.9 mL of the serum sample was injected onto the column via a 2 mL sample injection loop while the equilibration buffer passed through the column at very low speed (0.1 mL/min). When the entire sample had been loaded to the column, the flow rate was adjusted to 0.5 mL/min, and the UV absorbance reading was blanked to a value of 0 absorbance units. The column was washed with the equilibration buffer until all other unnecessary proteins pass through the column as a peak rise and flat-lined to zero again. When the UV absorbance reading had returned to its baseline value (usually after 15 mL),

the equilibration buffer was replaced with a chelating buffer (table 4-3) to elute the CRP that bound to the column. One column volume of chelating buffer was passed through the column at a flow rate of 0.5 mL/min with the absorbance measured at 280 nm. The fractions that correspond to eluted CRP were collected in 5 mL tubes. Once all CRP had been eluted, the column was washed with 20 column volumes of equilibration to be prepared for storage.

Table 4-3 A table of the components of the buffers required to purify CRP from human serum samples using the phosphocholine-bound affinity chromatography column. Concentrations of each component used are also given

Buffer (pH 7.4)	Components
Regeneration buffer 1	<ul style="list-style-type: none"> <li>• 0.5 M sodium chloride</li> <li>• 0.2 M tris</li> <li>• 0.01 M EDTA</li> </ul>
Regeneration buffer 2	<ul style="list-style-type: none"> <li>• 0.5 M sodium chloride</li> <li>• 0.05 M tris</li> <li>• 0.01 M calcium chloride</li> </ul>
Equilibration buffer	<ul style="list-style-type: none"> <li>• 0.01 M calcium chloride</li> <li>• 0.05 M tris</li> <li>• 0.15 M sodium chloride</li> <li>• 0.01% sodium azide</li> </ul>
Chelating buffer	<ul style="list-style-type: none"> <li>• 0.01 M EDTA</li> <li>• 0.05 M tris</li> <li>• 0.150 M sodium chloride</li> <li>• 0.01% sodium azide</li> </ul>

#### 4.2.9 Concentrating the eluted fractions

The eluted affinity chromatography fractions that correspond to CRP were concentrated using a Vivaspin™ Turbo 15 mL Centrifugal Concentrators (Fisher Scientific), 10 kDa MWCO. The centrifugal concentrator was initially washed with DW at 1,000 x g, 4 °C for 10 minutes.



After the washing step, the protein sample was then concentrated down at 1,000 x g, 4 °C for 15 minutes or until the desired volume of less than 1 mL had been achieved.

#### 4.2.10 Separation of pentameric and monomeric CRP by size exclusion chromatography

Size exclusion chromatography was used to separate and purify the *in vitro* produced mCRP, and the pCRP and the potential mCRP from the human serum samples based on their elution volumes. The system, column and software used were the same as described in section 4.2.7, and the size exclusion column had been calibrated before running the samples in order to accurately determine protein sizes via their reported elution volumes. The elution volume of all the samples was determined by measuring the absorbance at 280 nm.

Before running the samples, the tubing, pumps and all valves of the system were purged with 30 mL of DW at 4 mL/min flow rate and bypassing the column. The column was stored in 20% ethanol, three column volumes (120 mL per column volume) DW were used to purge the ethanol storage buffer from the column at 1.5 mL/min flow rate. The column was then equilibrated using three column volumes of elution buffer containing 280 mM NaCl, 20 mM tris, 5 mM CaCl<sub>2</sub>, 0.01% sodium azide. The pH of the elution buffer was adjusted to pH 7.4 by the addition of HCl. The buffer and DW used on the column were pre-filtered through a 0.2 µm cellulose acetate filter (Sartorius Stedim Biotech, Goettingen, Germany) to remove particulate matter that may clog the system, and de-gassed to avoid introducing air into the column.

The samples subjected to size exclusion column were all the human serum samples that had been purified by affinity chromatography and concentrated to a volume of less than 1 mL, and the *in vitro* produced mCRP dissociation sample which was in a final volume of 1

mL. The volume of samples was within 0.1-2% of the column volume (120 mL) which is recommended to achieve the best resolution.

The sample was loaded onto the column (HiLoad 16/60 Superdex 200 pg column (GE Healthcare, 2011)) via a 2 mL sample injection loop while the elution buffer passes through the column at deficient speed. The proteins were eluted with one column volume of elution buffer at a flow rate of 0.5 mL/min while the absorbance was measured at 280 nm. The eluted proteins were collected as 5 mL fractions in 20 mL universal tubes. The pCRP containing fractions and the fractions thought to contain mCRP were concentrated, as described in section 4.2.9, to give final volumes ranging from 0.4 to 1.1 mL. The variation in the volumes of the purified samples was considered when determining the protein concentration of the sample. After the proteins were eluted, the size exclusion column was washed with 3 column volumes DW, followed by 3 column volumes of 20% ethanol to prepare the column for storage.

#### 4.2.11 SDS-PAGE of the size exclusion eluted fractions

SDS-PAGE was performed to evaluate the purity and the protein content of fractions obtained from size exclusion chromatography. Size exclusion eluted fractions containing pCRP from the human serum samples labelled CRP01, 02, 06, 10, 11, 13, 14, 15, 17, 18, 21, 22, 23 and 24 were selected and subjected to SDS-PAGE according to the previously described procedure (section 2.2.3). The gel was cast as described in section 2.2.3 with acrylamide content of 4% and 15% (v/v) in the stacking and resolving gels, respectively. The samples were mixed 1:1 (v/v) with non-reducing loading buffer and heated at 95 °C for 5 minutes before being loading to the gel. A total volume of 20 µL was loaded to each well to maximise the amount of the protein. Low range molecular weight markers were run

beside the samples. Electrophoresis was run at 200 V for 45 min, and the samples were visualised by Coomassie Instant Blue stain.

#### 4.2.12 Enzyme-linked Immunosorbent assay of the eluted serum fractions

Enzyme-linked immunosorbent assay (ELISA) was used to detect and quantify both pCRP and mCRP isoform in human serum samples. pCRP containing fractions and fractions thought to contain mCRP obtained from size exclusion chromatography (section 4.2.10) were tested separately. All ELISA analysis was carried out using 96-well flat-bottom plates with a low evaporation lid. A monoclonal anti-C-reactive protein antibody (clone-8), produced in a mouse, was used to detect both CRP isoforms. Anti-mouse IgG (whole molecule)-peroxidase antibody, was used as a secondary antibody. Both primary and secondary antibody was diluted as recommended by the supplier. The buffers and solutions used for ELISA experiments are detailed in Table 4-4.

Table 4-4 Table of the components of the buffers and solution used in the ELISA experiments for the pCRP and mCRP containing fractions. The components concentrations are also given.

Buffer/Solution	Components	
<b>PBS pH 7.4 with HCl</b>	Sodium chloride	137 mM
	Potassium chloride	2.7 mM
	Disodium phosphate	10 mM
	Monopotassium phosphate	2 mM
<b>Wash buffer (PBS – Tween)</b>	PBS	1 L
	Tween	0.5 mL
<b>Blocking solution (PBS-Tween with 3% BSA)</b>	PBS – tween	20 mL
	BSA	0.6 mg
<b>Calcium buffer pH 7.4 with HCl</b>	Tris	20 mM
	Sodium chloride	280 mM
	Calcium chloride	5 mM
<b>Primary monoclonal antibody (Dilution 1:40,000)</b>	Primary antibody	1 µL
	PBS–tween	up to 40 mL
<b>Secondary antibody (Dilution 1:40,000)</b>	Secondary antibody	1 µL
	PBS–tween	up to 40 mL

For the pCRP containing fractions ELISA assay, native pCRP from SCRIPPS was used to prepare the calibration standards. An aliquot was diluted to a concentration of 0.2 mg/L (20 ng in 100  $\mu$ l). Subsequently, a serial dilution was carried out (eight dilutions) using calcium buffer (280 mM NaCl, 20 mM Tris, 5 mM CaCl<sub>2</sub>, 0.01% sodium azide), with the last dilution include only the calcium buffer (negative control). 100  $\mu$ L of each dilution was added in triplicate to the columns 1 to 3, rows A-H (see table 4-5). Each pCRP containing fractions was first diluted in calcium buffer at 1:1000, and 100  $\mu$ L of each serum sample was added, in triplicate, to a 96-well plate, see table 4-5.

Table 4-5 The amounts of pCRP (ng) from SCRIPPS used in the preparation of the ELISA calibration standards. The dilution and the volumes of the pCRP containing samples added to the wells are also given

	1	2	3	4	5	6	7	8	9	10	11	12
<b>A</b>	20	20	20	Fractions thought to contain pCRP from human serum samples CRP01 to CRP24 were diluted at 1:1000 and 100 $\mu$ L were added in triplicate to the columns 4-12, row A-H.								
<b>B</b>	10	10	10									
<b>C</b>	6	6	6									
<b>D</b>	3.6	3.6	3.6									
<b>E</b>	2.16	2.16	2.16									
<b>F</b>	1.296	1.296	1.296									
<b>G</b>	0.778	0.778	0.778									
<b>H</b>	0	0	0									

For the mCRP containing fractions ELISA assay, the mCRP dissociated by 3 M urea was used to prepare the calibration standards. An aliquot was diluted to a concentration of 0.125 mg/L (12.5 ng in 100  $\mu$ l). Subsequently, a serial dilution was carried out (seven dilutions) using calcium buffer, with the last dilution includes only the calcium buffer (negative control). 100  $\mu$ L of each dilution was added, in triplicate, to the columns 1 to 3, rows A-G

(see table 4-6). 100  $\mu$ L of each mCRP containing samples was added without dilution, in triplicate, to a 96-well plate, see table 4-6.

Table 4-6 The amounts of mCRP (ng), from 3 M urea dissociation experiment, used in the preparation of the ELISA calibration standards. The dilution and the volumes of the mCRP containing samples added to the wells are also given.

	1	2	3	4	5	6	7	8	9	10	11	12
A	12.5	12.5	12.5	Fractions thought to contain mCRP from human serum samples CRP01 to CRP24 were added in triplicate to the columns 4-12, row A-H.								
B	6.25	6.25	6.25									
C	3.125	3.125	3.125									
D	1.563	1.563	1.563									
E	0.781	0.781	0.781									
F	0.391	0.391	0.391									
G	0	0	0									
H	0	0	0									

Both plates were left to incubate overnight at 4°C on the rocker. After incubation, the following procedure was used for both plates. The plates were washed thoroughly for three cycles with washing buffer of phosphate buffered saline with tween, PBS-T (see Table 4-4). Each cycle the washing buffer was left in the well for 2-3 min then discarded by blotting the plate onto absorbent material. The plates were then blocked with 300  $\mu$ L of a freshly prepared blocking solution (PBS-T supplemented with 3% of bovine serum albumin) for 2 hours at 4°C on the rocker. The plates were washed three times with PBS-T, and each time the washing solution was left in the well for 2-3 min before being discarded. The primary monoclonal antibody solution was thawed before use. The primary monoclonal anti-C-reactive protein antibody (clone-8) solution was diluted at 1:40,000 in wash buffer and 100  $\mu$ L of the diluted antibody solution was added to each well. The plate was incubated for 5 hours at 4 °C on the rocker. The plates were washed three times with PBS-T, and each time

the washing solution was left in the well for 2-3 min before being discarded. The secondary antibody was thawed before use and diluted at 1:40,000 in wash buffer. 100  $\mu$ L of the diluted secondary antibody was added to each well then incubated for 2 hours at 4 °C on the rocking machine. After incubation, the plates were washed again as with the primary antibody. Afterwards, the developing substrate (3,3',5,5' tetramethylbenzidine (TMB)) was added to each well (100  $\mu$ L/well) and allowed to develop for 10 minutes at room temperature, or until the required colour intensity was reached. Finally, the reaction was stopped by adding 100  $\mu$ L/well of 2 M sulfuric acid; the optical density for each sample was immediately determined at 450 nm using a BioTek EL800 plate reader operated by Gen5 Software. The standard curves were then generated and used to determine the protein concentration in each sample, considering the variation in the initial sample volume.

#### 4.2.13 Intact mass spectrometry of the eluted fractions

Purified samples containing pCRP, obtained from size exclusion chromatography described in section 4.2.10, were subject to LC-MS analysis to test for protein glycation as a function of the net gain in the molecular weight of the protein as a result of glucose conjugation. A portion from each sample was buffer exchanged and further concentrated to about half of its volume using Amicon ultra centrifugal filter devices (Millipore) fitted with a 10 kDa molecular weight cut-off membrane. The Amicon devices were first washed with 500  $\mu$ L of DW at 1000 x g for 5 minutes. After the washing step, the sample was added, and the volume completed to 500  $\mu$ L with solvent A (0.1% F.A, 5% ACN) and then concentrated at 14000 x g for 15 minutes. This process (buffer exchanging and concentration) was repeated three times; the centrifugation steps were performed at 4 °C.

The LC-MS setup including the system, buffers, and the software is the same as described in section 2.2.5. For each HPLC run, 2  $\mu\text{L}$  of protein solution (containing an average of  $>5$  pmole of pCRP) was injected into the 20  $\mu\text{L}$  loop and loaded to the C4 guard column over 5 minutes using 95% solvent A, 5% solvent B (30  $\mu\text{L}/\text{min}$ ). After the trap/wash period, the sample was loaded onto the C4 analytical column at a split flow rate of 0.35  $\mu\text{L}/\text{min}$ . The gradient consisted of 2 ramps: 15 minutes (5% - 60% B), then, 5 minutes 60-95% B. The column was then washed at 95% B for 5 minutes before re-equilibration to the starting conditions. Sample runs were bracketed by blank runs to allow full equilibration of columns and to minimise system carryover and consequent column contamination. The HPLC column was coupled online to the Q-TOF mass spectrometer via nanoESI, with typical capillary voltages of 1-2 kV being applied. The Q-TOF was operated in MS only mode. The Q-TOF parameters were the same as described in section 2.2.5. The scan  $m/z$  range started at 500 and ended at 2500 in positive ion mode, scan time 1 s, 0.1 s interval. Data were collected and processed using MassLynx V4.1 (Waters). The multiply-charged electrospray spectra were deconvoluted (after background subtraction) using the Maximum Entropy 1 package of the MassLynx software, with a resolution of 1 Da/channel, a uniform Gaussian width at half height of 0.5 Da, and minimum intensity ratios of 33%.

In addition to pCRP samples, the mCRP containing samples with high protein content, as revealed by ELISA, were also analysed by LC-MS as an additional measure to confirm the identity and the mass of the isolated isoform. The mCRP containing fractions from human serum samples labelled CRP04, CRP17, and CRP21 were selected. 200  $\mu\text{L}$  from each selected sample was buffer exchanged into solvent A and concentrated down to 40  $\mu\text{L}$  using Amicon ultra centrifugal filter devices, then 5  $\mu\text{L}$  was injected into the system and analysed as with pCRP samples.

#### 4.2.14 Bottom-up mass spectrometric analysis of pCRP containing samples

Purified pCRP containing samples were digested with trypsin and Glu-C and subjected to bottom-up mass spectrometric analysis with the aim to identify glycation, if any, at the peptide level. Purified pCRP containing samples from all 24 human serum samples were treated, digested and analysed using the same reagents, protocol, and equipment described in section 3.2.2 and 3.2.3. The purified pCRP samples were in calcium buffer (see section 4.2.10) with an average protein content of 0.04  $\mu\text{g}/\mu\text{L}$  as determined by ELISA. An aliquot (200  $\mu\text{l}$ ) from each sample was concentrated, using Amicon ultra centrifugal filter devices, to a final volume of 50  $\mu\text{L}$  (assumed to contain an average of 8  $\mu\text{g}$  of protein). The concentrated pCRP samples were denatured with Rapigest, reduced with DDT, and alkylated by IAA as described in section 3.2.2. The treated samples were divided into two equal portions (each containing  $\approx 4$   $\mu\text{g}$  of the protein). The first portion was used for the trypsin digestion while the second portion was digested with endoproteinase Glu-C.

Both trypsin (proteomics grade, Sigma-Aldrich) and Glu-C (protease V8 sequencing grade, Roche) were freshly reconstituted, and working solution of 0.1  $\mu\text{g}/\mu\text{L}$  was prepared as described in section 3.2.2. Trypsin (at 0.1  $\mu\text{g}/\mu\text{L}$ ) was added to the first portion of each CRP sample to give substrate to enzyme ratios of 30 to 1, and an equal volume of Glu-C (at 0.1  $\mu\text{g}/\mu\text{L}$ ) was added but to the second portion to give the same substrate to enzyme ratio. The solutions were incubated for 24 hours at 37 °C with gentle agitation. A second application of each enzyme was added to the corresponding aliquot after 8 h of incubation, this to aid complete protein digestion. The digestion was stopped by lowering the pH to <2 by adding TFA to a final concentration of 0.5% (v/v). The acidified solutions ( $\approx 65$   $\mu\text{L}$  samples)



were then incubated at 37 °C for 30-45min before centrifugation and removal of the supernatant.

The CRP peptides digests were then desalted and concentrated using C18 Zip-tip (Millipore) and the same protocol described in section 3.2.3. The zip-tip was wetted ten times by 10ul portions of 100% ACN and then equilibrated ten times by a 10 µL portions 0.1% F.A solution. A 30 µL aliquot of CRP digest (assumed to contain  $\approx 1$  µg of the protein) was absorbed onto the reversed-phase medium by 10 repeated aspirate-dispense cycles of sample loading. The tip was washed immediately with 100 µL of 0.1% F.A. Finally, A 10 µL portion of 80% ACN in 0.1% F.A solution was placed into a microcentrifuge tube, and peptide elution was performed by passing this solution through the Zip-tips with 3 aspirate- dispense cycles. The desalted and concentrated protein digest (with an average concentration of 3 pmole/µL if assuming 80% recovery from the zip-tip process) were then transferred to autosampler vial inserts (Waters Corporation).

The vials were loaded onto HPLC autosampler, and an aliquot of solvent A was also loaded to be used as a blank. The samples run was performed after three blank runs to ensure that the LC-MS system is stable before samples analysis. Moreover, two blanks run was added between the samples as well to minimise system contamination and carryover. The LC-MS setup including the system, buffers, and the software is all the same as described in section 3.2.3. For each HPLC separation, digested protein (typically 2 µL) were loaded onto C18 trap column throughout 5 minutes using 95% solvent A, 5% solvent B at 30 µL/min to desalt and concentrate the peptides. After the trapping period, the peptides were separated on a C18 analytical column at a flow rate of 0.35 µL/min. The HPLC gradients consisted of 2 ramping steps; a 35 minutes ramp from 5% to 60% solvent B, followed by a 10 minutes ramp to 95%

solvent B then preserving on 95% B for additional 5 minutes before returning to the initial starting conditions at the end of the run. After separation, samples eluting from the HPLC column were delivered to a Q-TOF mass spectrometer and introduced into the gas phase through the nanoESI source by applying a DC potential of 1-2 kV to the sample solution within the Nano-spray emitter capillary. The Survey spectra were acquired in the positive ion mode over the mass range of 400–1600 amu, with data-dependent product ion analysis of multiply charged precursors as described in 3.2.3.

The data were collected using MassLynx version 4.1 (Waters Ltd., Manchester, UK). The collected data were extracted as text formatted mgf files using Mascot Distiller (Matrix Science, London). The mgf files were subjected to database searching using the Mascot search engine (Matrix Science, London, UK) as described in section 3.2.4. The MS/MS spectra of the  $m/z$  values assigned to glycated CRP peptides through Mascot, if present, were analysed manually to interpret unassigned peaks within the spectrum and confirm the glycation site. The  $m/z$  values that did not correspond to any peptide sequence in the selected database during the Mascot search were considered as possible CRP glycated peptides and further analysed as described in section 3.2.4.

#### 4.2.15 Statistical Analysis

All graphs and tables produced from the data collected from the chromatographic separation and ELISA quantification of CRP isoforms were created by Microsoft Excel 365 ProPlus. Statistical analysis including descriptive statistics and Mann-Whitney test was carried out in Minitab® Statistical Software (Version 18). Central tendencies are reported as mean  $\pm$ SEM for normal data and median + interquartile range (IQR) for non-normal data.

## 4.3 Results

### 4.3.1 Experimental design

The primary aim of the study was to establish whether modified forms of CRP (glycated and/or monomeric) are present in the serum of diabetic patients with raised levels of CRP (>100 mg/L). After the required approvals were in place, the participant recruitment process was started. All the participants were inpatients at UHNM with markedly raised levels of CRP (>100 mg/L) and classified as having type 1 or type 2 diabetes. During the time available for the sample collection, it was possible to recruit 24 diabetic patients; 20 patients with type II diabetes and only 4 patients with type I diabetes, since type 1 diabetes is less common in adults than type 2 (Hex *et al.*, 2012). With a sample size of 24 patients, the ME increases from  $\pm 13\%$  (when  $n=40$ ) to  $\pm 17\%$  (with  $n=24$ ), meaning that the modified form of CRP would be expected to be identified in 14 to 22 (58%-92%) serum samples at 95% confidence level. A full list of the sample IDs, type of diabetes and the initial CRP concentrations (mg/L) as provided by UHNM can be found in Appendix 3.

The methodology used for the analysis of human serum samples is summarised in Figure 4-1: CRP was first purified from human serum via affinity chromatography, followed by separation of CRP isoforms (pCRP and mCRP) using size exclusion chromatography. The pCRP containing fractions were quantified by ELISA and subjected to LC-MS and LC-MSMS analysis to test for the glycated isoform of CRP. The mCRP containing fractions were subjected to LC-MS and ELISA for mCRP identification and quantification, respectively.

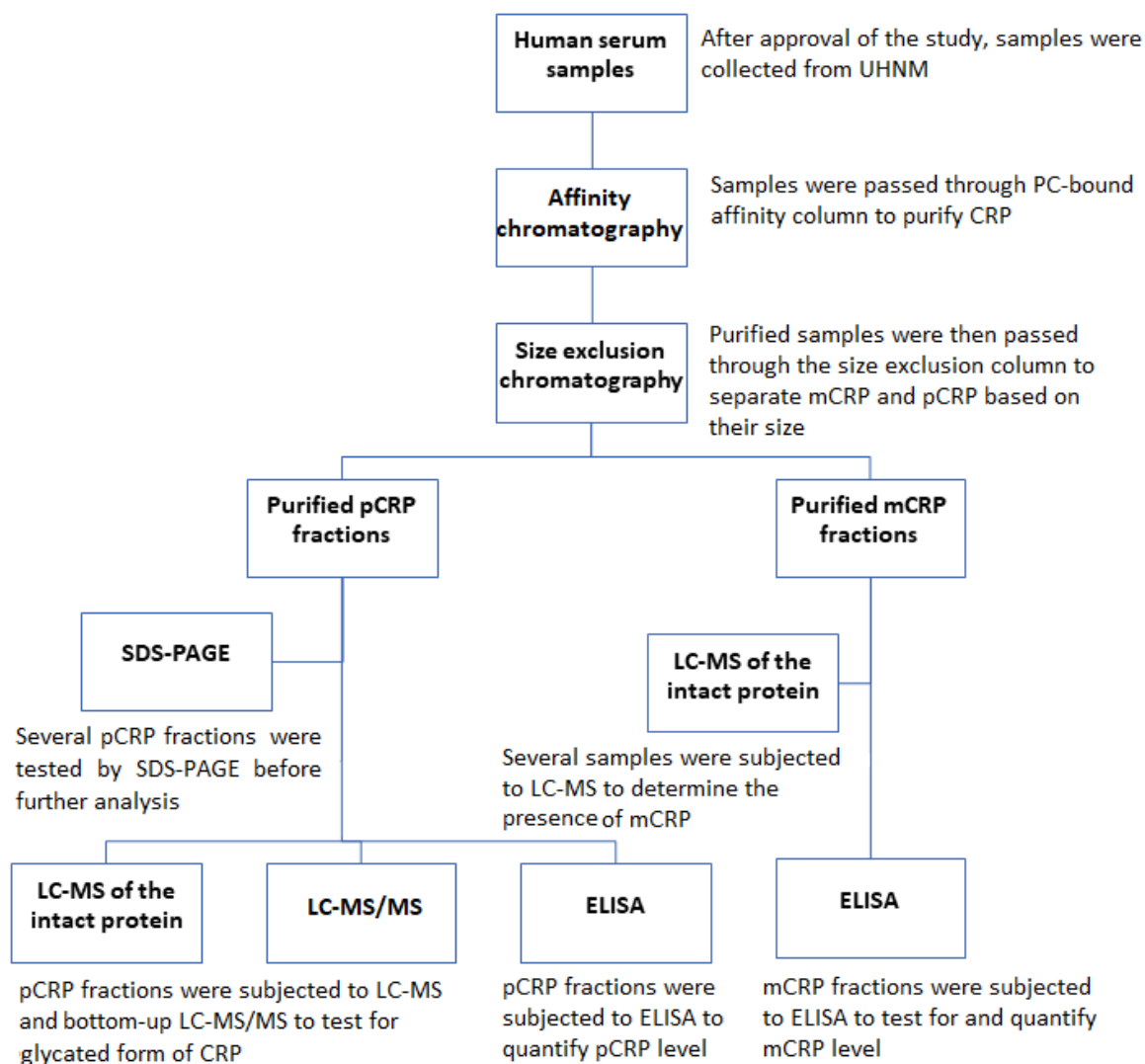


Figure 4-1 A flow diagram of the methodology used for the analysis of human serum samples. Unless otherwise indicated, each experimental method was performed on all samples

#### 4.3.2 Calibration of the size exclusion chromatography column

Prior to the use of size exclusion chromatography, the column was calibrated with known molecular weight proteins including: aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumin (44 kDa), conalbumin (75 kDa), in addition to CRP (115 kDa) to further expand the molecular weight range of the calibration standards. Figure 4-2 shows the chromatography trace for the molecular weight standards; the sample was

injected into the column at 0 mL, and six peaks were observed at elution volumes of 70.93, 75.64, 81.47, 89.12, 96.99 and 105.92 mL for CRP, conalbumin, ovalbumin, carbonic anhydrase, ribonuclease A and aprotinin, respectively.

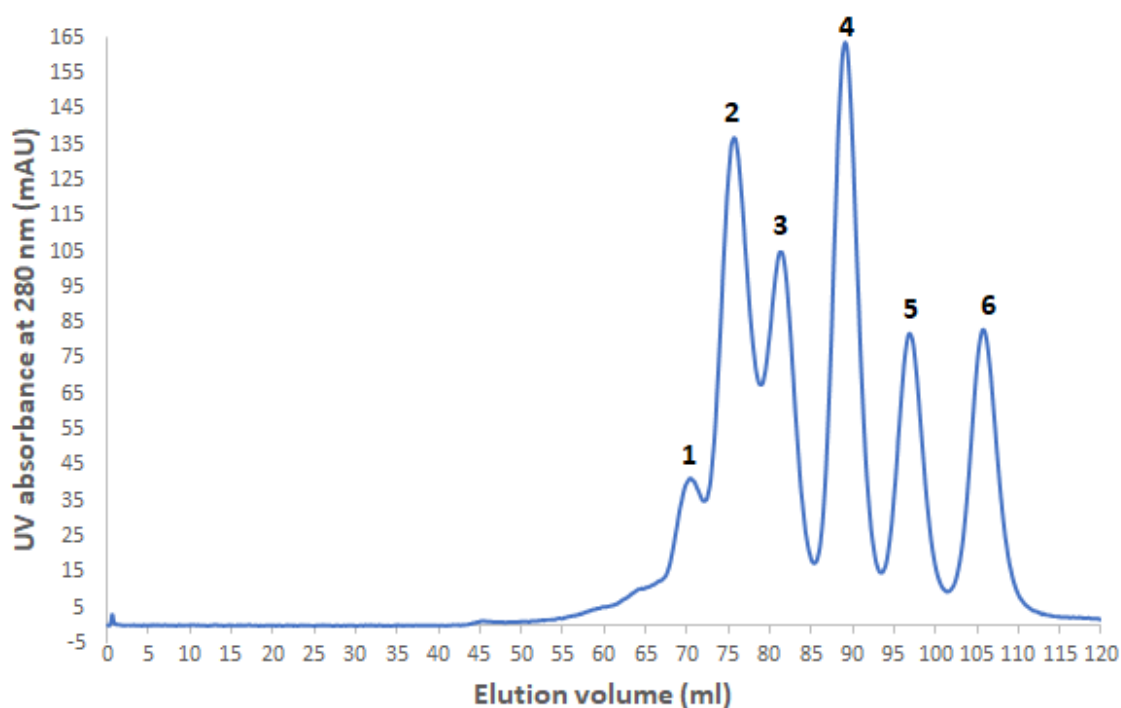


Figure 4-2 A size exclusion chromatography trace of the molecular weight standards. Peak 1, CRP (115 kDa), has an elution volume of 70.93 mL. Peak 2, conalbumin (75 kDa), has an elution volume of 75.64 mL. Peak 3, ovalbumin (44 kDa), has an elution volume of 81.47 mL. Peak 4, carbonic anhydrase (29 kDa), has an elution volume of 89.12 mL. Peak 5, ribonuclease A (13.7 kDa), has an elution volume of 96.99 mL. Peak 6, aprotinin (6.5 kDa), has an elution volume of 105.92 mL.

A second chromatography run was performed with Blue Dextran, a high molecular weight glucose polymer, in order to determine the column void volume. The chromatography trace of Blue Dextran shows a single sharp peak with maximal absorption at 44.98 mL; which represents the column void volume (see Figure 4-3).

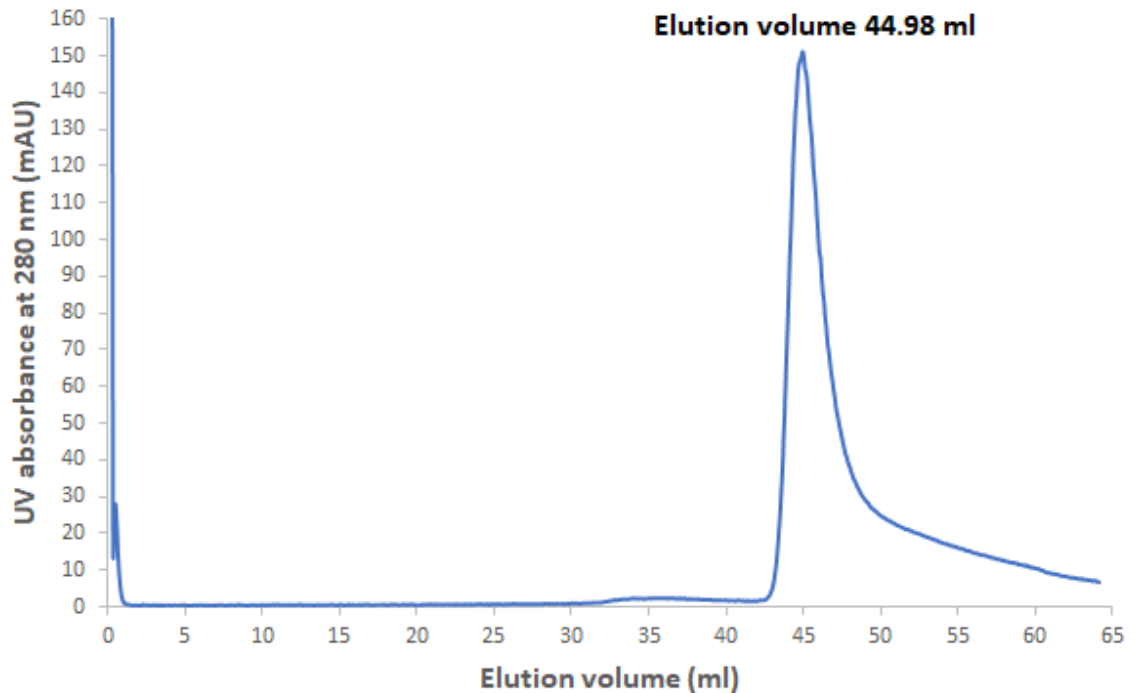


Figure 4-3 A size exclusion chromatography trace for Blue Dextran. Single sharp peak is observed at 44.98 mL elution volume. Absorbance was measured at 280 nm.

The elution volumes of the protein standards obtained from the chromatography trace in Figure 4-2 were used to calculate the proportion of pores available for each protein standard ( $K_{av}$ ) using the following equation:

Equation 4-2 The equation used to calculate  $K_{av}$  value for the protein standards

$$K_{av} = \frac{V_e - V_o}{V_t - V_o}$$

In equation 4-2,  $V_e$  is the elution volume for the protein standard,  $V_o$  is the column void volume (44.98 mL from Figure 4-3),  $V_t$  is the total column volume, which was calculated to be 120.63 mL ( $\pi \times \text{radius} \times \text{length}$ ).

The calculated  $K_{av}$  values for each of the protein standards are presented in Table 4-7 in addition to the  $\log_{10}$  of the molecular weight of each protein standard. The calibration

graph was constructed by plotting the  $K_{av}$  versus the  $\log_{10}$  of their molecular weights as shown in figure 4-4.

Table 4-7 A table of the calculated  $K_{av}$  values for the molecular weight protein standards. The molecular weights, elution volumes, and  $\log_{10}$  of molecular weights for all standards are also shown.

Protein	Molecular Weight (kDa)	Elution Volume (mL)	$K_{av}$	$\log_{10}$ of Molecular Weight
Aprotinin	6.5	105.92	0.806	3.813
Ribonuclease A	13.7	96.99	0.688	4.137
Carbonic anhydrase	29	89.12	0.583	4.462
Ovalbumin	44	81.47	0.482	4.643
Conalbumin	75	75.64	0.405	4.875
pCRP	115	70.93	0.343	5.061

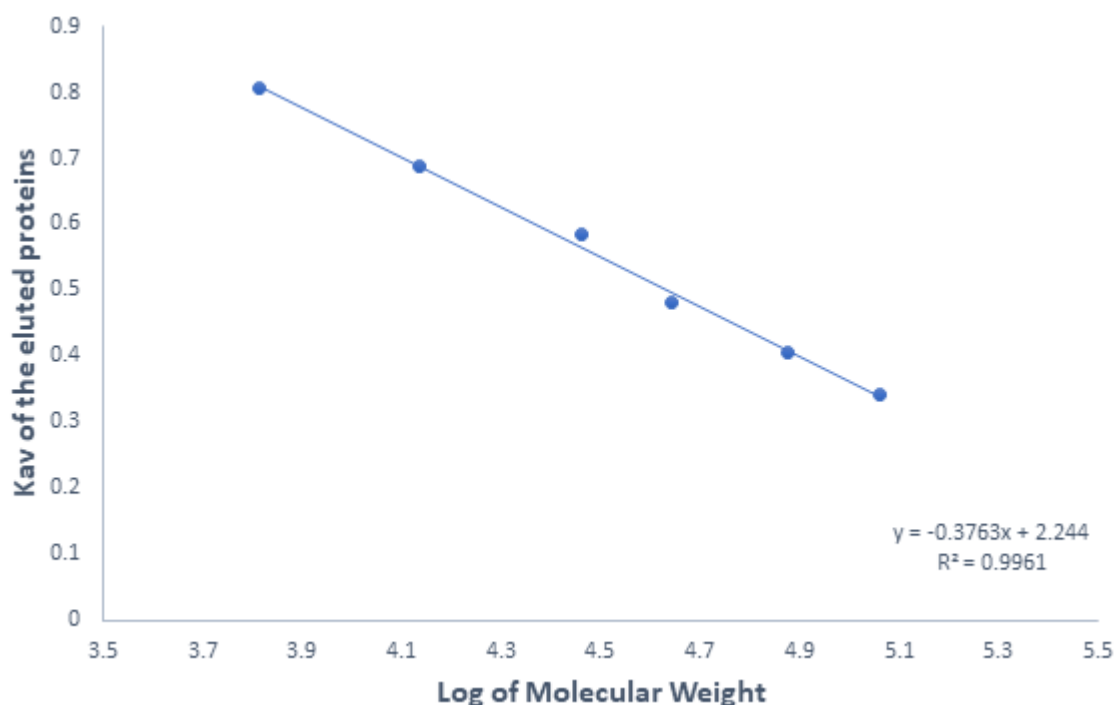


Figure 4-4 Calibration graph of the  $K_{av}$  values of the molecular weight standards versus the  $\log_{10}$  of their molecular weight. The  $R^2$  value is 0.9961, and line equation for the graph is  $y = -0.3763x + 2.244$

The calibration graph displays an inverse relationship between  $\log_{10}$  molecular weight and  $K_{av}$  meaning that the larger the  $K_{av}$  (representative of elution volume), the smaller the molecular weight of the protein. The high  $R^2$  value of 0.9961 indicates the strong correlation between the two data sets. The line equation of the calibration graph ( $y = -0.3763x + 2.244$ ) describes the inverse relationship between  $K_{av}$  and  $\log_{10}$  of the molecular weight of the proteins. From this equation, the molecular weight of the proteins in the sample can be determined based on their elution volumes from the size exclusion column. The ability to estimate the molecular weight of the eluted proteins based on their elution volumes enable reliable identification of CRP forms (whether mCRP or pCRP form) in the *in vitro*-produced mCRP sample or the purified human serum samples.

#### 4.3.3 Purification of *in vitro* prepared monomeric C-reactive protein

mCRP was prepared *in vitro* from the dissociation of native pCRP using 3 M urea in the presence of EDTA as calcium chelator (see section 4.2.6). After 10 weeks of incubation in dissociation buffer, size exclusion chromatography was used to separate mCRP from residual undissociated pCRP. Figure 4-5 shows a size exclusion chromatography trace of the dissociation trial; two peaks are observed on the chromatography trace, labelled peak 1 and peak 2. Peak 1 is broad in shape and small, eluted at volume ~73 mL with absorbance not exceeding 2.9 mAU. Applying the calibration equation (figure 4.4), the elution volume of peak 1 represents a protein with a molecular weight of 95.2 kDa. Peak 2 is sharp and significantly larger with an absorbance of 73.6 mAU and elution volume of 89.9 mL (peak start-end; 86-93 mL). Peak 2 represents fraction with molecular weight 24.2 kDa, consistent with the molecular weight of mCRP. Fractions corresponding to the mCRP (peak 2) were collected and concentrated to 0.5 mL. The concentration was measured by Nanodrop and found to be 0.4 mg/mL. The protein was stored at 4 °C for further experimental analysis.



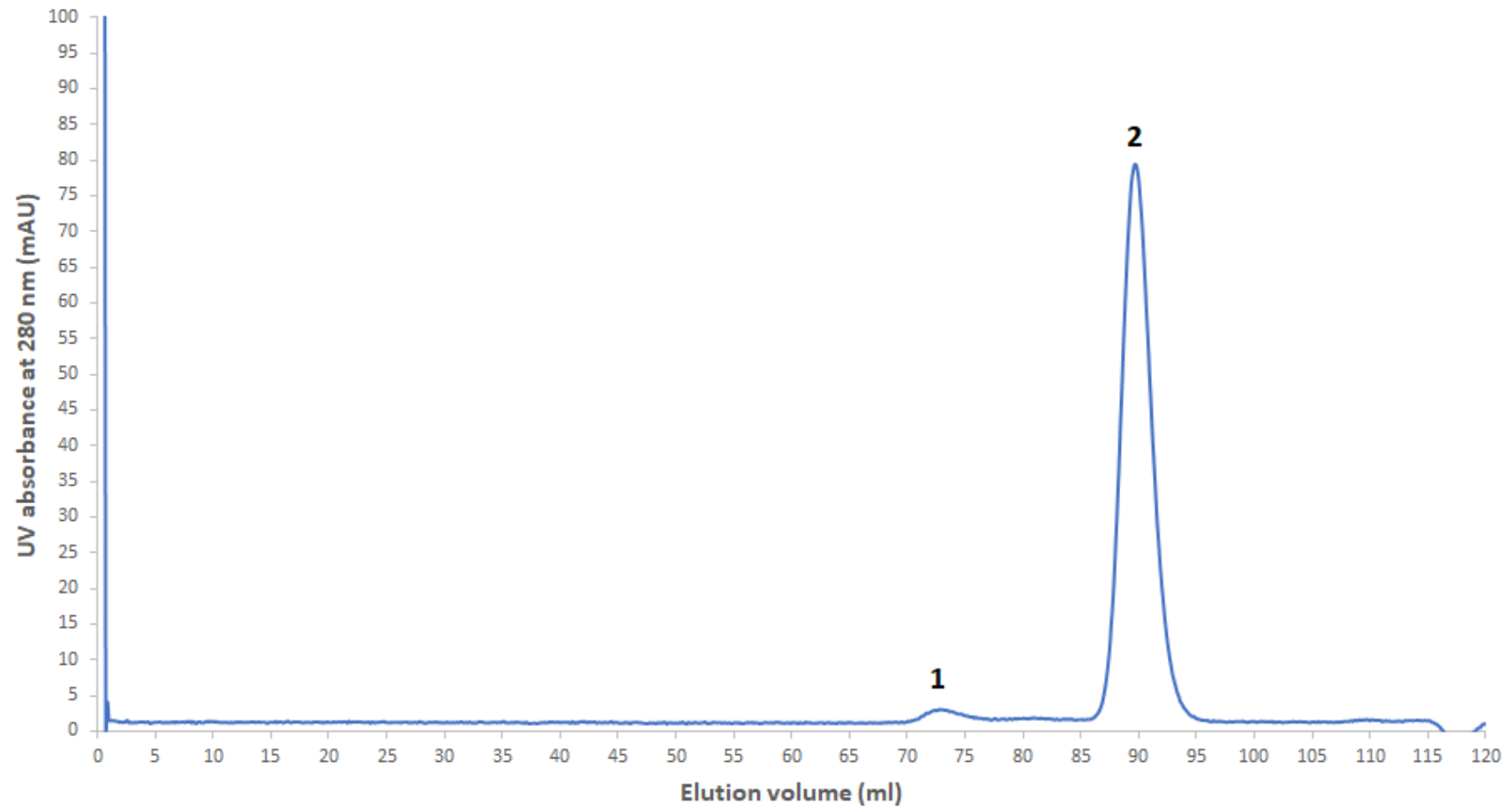


Figure 4-5 A Size exclusion chromatography trace of CRP treated with 3 M urea in the presence of EDTA for 10 weeks. Peak 1 has an elution volume of ~73 mL and UV absorbance of 2.9 mAU. Peak 2 is significantly larger with UV absorbance of 73.6 mAU, and elution volume of 89.9 mL consisting with the elution volume of mCRP. The sample was injected at 0 mL and the UV absorbance was measured at 280 nm.

#### 4.3.4 Purification of the human serum samples by affinity chromatography

Affinity chromatography was performed on collected serum samples to purify CRP from other serum constituents. Samples were passed through a phosphorylcholine (PC) affinity column in the presence of calcium to enable CRP binding to its natural ligand, PC, while the other constituents were washed out. The CRP bound to the column was then eluted by using an EDTA-containing buffer. All 24 samples were purified by the same method. A representative trace for the purification of human serum by affinity chromatography is shown in Figure 4-6. All results obtained from affinity chromatography followed the same trend.

After loading the sample, while the calcium buffer passes through the column at a low flow rate, a peak started to develop just after 2.7 mL of the calcium buffer (peak 1 in the figure 4-6). This peak reached an absorbance reading of 3060.9 mAU at 6.7 mL, with absorbance reading not returning to 0 mAU until 15.9 mL of calcium buffer had passed. Peak 1 is due to the elution of serum constituents that were unable to bind to the column. The calcium buffer was allowed to run through the column for an additional time after the absorbance reading had returned to the baseline level to aid removal of unbound constituents. The calcium buffer was then replaced with the calcium chelator as indicated by the marker at the volume point of around 41 mL. A second peak started to develop after several mL of the calcium chelator passed through the column, and extended for around 4.5 mL with absorbance reading reached 73.9 mAU (see peak 2 in the Figure 4-6). Peak 2 on the chromatography trace is due to the elution of CRP from the PC-column due to the removal of calcium by EDTA since the calcium removal results in elution of calcium-dependent PC-bound proteins, such as CRP. The slight background increase that seen before peak 2 in the magnified section of the Figure 4-6 is due to loading of the chelator buffer to the column.

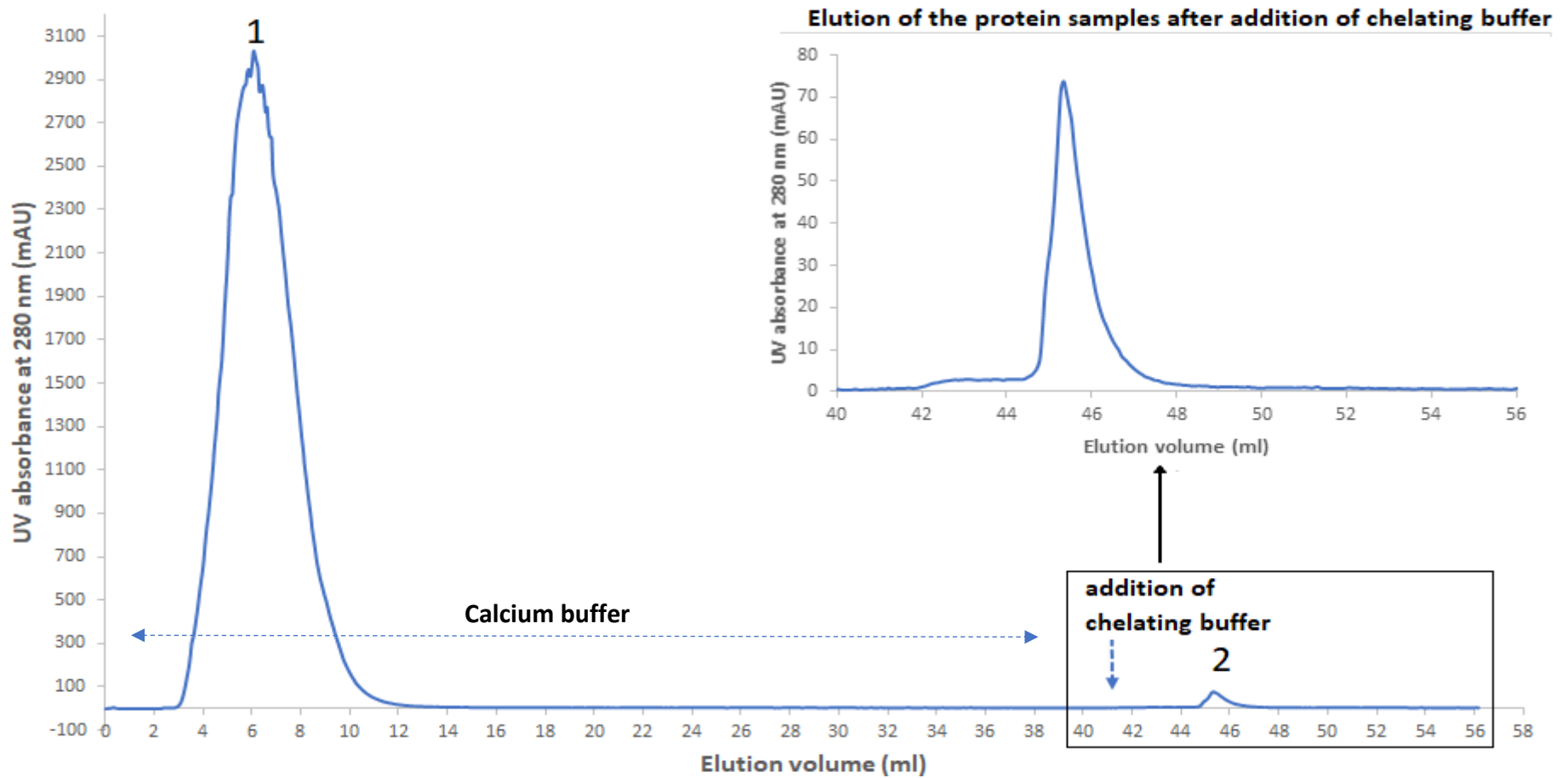


Figure 4-6 A chromatography trace of human serum passed through PC-bound affinity column. Peak 1 eluted at 6.7 mL when the calcium buffer passes through the column, this peak represents the serum constituents which did not bind to the column. Peak 2 was eluted after addition of the chelating buffer, it has an absorbance reading of 73.96 mAU and represents elution of the calcium-dependent PC-bound CRP.

#### 4.3.5 Isolation of CRP isoforms by size exclusion chromatography

The calcium-dependent proteins from human serum samples retained by PC-bound affinity column were subjected to size exclusion chromatography to further purify and separate the potential CRP isoforms based on their sizes. Affinity chromatography fraction from peak 2 of the chromatography trace in Figure 4-6 was collected, concentrated to 1 mL volume, and loaded onto a size exclusion column. A representative trace for the results acquired from the size exclusion chromatography is shown in Figure 4-7. Three peaks are observed on the chromatography trace and labelled peaks 1, 2 and 3. Peak 1 was eluted before the column void volume at 39.6 mL with an absorbance of 1.9 mAU. Peak 2 was sharp and larger than peaks 1 and 3 with an absorbance reading of 9.5 mAU. Peak 2 had an elution volume of 71.8 mL (peak start-end; 69.5-76.7 mL). Based on the calibration graph equation, peak 2 represents a protein with a molecular weight of 105 kDa which is consistent with the molecular weight for pCRP. Peak 3 was eluted at ~109.5 mL with an absorbance reading of 2.3 mAU. Based on the calibration graph equation, the elution volume of peak 3 represents a protein with a molecular weight of 4.98 kDa.

For all the samples, a peak corresponding to elution of pCRP was observed, with an elution volume around 72 mL and UV absorbance in the range of 4.89-15.23 mAU with a median of 10.34 mAU. The vast majority of the samples (18 out of 24) show a chromatography trace typical to the example presented here with three peaks; one eluting before the column void volume, a second peak corresponding to pCRP, and a third peak at ~110 mL. Three of the studied samples show only one peak corresponding to pCRP, while the other three samples showed two peaks: one peak at pCRP elution area and another peak either before the column void volume or at ~110 mL. The peaks that were eluted before the

column void volume or at ~110 mL are outside of the column resolution and were not subjected to further experimental analysis and disregarded.

Based on calibration graph and the size exclusion run of the *in vitro* produced mCRP (Figure 4-5), mCRP isoform should be eluted at a volume of ~90 mL (90 mL elution volume represent a protein with the molecular weight of around 24 kDa). However, with all samples, the absorbance reading remain around the baseline level with no evidence of any peak at the region that correspond to the elution of mCRP.

The fractions containing pCRP (typically from elution volume of 65 to 80 mL), in addition to fractions at elution volume between 85 to 95 were collected separately, concentrated, and stored at 4 °C for further experimental analysis. The reason behind collecting the fractions eluting between 85 to 95 mL, although no peak was observed, is to exclude mCRP levels being too low to be detected the UV absorbance detector and inducing an elution peak.

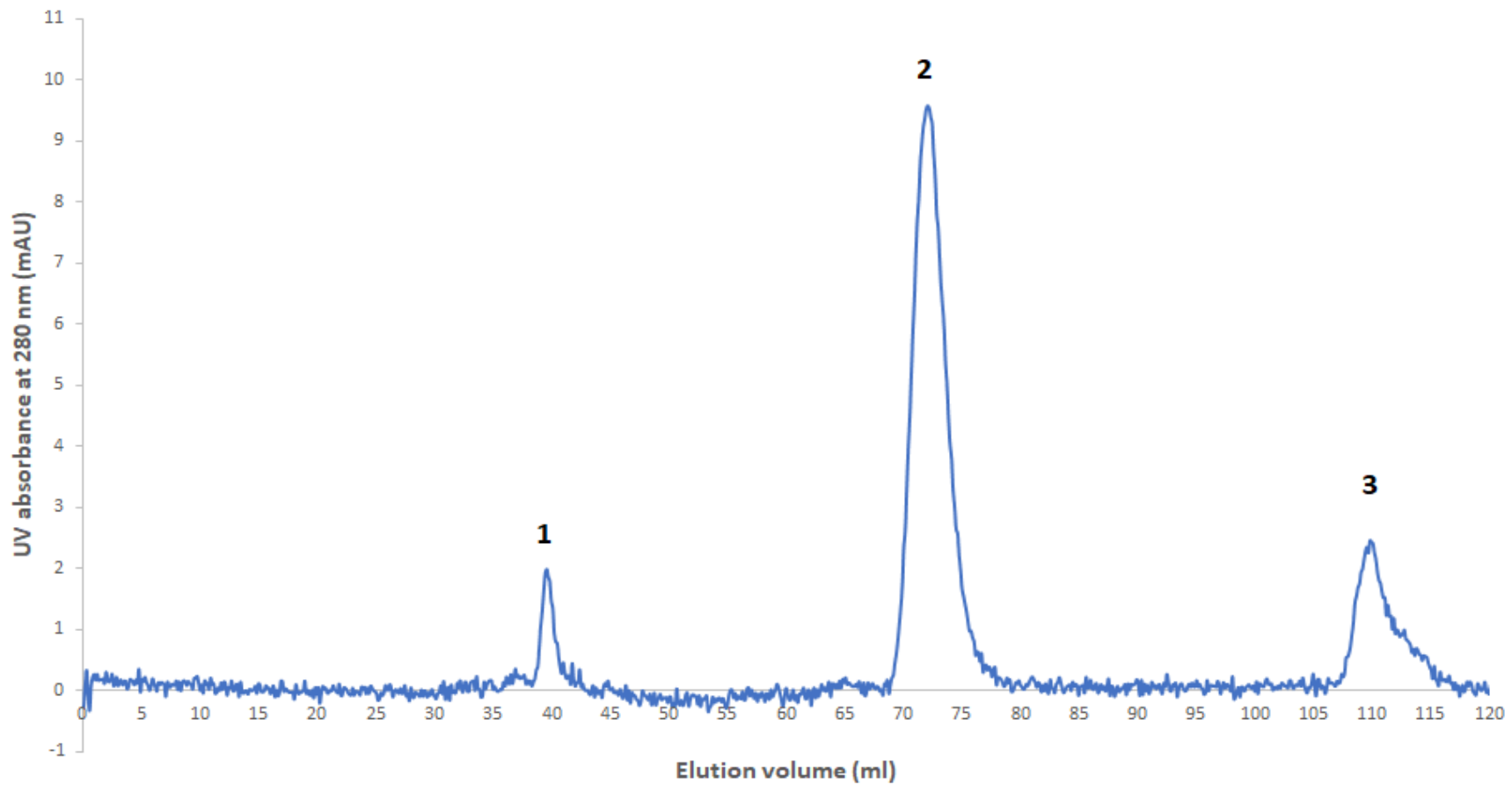


Figure 4-7 Size exclusion chromatography trace of the proteins from human serum sample retained by the PC-bound affinity column. The presented trace represent the typical image of the purified human serum samples when subjected to size exclusion column. Three peaks were observed: the largest peak (labelled 2) displayed an elution volume of 71.8 mL and corresponds to pCRP, peaks labelled 1 and 3 displayed elution volumes of 39.6 mL and 110 mL, respectively.

#### 4.3.6 Monitoring size exclusion chromatography by SDS-PAGE

SDS-PAGE was carried out to determine the presence and purity of CRP in the fractions obtained from size exclusion chromatography. Only the pCRP fractions were tested, as we anticipated, based on prior data, that the amount of mCRP would be insufficient to be detected by SDS-PAGE. Figure 4-8 shows the SDS-PAGE result of the size exclusion fractions from fourteen human serum samples. All tested samples produced a single band at a molecular weight of  $\sim 23$  kDa, confirming the presence of CRP in the eluted fractions. For all the samples, no bands other than the CRP band were observed. Staining intensity indicated varying protein concentration.

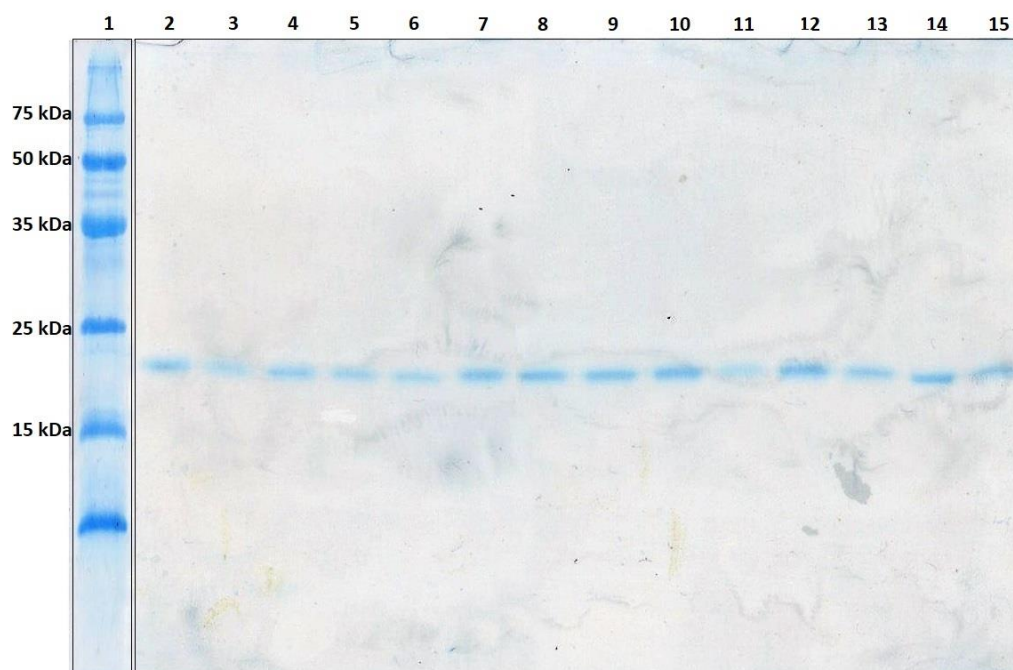


Figure 4-8 SDS-PAGE analysis of pCRP containing fractions after size exclusion chromatography separation. Lane 1 represents the molecular weight markers, while the lanes from 2 to 15 are the eluted fractions that thought to contain pCRP from the fourteen human serum samples. A total volume of 20  $\mu$ L of each sample was loaded to the gel, and the gel was stained with Coomassie brilliant blue.

#### 4.3.7 Enzyme-linked immunosorbent assay of the eluted serum fractions

Fractions obtained from size exclusion chromatography corresponding to the elution volume of pCRP (typically from elution volume of 65 to 80 mL) or mCRP (typically from elution volume 85 to 95) were concentrated and subjected to ELISA analysis in order to detect and quantify the levels of these isoforms, if present, in the corresponding fractions. Analysis via ELISA offers an increased level of sensitivity to detect a level of the protein which may be too low to be detected by the UV detector of the size exclusion column.

Fractions contain pCRP were tested in triplicate with samples also triplicated upon the plate. Known amounts of pCRP, purchased from SCRIPPS, were analysed beside the pCRP containing fractions to produce a calibration curve. Figure 4-9 shows the calibration graph generated from plotting the known concentrations of pCRP versus their absorbances. The calibration graph displays a strong positive correlation between the concentration of the protein and its absorbance as indicated by  $R^2$  value 0.9987. The line equation of the calibration graph ( $y=0.0378x+0.0225$ ) was used to determine the respective pCRP concentrations in the tested samples.

The result of ELISA analysis for the pCRP containing fractions is presented in Figure 4-10. The concentration of the protein samples is expressed in term of mg/L, and the values have been corrected to account for pCRP sample volume post purification process to give an estimate concentration of pCRP in the human serum samples (see 4.2.10 and Appendix 4). The average value of pCRP across all 24 human serum samples was calculated to be 30.03 mg/L (with SD equal to 13.85, and SE  $\pm 2.83$ ), the lowest recorded pCRP value was 4.09 mg/L (CRP05) and the highest recorded value 57.75 mg/L.



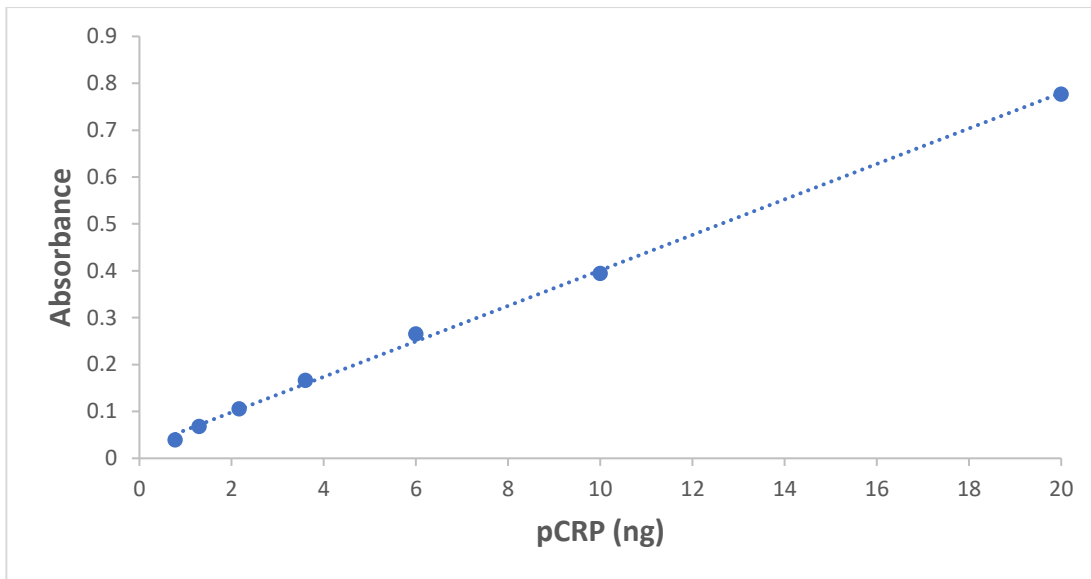


Figure 4-9 The calibration graph for pCRP ELISA assays. The line equation is  $y=0.0378x+0.0225$  and the  $R^2 = 0.9987$

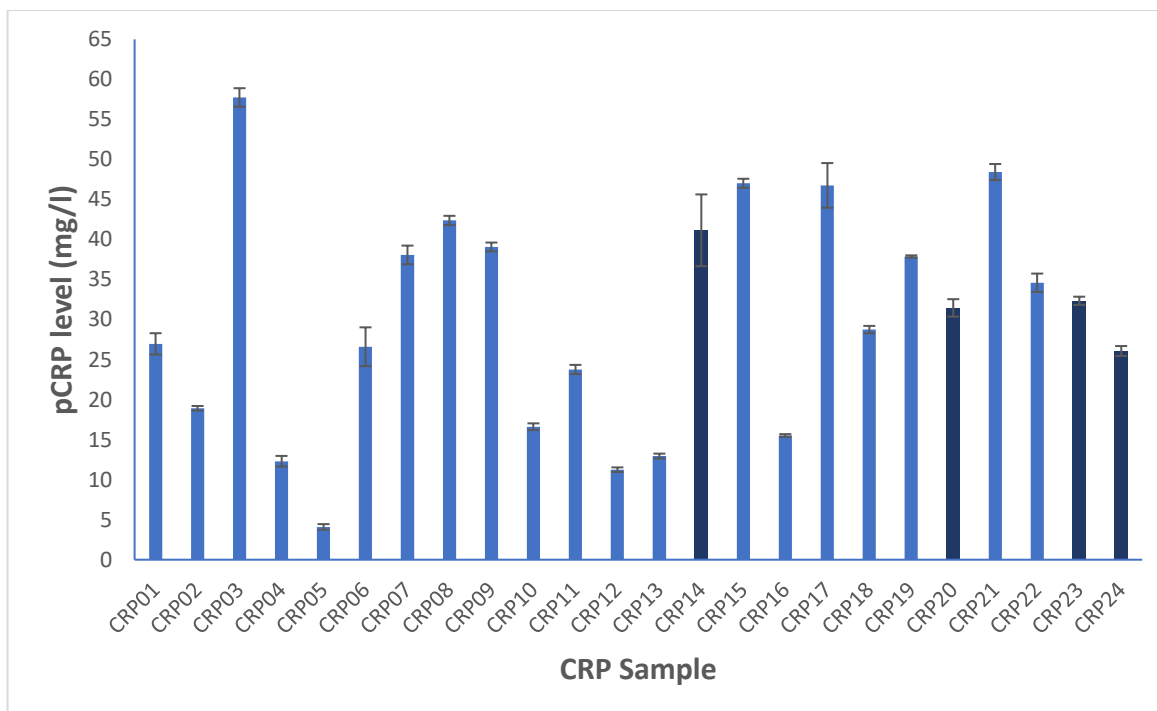


Figure 4-10 Pentameric CRP concentrations purified from serum samples as determined by ELISA. The fractions containing pCRP were diluted at 1:1000 and tested in triplicate over 3 separate ELISA assays. The primary and secondary antibody dilutions were 1:40,000, the samples were developed with TMB and measured at 450 nm. The bars represent the average concentrations in term of mg/L; the error bars are the SEM for each sample. Darker bars are samples from type 1 diabetic patients, paler are from type 2 diabetic patients.

The results of the ELISA quantification of pCRP-containing samples indicates a significant loss of protein post purification compared to the initial sample concentrations, where all samples were over 100 mg/L. The initial samples concentrations, as provided by UHNM, were plotted against the pCRP concentrations obtained via the ELISA analysis to provide an overview of the relationship between these two data sets. This comparison is shown in Figure 4-11. The correlation coefficient between the two data set is 0.6 (p-value 0.002), indicating a moderate positive correlation between the purified pCRP concentration and the initial CRP concentration.

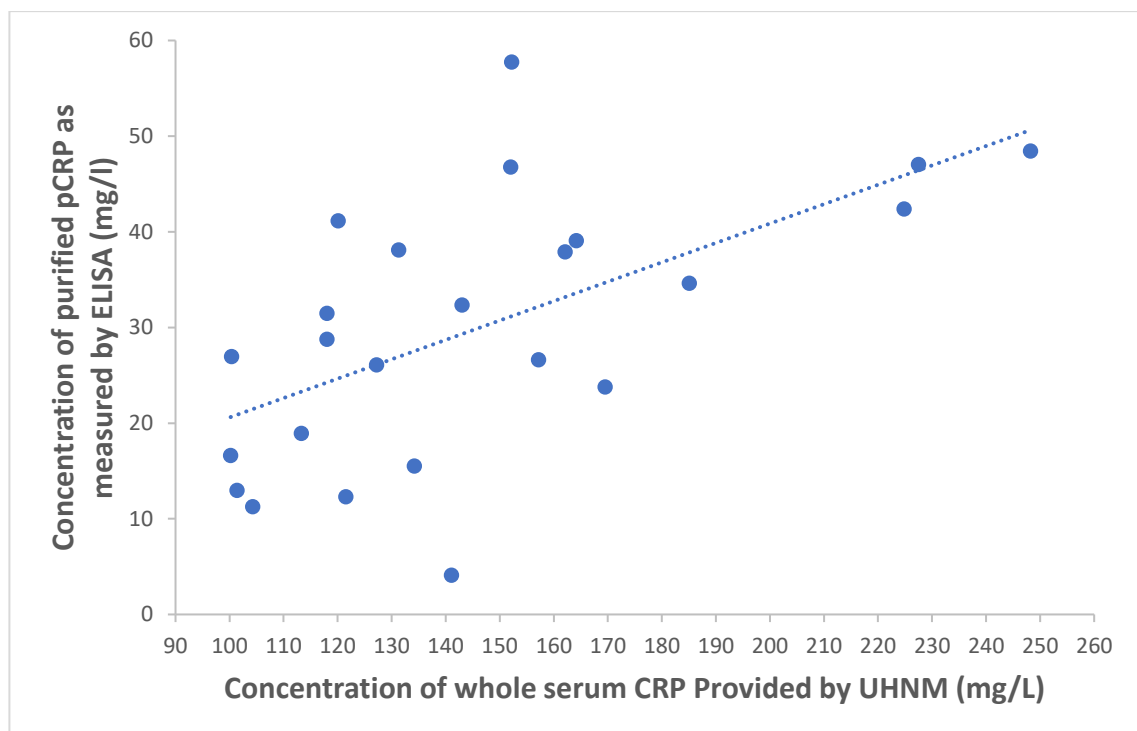


Figure 4-11 A comparison of the pCRP concentrations obtained from the ELISA analysis of purified pCRP against the data provided from UHNM as whole serum CRP levels. A moderate positive correlation between the data sets was found with a correlation coefficient of 0.6

Size exclusion fractions corresponding to the elution volume of mCRP were analysed in the same way, with samples being triplicated upon the ELISA plate. Known amounts of mCRP, produced *in vitro* from native pCRP by urea and chelator treatment, were included alongside mCRP-containing samples to produce the calibration curve. Figure 4-12 shows the calibration graph generated from plotting the mCRP standards versus their absorbances. The calibration graph displays a strong positive correlation between the concentration of the protein and its absorbance as indicated by  $R^2$  value 0.9961. The line equation of the calibration graph ( $y=0.0616x-0.0174$ ) was used to determine the respective mCRP concentrations in the tested samples.

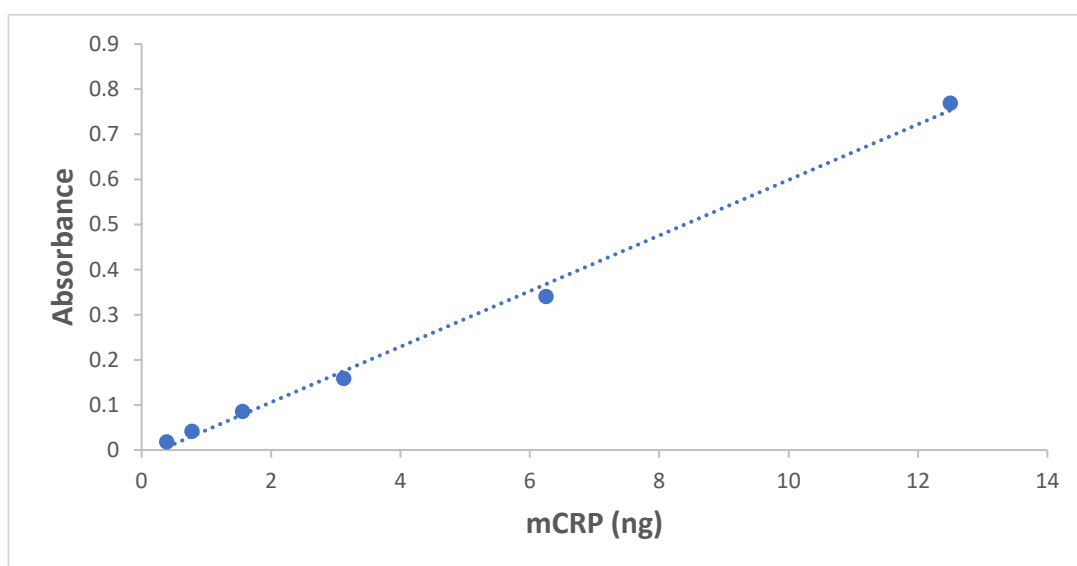


Figure 4-12 Calibration graph for mCRP ELISA assays. The line equation is  $y=0.0616x-0.0174$  and the  $R^2 = 0.9961$

The result of ELISA analysis for the mCRP-containing fractions is presented in Figure 4-13. The concentration of mCRP in the samples is expressed in term of mg/L, and the values have been corrected to account for the mCRP sample volume post purification process to give an estimate concentration of mCRP in the serum samples (see 4.2.10 and Appendix 3).

The entire 24 human serum samples show a detectable level of mCRP, with an average mCRP content of 0.093 mg/L (SD of mean equal to 0.0461, and SE mean  $\pm 0.0094$ ). The lowest recorded mCRP value was 0.043 mg/L and the highest recorded value 0.219 mg/L. The median value across all 24 samples was calculated to be 0.079 mg/L, IQR 0.058-0.109 mg/L. For all total 24 human serum sample, all mCRP values were significantly lower than those for pCRP.

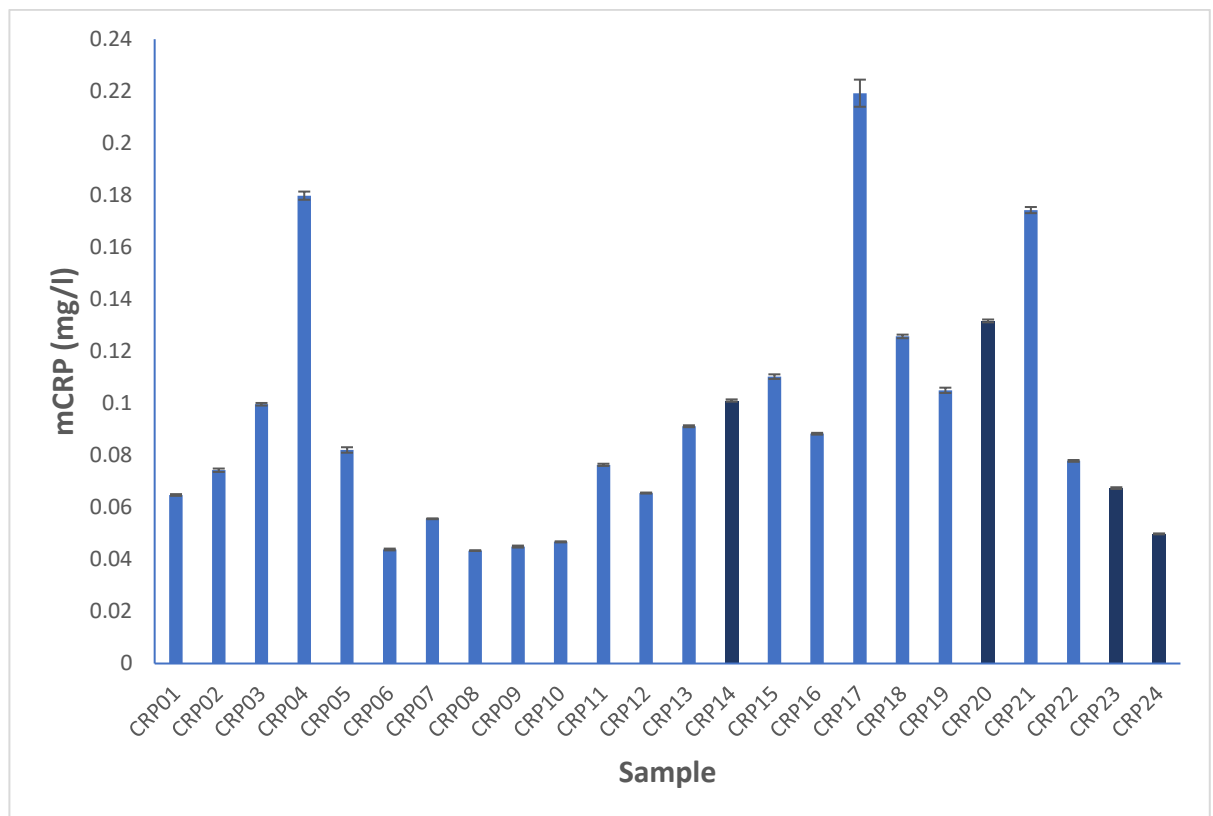


Figure 4-13 Monomeric CRP concentrations purified from human serum samples as determined by ELISA. The fractions containing mCRP were measured in triplicate. The primary and secondary antibody dilutions were 1:40,000, the samples were developed with TMB and measured at 450 nm. The bars represent the average concentrations in term of mg/L, the error bars are the SEM for each sample. Darker bars are samples from type 1 diabetic patients, paler were type 2 diabetic patients.

Since the serum samples were collected from both type 1 and type 2 diabetes participants, the differences in the level of mCRP between the two groups were investigated. Type 1 diabetes group comprised 4 participants with a median mCRP value of 0.084 mg/L, while the type 2 diabetes group comprised 20 participants with a median mCRP value of 0.080 mg/L. Comparing the mCRP levels between the two groups using the Mann-Whitney test for not normally distributed data, show no significant difference between the groups ( $P > 0.908$ ). However, due to the small sample size of the groups, particularly that of type 1 diabetic participants, it is not possible to draw a meaningful predictive conclusion.

The differences in the level of pCRP between the two groups (type 1 and type 2) were also investigated using Mann-Whitney test, no significant difference in the level of pCRP were found between type 1 and type 2 diabetes groups ( $P > 0.7$ ).

The correlation between the levels of mCRP and pCRP within the serum sample for the entire 24 participants was also investigated. No correlation was found between levels of mCRP and pCRP as illustrated in Figure 4-14 with  $R^2 = 0.0574$ ,  $P > 0.26$ .

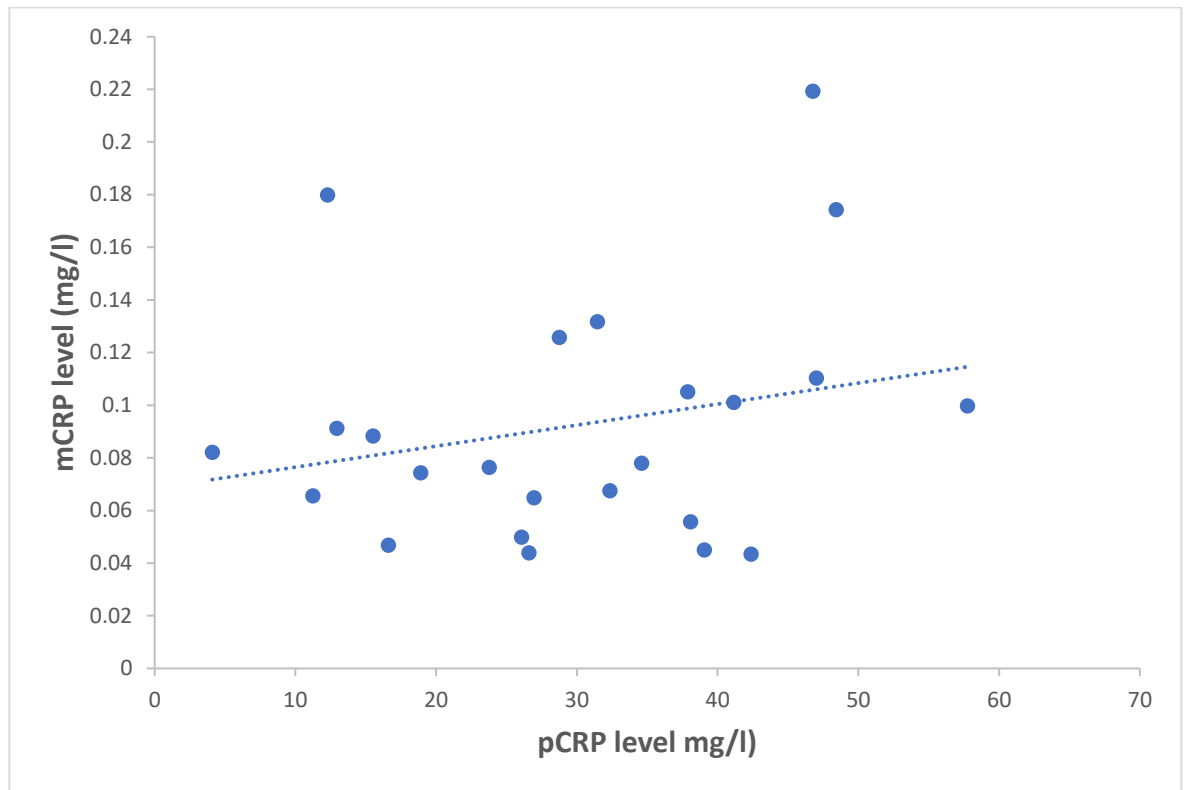


Figure 4-14 Scatter plot illustrating the correlation between the pCRP and mCRP levels purified from human serum samples as determined by ELISA. The line equation is  $y = 0.0008x + 0.0685$ , with  $R^2 = 0.0574$  showing no correlation between levels of mCRP and pCRP  $P > 0.26$ . The concentration for each sample is expressed in terms of mg/L.

## 4.3.8 Intact mass analysis of mCRP in diabetic serum samples

Monomeric CRP samples with high protein content, as revealed by ELISA, were also subjected to LC-MS analysis at the intact level to demonstrate the identity and the mass of the isolated isoform. The mCRP-containing fractions from three human serum samples were tested. All three samples produced very similar spectra. An example is shown in figure 4-15.

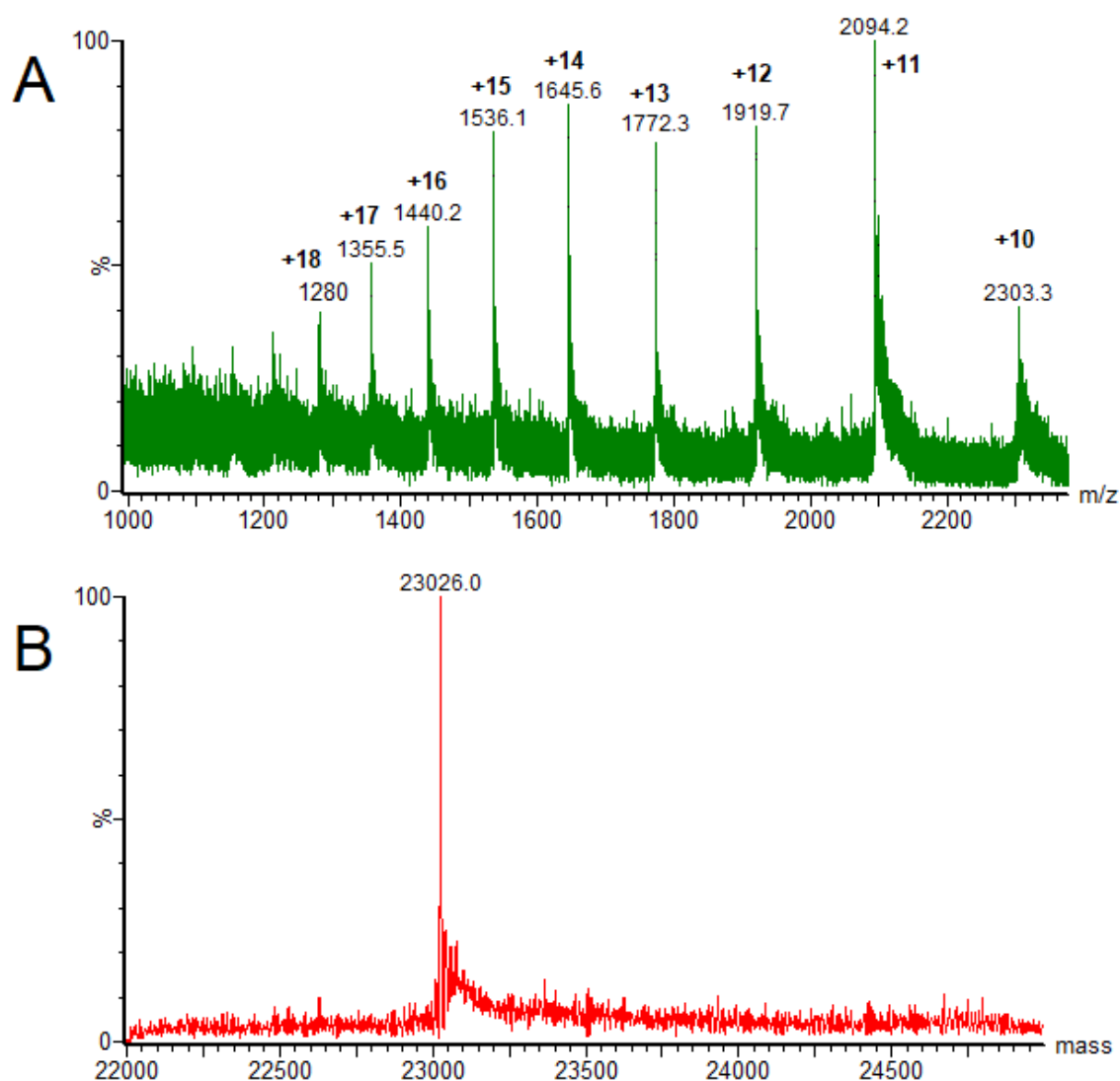


Figure 4-15 Mass spectra of mCRP purified from human serum sample and measured by nESI-QTOF-MS in positive ion mode. A, combined  $m/z$  spectrum after LC-MS analysis of 5  $\mu$ l of mCRP isolated from a diabetic human serum sample. The charge states are indicated on the spectrum, and the intensities are scaled to the highest peak, B: Deconvoluted LC-MS mass spectrum of mCRP based on the  $m/z$  spectrum presented in Figure A.

The low intensities of the peaks in  $m/z$  spectrum are due to the low mCRP concentration in the sample. The protein produced a charge state distribution similar to that produced by standard CRP preparation (figure 2-5). After software-assisted deconvolution of the spectrum, the mass spectrum depicted in Figure 14-5, B was obtained. The deconvoluted spectrum shows a mean species at 23,026 Da, which corresponds to the expected mass of unmodified monomeric CRP.

#### 4.3.9 Intact mass analysis of pCRP in human serum samples

Mass spectrometry has been utilised to investigate the presence of a glycated form of CRP in diabetic serum samples; pCRP isolated from human serum samples were analysed by Q-TOF mass spectrometer at the intact level. LC-MS analysis of pCRP shows a main peak at 23027 Da corresponding to the expected mass of the unmodified CRP subunit. However, none of the samples shows any evidence of peak corresponding to the condensation of glucose molecules. All 24 human serum samples were tested and produced the same spectra as shown in Figure 4-16. Low-intensity signals, besides the main peak at 23027 Da, were seen with all analysed human samples but also found in the standard CRP preparation which analysed prior to running the human samples. These peaks are likely due to artefacts causing a mass shift of -18 Da, +17 Da, +34 Da and +221 Da to protein but not related to the protein glycation.

Further LC-MS attempts were also carried out with five samples using SYNAPT G2-S mass spectrometer (Waters) which offers improved sensitivity than the MICROMASS Q-TOF premier (Waters) which was used in our study (Holčapek *et al.*, 2012). However, near identical results were obtained (data not shown).



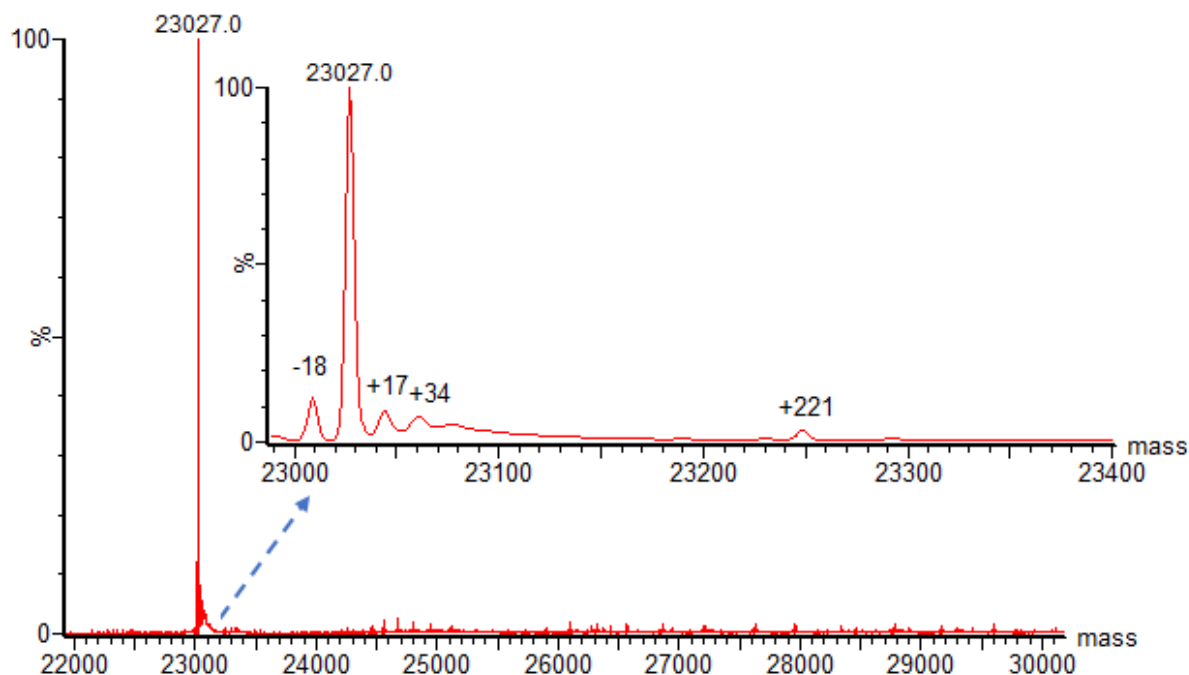


Figure 4-16 The deconvoluted mass spectrum of pCRP isolated from a diabetic serum sample. The main peak at 23027 Da corresponds to the expected mass of unmodified CRP subunit. No peaks corresponding to the condensation of glucose molecules were identified.

#### 4.3.10 Bottom-up tandem mass analysis of pCRP in human serum samples

The investigation of the presence of the glycated form of CRP in the diabetic serum samples was extended to include bottom-up tandem mass analysis. The pCRP purified from diabetic human serum samples were enzymatically digested by trypsin and Glu-C separately, and the resulting digests were analysed by QTOF-MS/MS. The obtained data were submitted to Mascot, enabling identification of CRP peptides and, if present, the possible CRP glycated peptides. For most of the samples (16 out of 24), sequence coverage of >50% was achieved with trypsin alone, and >80% from the Glu-C digest alone (table 4-8). Combining the mapping results of both enzymes generated >90% sequence coverage in 17 out of 24 samples. In 16 out of 24 samples, we were able to cover all the probable glycation sites, while the rest of the digested samples one or more of the lysine residues were not covered.

Table 4-8 Protein sequence coverage obtained after trypsin and Glu-C digestion of pCRP purified from diabetic serum samples, the lysine residues that were not covered by tryptic and Glu-C digest were also given.

Sample ID	% of sequence covered by tryptic digest	% of sequence covered by Glu-C digest	Total sequence coverage obtained from both trypsin and Glu-C digest	Lysine residues that not covered in the total sequence coverage
CRP01	52%	58%	96%	K23
CRP02	53%	94%	100%	
CRP03	58%	81%	100%	
CRP04	52%	62%	81%	K7, 13 and 23
CRP05	21%	39%	52%	K7, 13, 69 and 191
CRP06	28%	91%	97%	
CRP07	53%	94%	100%	
CRP08	58%	81%	100%	
CRP09	21%	46%	55%	K28, 31 and 69
CRP10	58%	91%	100%	
CRP11	52%	94%	100%	
CRP12	58%	91%	100%	
CRP13	62%	94%	100%	
CRP14	22%	46%	55%	K23, 28 and 31
CRP15	21%	57%	68%	K69
CRP16	35%	81%	97%	
CRP17	58%	91%	100%	
CRP18	52%	94%	100%	
CRP19	58%	91%	100%	
CRP20	26%	50%	65%	K7 and 13
CRP21	21%	46%	65%	k7 and 13
CRP22	52%	91%	100%	
CRP23	53%	94%	100%	
CRP24	58%	81%	100%	

For all the 24 human serum samples, only non-glycated CRP peptides were identified through the Mascot search of the data obtained from the digest of trypsin and Glu-C enzymes, no CRP glycosylated peptide was identified. For each sample tested, the  $m/z$  values that did not correspond to any peptide sequence in the selected database during the Mascot search were considered as possible CRP glycosylated peptides and further analysed as explained in section 3.2.4. However, none of these unmatched  $m/z$  values corresponds to glycosylated CRP peptide.

## 4.4 Discussion

### 4.4.1 Production and Purification of Monomeric C-reactive Protein *in vitro*

Treatment of the native pentameric CRP with urea, in the presence of a calcium chelator, yields a monomeric form of CRP (Potempa *et al.*, 1987). While the high concentration of urea (8 M) can achieve dissociation in a short period (Potempa *et al.*, 1987), a less harsh condition with lower urea concentrations can achieve complete dissociation after more extended exposure (Williams, 2017). In this work, mCRP was produced by treating the native pCRP with 3 M urea in the absence of calcium for 10 weeks, in order to minimise harsh conditions while producing mCRP. The size exclusion eluted fraction at ~73 mL corresponded to remaining undissociated protein, whereas the eluted fraction at 89.9 mL represents the *in vitro* produced mCRP (see figure 4-5) based on the calibration graph equation. Consistent with the recent study by Williams study (Williams, 2017), this treatment results in dissociation of >95% of CRP into mCRP (see Figure 4-5).

It is necessary to remove calcium to achieve a high dissociation yield (Kresl *et al.*, 1998). Removal of calcium results in conformational changes within the protein structure, whereby a large loop, composed of residues 140-150, moves away from the body of the protein, revealing hidden sites susceptible to proteolysis (Ramadan *et al.*, 2002). It is possible that these conformational changes as a result of calcium removal promote pentamer instability and weaken the interprotomer contacts making the protein more labile to urea treatment.

Urea is a powerful denaturant, commonly used in proteomic studies due to its ability to interact with both polar and nonpolar groups (Rossky, 2008, Zou *et al.*, 1998). It is accepted that the denaturation of proteins by urea occurs through the direct interaction between

urea and the protein backbone and/or side chains via hydrogen bonds and other electrostatic interactions. This interaction will weaken the internal hydrogen bonds and promote protein unfolding. Upon unfolding and intrusion of urea into the hydrophobic core, both urea and water solvate the exposed hydrophobic, polar, and the peptide backbone (Hua *et al.*, 2008, Rossky, 2008). Another possible mechanism is that urea may exert its effect indirectly by disturbing the water shell around non-polar groups in the protein, thereby weakening the hydrophobic effect of these groups, inducing protein unfolding (Zou *et al.*, 1998, Vanzi *et al.*, 1998). Although the direct interaction mechanism is more widely accepted, it is also plausible that both mechanisms are operating (Stumpe & Grubmüller, 2007, Vanzi *et al.*, 1998). It is unclear by which mechanism urea induces dissociation of CRP, and what the modifications that may have on the function and structure of this protein. CRP units are held together by electrostatic intersubunit interactions (that is, salt bridges and hydrogen bonds); (Shrive *et al.*, 1996). These interactions between CRP protomers may be targeted by urea, causing dissociation. Dissociation of CRP with a high concentration of urea (8 M) results in a loss of the capacity of the protein for calcium-dependent binding to the C-polysaccharide (Potempa *et al.*, 1983). In contrast, mCRP produced from the treatment of the native pCRP with 3 M urea over 10 weeks retains its ability to bind phosphocholine (Williams, 2017). It is worth mentioning that once CRP dissociates into its monomeric form, is able to re-associate even in the presence of calcium (Williams, 2017), perhaps because the ability of urea to carbamylate the free amino groups of protein (N termini of proteins or the  $\epsilon$ -amino group of lysine and arginine residues); (Sun, S. *et al.*, 2014). It is possible that interprotomer contacts become carbamylated by urea, preventing reassociation of subunits.

#### 4.4.2 Purification and isolation of CRP isoforms from human serum samples

Purification of CRP from human serum was carried out by using affinity chromatography based on the calcium-dependent binding of CRP to phosphorylcholine (PC) that has been immobilised on the chromatography column (Volanakis *et al.*, 1978). The serum components that are unable to interact with PC pass freely through the column as represented by a significant peak in figure 4-6. Since the binding of CRP to phosphorylcholine is calcium-dependent, its elution could be achieved merely by chelating the calcium with EDTA (Volanakis *et al.*, 1978). Upon addition of EDTA, a smaller peak was observed, which correspond to PC-bound proteins including CRP. A number of other serum proteins are known to bind PC or to PC-bound CRP (Volanakis *et al.*, 1978, Christner & Mortensen, 1994). However, the bulk of the peak is thought to represent CRP because the serum samples had high CRP content (>100 mg/L). Since the *in vivo* generated mCRP retains its specificity to bind PC (Williams, 2017), the affinity-purified samples should contain both monomeric and pentameric CRP.

Size exclusion chromatography further purified the affinity eluted samples and is able to isolate the different forms of CRP based on their elution volumes. All the 24 patient serum samples produced a peak at an elution volume of ~72 mL (figure 4-7). Based on the results of column calibration, coupled with the results of SDS-PAGE of the eluted fractions (figure 4-8), it was concluded that this peak corresponds to elution of pCRP. Furthermore, SDS-PAGE established the purity of the eluted fractions, by producing a single band per tested sample, and hence the efficiency of the purification process. The chromatography trace in figure 4-7 is a typical example of that seen for most human serum samples. Besides the peak that corresponds to pCRP, additional smaller peaks were commonly observed. Peak 1 in figure 4-7 was eluted before the column void volume, suggesting a molecule with a large

molecular size, which could be a protein complex or CRP aggregate. However, peak elution before the column void volume may also be as a result of a carryover from a previous run, although this is less likely as the column was washed adequately before and after each sample application. The other peak commonly observed was the peak at elution volume of ~110 mL, based on the calibration graph equation the elution volume of this peak (peak 3 in figure 4-7) represent a protein with a molecular weight of 4.98 kDa. Similar to peak 1, it was unclear what peak 3 may represent, fractions correspond to these peaks (peak 1 and 3 in figure 4-7) were not pursued through additional experimental analysis.

Based on the size exclusion column calibration equation, in addition to the result of purification of the *in vitro* produced mCRP (figure 4-5), the mCRP is expected to elute at ~90 mL, and the peak could extend for 10 mL volume. However, none of the purified samples produced a peak for mCRP. This failure to produce a peak on the chromatogram could be as a result of mCRP levels being below the limit of detection. Therefore, the fractions eluting between 85 to 95 mL were collected for each serum sample and tested regardless.

#### 4.4.3 Enzyme-linked immunosorbent assay of the purified serum samples

ELISA was carried out on the size exclusion eluted fractions corresponding mCRP, as well as pCRP-containing fractions. The mean level of pCRP across all 24 diabetic serum samples was calculated to be 30.03 mg/L, with the lowest recorded pCRP value of 4.09 mg/L and the highest recorded value of 57.75 mg/L. The results of the ELISA quantification of pCRP indicates a significant loss of protein during the purification processes as compared to the initial sample concentrations where all samples were over 100 mg/L. This loss is primarily attributed to the purification process which included multiple chromatographic and

concentrating steps to isolate CRP isoforms. Another factor that may have contributed to this loss, although to a lesser extent, is the fact that the ELISA reading represented only the levels of the pCRP portion, excluding other CRP isoforms, aggregates, or CRP complexed with other proteins that have different elution volumes. Although some of the samples displayed more significant losses than others, it was expected to find a significant positive correlation between the initial CRP levels and the levels of the purified pCRP (Figure 4-11) because the majority of CRP is pentameric (native) isoform. Williams (2017) used a similar methodology to purify CRP isoforms from human serum samples with initial CRP values above 100 mg/L. The percentage of yield for pCRP in the Williams study were ranged from 5-30% of the initial sample concentrations, indicating a large protein loss during the purification process (Williams, 2017).

Despite the loss that has been seen with pCRP-containing samples, the ELISA analysis of the purified mCRP-containing samples shows a detectable level of mCRP across all 24 diabetic serum samples, see figure 4-13. The levels of mCRP ranged from 0.043 mg/L to 0.219 mg/L, with mean value of 0.093 mg/L. The literature suggests that mCRP produced *in vivo* could originate from dissociation of pCRP when it binds to damaged membrane surfaces such as apoptotic/necrotic cells (Gershov *et al.*, 2000), activated platelets (Eisenhardt *et al.*, 2009), circulating microparticles (Habersberger *et al.*, 2012, Van der Zee *et al.*, 2010), liposomes containing lysophosphatidylcholine (LPC); (Ji *et al.*, 2007), and oxidised low-density lipoprotein (Chang *et al.*, 2002). The dissociation requires binding of circulating pCRP to the bioactive lipid; LPC, expressed on the membrane surface (Wu *et al.*, 2015, Eisenhardt *et al.*, 2009). However, the curvature of the membrane and the hydrophobic microenvironment were also found to be additional factors affecting the binding and dissociation process of pCRP (Ji *et al.*, 2007, Wang *et al.*, 2012). In addition to the dissociation of pCRP on the

damaged membrane surfaces, independent mCRP synthesis may also be another source of this form. The expression of CRP mRNA has been identified in various extra-hepatic cells including vascular smooth muscle cells in atherosclerotic plaque and human adipose tissue (Yasojima *et al.*, 2001, Devaraj *et al.*, 2003, Jabs *et al.*, 2003). However, whether the CRP that is synthesised by extra-hepatic cells is processed into pCRP or expressed as mCRP remains unclear (Eisenhardt *et al.*, 2009). In fact, CRP mRNA in U937 macrophages is translated into mCRP rather than pCRP (Ciubotaru *et al.*, 2005). Furthermore, studies have also indicated that pCRP can spontaneously dissociate upon binding to a fixed surface such as an ELISA plate (Eisenhardt *et al.*, 2011). However, in this work, the tested samples were purified CRP isoforms, separated according to size by size exclusion chromatography.

The levels of circulating mCRP are significantly low compared to the pCRP level of the same serum sample, this is likely because of the limited solubility of this isoform, rendering it tissue based protein rather than soluble serum-based protein (Molins *et al.*, 2011). Indeed, mCRP was not identified in the peripheral circulation until recently (Habersberger *et al.*, 2012). In Habersberger *et al.*; (2012) study, the mCRP was qualitatively identified, by flow cytometry, on circulating microparticles of patients with acute myocardial infarction (AMI) (Habersberger *et al.*, 2012). Another study by Wang *et al.*; (2015), the researchers were able to quantitatively identify mCRP in the plasma of the patients with AMI by using a monoclonal antibody specific for mCRP (Wang *et al.*, 2015). However, Wang *et al.* did not give details of whether the circulating plasma mCRP reflects free mCRP or mCRP complexed with microparticles (Wang *et al.*, 2015). The methodology used in this study included purification of CRP isoforms through both affinity and size exclusion chromatography before quantification by ELISA, suggested that the identified mCRP is a soluble free circulating form with a molecular size consistent with the size of a single monomer of CRP,



as demonstrated by intact mass analysis, see figure 4-15. The methodology used also shows that the circulating mCRP retains its ability to bind PC, which is consistent with the previous findings by Williams (Williams, 2017).

The average value of mCRP determined in this study was considerably higher than those reported by Wang *et al.*, where the mean level of mCRP in AMI group (n=101) was found to be 0.021 mg/L and significantly higher than the control group (Wang *et al.*, 2015). In Wang *et al.* study, the pCRP levels of the analysed samples fall between 1.9-7.9 mg/L (Wang *et al.*, 2015). In our study, however, only the patients with a high level of pCRP (>100 mg/L) were enrolled. Despite no correlation being found between pCRP and mCRP levels (as shown in Figure 4-14), it is still possible that the higher levels of pCRP result in increasing the chance of pCRP binding to the exposed bioactive lipid present on the surface of damaged membranes and microparticles, a process which is necessary for the conversion of pCRP to mCRP (Eisenhardt *et al.*, 2009, Habersberger *et al.*, 2012). In agreement with this assumption, Wang *et al.* study show that a proportion of patients with high pCRP levels (six patients out of 42 with a pCRP level ranged between 3-285 mg/L) displayed increased mCRP levels (Wang *et al.*, 2015). In another study, Moran tested the presence of mCRP in the serum of patients with rheumatoid arthritis (RA) using a specific anti-human mCRP detection antibody (Moran, 2018). The results of mCRP in diabetic patients are comparable to those in RA patients of Moran's study, where the mean level of mCRP in RA group was 0.092 mg/L (range between 0.053 mg/L to 0.133 mg/L, n=30) with a significant difference from the healthy control participant (Moran, 2018). It worth mentioning that the average pCRP levels in the RA group of Moran study were also higher than the AMI group in Wang *et al.* study (22.85 mg/L vs 3.6 mg/L). On the other hand, the level of circulating mCRP determined in our study was considerably lower than the higher range reported by

Williams, who demonstrated mCRP levels fall between 0.04 mg/L and 3.52 mg/L in 40 participants identified with high sera pCRP levels (>100 mg/L). The mCRP within Williams' study was purified using the same methodology used in our study however, the underlying pathology of the enrolled participants was not known (Williams, 2017).

Inflammation is the feature uniting the studies of mCRP in the circulation. Numerous studies hypothesised that chronic low-grade inflammation plays a role in the pathogenesis of both type 1 and type 2 diabetes, and moderately increased levels of serum CRP are found in both types of diabetes (Schalkwijk *et al.*, 1999, Kilpatrick *et al.*, 2000, Pickup & Crook, 1998, Duncan *et al.*, 2003). The increased levels of serum CRP in diabetes are not caused by the development or presence of microvascular complications but might precede these (Schalkwijk *et al.*, 1999, Chase *et al.*, 2004). Hyperglycaemia itself or through the formation of AGE are known to stimulate the release of the inflammatory cytokines including TNF- $\alpha$  and IL-6 from adipocytes, monocytes and other cells. These released cytokines, in turn, can affect the production of CRP in the liver (Ingle & Patel, 2011, Goh & Cooper, 2008). Moreover, diabetic individuals are at high risk of developing other inflammatory conditions including atherosclerosis and heart disease correlated to high levels of CRP (Messier *et al.*, 2004, Lee *et al.*, 2015). Endothelial dysfunction and platelet activation together with increasing circulating microparticles are a common pathological features found in diabetic patients (Ingle & Patel, 2011, Ouviaña *et al.*, 2001, Tramontano *et al.*, 2010), and are known to dissociate CRP, and may be responsible for the mCRP occurring in these patients (Habersberger *et al.*, 2012, Eisenhardt *et al.*, 2009). The resulting mCRP can mediate actions that are different from that produced by pCRP through activation of Fc $\gamma$ RIII or integration into the lipid raft domain through a putative cholesterol binding domain that is hidden in the pentameric isoform (Ji *et al.*, 2009, Khreiss *et al.*, 2004). mCRP can induce pro-

inflammatory effects on cells such as monocytes and endothelial cells. More specifically, mCRP has been shown to bind principally to the apical of endothelial cells surfaces that are enriched in lipid raft domains, promoting the release of proinflammatory cytokines, generation of reactive oxygen species, and adhesion molecule expression (Ji *et al.*, 2009, Li *et al.*, 2014). Crawford *et al.* (2016) suggest that mCRP mediates chronic inflammation by modulating the polarisation of monocytes and T lymphocytes migrating across the endothelial layer, promoting a Th1 response and M1 macrophage differentiation (Crawford *et al.*, 2016).

It was expected to find a positive correlation between mCRP and pCRP levels. However, the results show that mCRP does not correlate with pCRP levels with, an uneven distribution throughout the cohort. Several samples show higher levels of mCRP regardless of their pCRP levels as compared to other samples in the cohort. It is possible that these participants have one or more other underlying conditions or undiagnosed disease which account for an independent increase of mCRP form. Diabetes mellitus is usually complicated by other conditions, such as obesity, dyslipidemia, chronic kidney disease, cardiovascular disease and others conditions, which may play a leading role in providing the essential or synergistic conditions that enhance the conversion of pCRP to mCRP or possibly the expression of mCRP rather than pCRP. Moreover, in this work, only diabetic patients with raised levels of pCRP have been enrolled (>100 mg/L), suggesting the presence of underlying acute inflammation or infectious diseases in those patients.

The absence of correlation between the circulating mCRP and the native pCRP was also reported in RA patients (Moran, 2018). In data published by Crawford *et al.* (2016), the circulating mCRP-bearing microparticles were also found not correlated to circulating pCRP

levels but strongly correlated with the amount of low-density lipoprotein cholesterol in patients with peripheral artery disease (Crawford *et al.*, 2016). Another study by Schwedler *et al.* (2003) shows increased tubular mCRP staining with declining renal function and the severity of the histological lesions in patients with diabetic nephropathy. However, no correlation was evident between the tubular mCRP staining and serum levels of CRP (Schwedler *et al.*, 2003), while in Wang *et al.* study, the circulating plasma mCRP levels were higher in deceased AMI patients than in living patients (Wang *et al.*, 2015). In diabetes mellitus, the levels of mCRP could be also correlated to the prognosis and complications associated with diabetes mellitus rather than the pCRP levels. However, it must mention that the small sample size of the study and the protein loss during the purification process, which could be uneven through the samples, could also affect the correlation between mCRP and pCRP levels.

Due to the potential for CRP losses during the purification process, the results of ELISA quantification of mCRP purified from diabetic serum samples could reflect mCRP lower than the actual levels in patient serum. Furthermore, the measured mCRP levels in this study represent the free unbound mCRP and did not count the mCRP bound to molecules such as MPs which could be lost during the chromatographic steps.

#### 4.4.4 Investigation of the glycated form of CRP using mass spectrometry

Mass spectrometry has been used to investigate the presence of a glycated form of CRP in serum samples of diabetic patients. The investigation was carried out at the intact level of the protein with LC-MS and the peptide level with LC-MS/MS. The results of intact mass analysis of all 24 human serum samples fail to show any evidence of glucose condensation on the protein in the form of mass increment of 162 Da per each attached glucose

molecule. Although the result of our *in vitro* study shows that CRP treated with 30 mM of glucose for a period equal to the half-life of the protein result in the formation of a mono-glycated isoform of CRP (see Figure 2-6). However, the stoichiometry of the reaction was low, where only 3.7% of the treated protein was mono-glycated isoform (see figure 2-9). The glucose concentration, even in the poorly controlled diabetic patients, is typically rather lower than 30mM, suggesting that the *in vivo* glycation processes of CRP is too low to induce a peak corresponding to glycated CRP isoform.

We have extended our investigation with trypsin and Glu-C enzymatic digestion of CRP containing samples with subsequent bottom-up tandem mass analysis of the resultant digestion products. Mascot search of the collected data against the customised database (comprising the CRP sequence and common laboratory-observed contaminants) and the manual inspection of the product ion spectra of the unmatched peptides could not identify any possible glycated peptide. The results of bottom-up tandem mass analysis of CRP from diabetic serum samples are consistent with our results of *in vitro* experiment; where the CRP incubated with glucose concentrations less than 100 mM for a period equal to the protein half-life in the circulation fail to show detectable glycated CRP peptides by tandem mass spectrometry (see table 3-4). However, when the incubation times were increased into three days, even with 30 mM of glucose, a single glycated peptide with glycation assigned to K13 was identified (see table 3-4). The glucose concentrations in diabetic patients and the exposure time of CRP to glucose which is limited by its half-life *in vivo* (Vigushin et al., 1993), were not enough to produce a detectable level of a glycated isoform of CRP.

Glycation reactions are, in general, relatively slow (Zhang *et al.*, 2008). Kinetic experiments on glycation of human serum albumin with glucose showed that the rate of formation of the Amadori product is 0.25% of the protein per day for each 1 mM of glucose (Baynes *et al.*, 1984), thus for the levels of glucose in well-controlled diabetes patients (<7.8 mM); around 1.9% of an individual protein could be glycated per day, so the longer the half-life of the protein the greater extent of glycation (Austin *et al.*, 1987). The rate of glycation is not consistent for different proteins; the rate of haemoglobin glycation is much slower than that of albumin (Garlick & Mazer, 1983), while complement C3, despite a short half-life of 2-5 days, shows moderate glycation as a result of its high intrinsic propensity for glycation (Austin *et al.*, 1987). For CRP, 3.7% of the protein is glycated *in vitro* when incubated at 30 mM glucose for 19h. If we hypothesise a linear relationship between the glucose concentration and extent of glycation of CRP, then <1% of the circulating CRP could be glycated *in vivo* when exposed to glucose levels of well-controlled diabetes patients (<7.8mM). This percentage (<1%) leads to a relatively small amount of glycated product when considering the low circulating concentration of CRP compared to other proteins such as albumin or complement C3 (Jolliff *et al.*, 1982).

Another possible reason for the failure to detect glycated CRP in the serum of diabetic patients is the significant CRP loss after the purification processes. The results of ELISA quantification of the pCRP fractions showed a significant protein loss post-purification compared to the initial serum samples concentration (see appendix 4). This loss is primarily due to the multiple chromatographic and concentrating steps to isolate the CRP isoforms. Furthermore, CRP might be lost during the purification process due to its binding to serum microvesicles (Tendler *et al.*, 2018, Van der Zee *et al.*, 2010). Serum microvesicles such as microparticles (small vesicles, with a diameter of 100–1000 nm, released from cells by

shedding) and exosomes (smaller membranous-derived vesicles secreted from intracellular multi-vesicular bodies and range in size from 30 to 90 nm) are known to contain substantial amount of CRP, where CRP bound to its natural ligand on the surface of these circulating microvesicles (Tendler *et al.*, 2018, Van der Zee *et al.*, 2010). The ligand-bound CRP on the surface of microvesicles will resist the purification process as it cannot bind to the PC-bound affinity column, therefore, will be lost during the affinity chromatography step. Another possible reason for failure to detect glycated CRP is that the glycation itself might affect the ability of CRP to bind to the PC-bound affinity column, however, this is less likely because the putative glycation sites, as determined in the *in vitro* glycation trails, were not involved in the phosphocholine binding site of CRP.

The short half-life of CRP *in vivo*, the glucose level in diabetic patients, the low abundance of CRP in serum compared to other proteins, and protein loss during the purification process may cumulatively account for CRP glycation remaining undetected.

#### 4.4.5 Conclusion

To our knowledge, there are no studies investigated the presence of modified forms of CRP (monomeric and/or glycated) in the serum of the patients with diabetes mellitus nor have quantified their presence. In the work described in this chapter, we demonstrate the presence of a freely circulating form of mCRP in the serum of diabetic patients with elevated levels of CRP. The identified mCRP has been demonstrated to be 23k Da in size via size exclusion chromatography and mass spectrometry. Furthermore, the circulating mCRP showed the ability to interact reversibly with phosphocholine, the natural ligand of pCRP, and this consistent with the recent finding of Williams (Williams, 2017). Both type 1 and 2 diabetes showed detectable levels of mCRP with no significant difference between the two

groups. However, due to the small sample size, especially with type 1 diabetic group, it was difficult to draw a conclusion and predict whether this finding would be observed within the larger population. mCRP showed broad range of values among the patients. Moreover, the circulating mCRP were not correlated with the pCRP level; while the same result was found in other studies (Crawford *et al.*, 2016, Schwedler *et al.*, 2003), the absence of correlation suggests that mCRP levels could be correlated to the diabetic complications or other underlying conditions that the patients may have rather than the pCRP itself.

We also investigated the presence of a glycated form of CRP in the serum of diabetic patients. We were unable to identify evidence for the presence of a glycated isoform in the serum of diabetic patients. While our *in vitro* study showed that CRP could be glycated after a 19h incubation period, the stoichiometry of the reaction was low, and glucose concentrations higher than those found in diabetic patients are required to produce a detectable level of glycated CRP isoform. The exposure time of CRP to glucose *in vivo* is limited by its half-life which is not enough to produce a detectable level of glycated CRP at glucose level of diabetic patients. This agrees with the results of the *in vitro* glycation trials where a longer incubation times or higher glucose concentrations are required to produce a detectable glycated CRP.



## Chapter 5: Discussion, Conclusion and Future Work

### 5.1 Overview

C-reactive protein is an acute phase protein that has long been employed in clinical practice as a marker for the presence of infection, inflammation, or tissue damage, with plasma levels increasing from baseline 1-2 mg/L to up to 1000 fold in the presence of inflammation (Ablij & Meinders, 2002, Hirschfield & Pepys, 2003). It is believed that elevated baseline levels (>1-2mg/L) can be a reliable indicator of the development of CVD (Pepys & Hirschfield, 2003). Increased baseline CRP levels have been linked to an increased risk of thrombotic events, including myocardial infarction (Ridker *et al.*, 2001, Kervinen *et al.*, 2001), and an increased risk of later development of diabetes (Barzilay *et al.*, 2001, Pradhan *et al.*, 2001). Furthermore, CRP levels are higher in people with diabetes compared with those without diabetes (Ford, 1999, Goldberg, 2000). Recent evidence suggests that CRP can dissociate to its monomeric subunit under physiological conditions after binding to the cell membrane (Ji *et al.*, 2007, Eisenhardt *et al.*, 2009). The dissociation has been linked with pro-inflammatory processes and the progression of the pathology commonly observed in inflammatory diseases (Wang *et al.*, 2015, Zhang *et al.*, 2012). Furthermore, the high level of glucose associated with diabetes mellitus results in glycation of serum proteins to varying extents, based on their half-life and glycability (Austin *et al.*, 1987). Protein glycation adversely affects the normal molecular structure, function and/or half-life (Brownlee, 2000, Cloos & Christgau, 2002, Stern *et al.*, 2002).

The two main aims of this research were to investigate the susceptibility of CRP to glycation *in vitro* and to clarify whether glycated and monomeric forms of human CRP were present,

detectable and quantifiable within serum samples obtained from participants with diabetes mellitus. The research presents experimental evidence contributing to a greater understanding of CRP isoforms within the human body.

## 5.2 Production & characterisation of *in vitro* glycated CRP

Protein glycation has been recognised as an important post-translational modification due to its impact on the structure and consequent potential to alter the physical and functional properties of the protein and its physiological effects (Rondeau & Bourdon, 2011, McCarthy *et al.*, 1998, Hunter *et al.*, 2003). To our knowledge, no studies have been conducted so far to illustrate the effect of this reaction on CRP. The findings presented in chapters 2 and 3 of this thesis shed light on CRP glycation *in vitro* with regard to the degree of glycation with different glucose concentrations and incubation times, the putative glycation sites, the microenvironment of the potential glycation sites (pKa, FAS, nearby acidic and basic residues), and the most susceptible residues toward glycation.

To elucidate susceptibility of CRP toward glycation with glucose, CRP was exposed to different concentrations of glucose (30-1000 mM) for periods ranging from 19 h to 6 days, at 37 °C and pH 7.4. The glycation reaction was monitored by mass spectrometry as a function of a net increase in the mass of intact protein due to the addition of glucose, a principle that has been utilised in several previous studies dealing with the analysis of protein glycation by mass spectrometry (Farah *et al.*, 2005, Lapolla *et al.*, 1993, Goodarzi *et al.*, 2004). The results of this *in vitro* study showed that CRP could be glycated, and higher states of glycation (number of attached glucose/CRP molecule) can be achieved with higher glucose concentrations as well as longer incubation times. Most importantly, CRP exposed to 30 mM glucose for a period equal to its half-life *in vivo* resulted in formation of a mono-

glycated isoform. While this glucose level (30 mM) is still far from the normal physiological levels, or the levels in well controlled diabetic patients, it is not uncommon to find such glucose levels in poorly controlled diabetic patients (McCarthy *et al.*, 1998). Identifying mono-glycated isoforms when treating CRP with 30 mM glucose for 19 h suggests the possibility of CRP glycation in patients with poorly controlled diabetes. However, based on the peak intensities of the MS spectra, the stoichiometry of the reaction is relatively low, where only 3.7% of the protein is estimated to be glycated. It is less likely that this degree of glycation could influence the biological functionality of CRP mainly because of the short exposure time of the protein to the glucose *in vivo* (Vigushin *et al.*, 1993). Further glycation of CRP could be achieved, during the period of the protein half-life, by exposing the protein to higher glucose levels; increasing the glucose concentration beyond 30 mM results in a parallel increase in the extent of glycation to reach 48.8% of the unmodified protein when exposed to 1 M glucose level. It was observed that CRP favours the formation of mono-glycated isoform despite increasing the glucose concentration to 1 M and extending the exposure time to 3 days, while other glycated isoforms with multiple condensed glucose molecules were also detected, their intensities were lower compared to mono-glycated CRP. When the incubation time was further increased to 6 days, a shift in the distribution of CRP glycoforms toward the highly glycated forms was observed in parallel with the increase in glucose concentration. Almost all the protein (97.6%) was glycated when exposed to 1 M glucose for 6 days, with an average of 3.6 glucose molecule/ CRP molecule, distributed between a mixture of species with varying glycation levels ranging from mono-glycated to octa-glycated variants. It is clear from the result of this study that CRP glycation could be detected on the intact protein by mass spectrometry, which offered a direct and reliable approach for monitoring CRP glycation and provides an image of sample

heterogeneity. The results of this study also show that with increased concentration of glucose and incubation time, both the extent of CRP glycation and site occupancy increase.

To elucidate the putative glycation sites, we have extended our study with trypsin and Glu-C enzymatic digestion of the *in vitro* glycosylated CRP samples with subsequent bottom-up tandem mass spectrometric analysis of the resultant digestion product. In agreement with a previous study (Taylor & Van Den Berg, 2007), CRP was found to be relatively resistant to enzymatic digestion; additional denaturation steps, higher CRP: enzyme molar ratio and longer incubation times with proteases were required to improve its degradation. Tandem mass analysis of the *in vitro* glycosylated CRP samples, by Q-TOF, produced characteristic fragmentation spectra for glycosylated peptides as a result of multiple neutral losses from the glucose-containing parent and product ions. This phenomenon has been reported previously with other glycosylated peptides analysed by CID (Lapolla *et al.*, 2004, Frolov *et al.*, 2006, Mennella *et al.*, 2006). The most common and intense ions detected in the MS/MS spectra of the glycosylated peptides were those arising from neutral loss of 3 H<sub>2</sub>O (pyrylium ions) and 3 H<sub>2</sub>O + HCHO (furylium ions) from the glycosylated parent and fragment ions, observations of these ions was strongly suggestive of the presence of a glycosylated peptide. The neutral loss signals in the product ion spectra of the glycosylated peptides were associated with weak or even undetectable glycosylated b- and y-ions which in turn result in an ambiguous assignment of the modified amino acid residue by Mascot, especially when multiple potential glycation sites were present on the glycosylated peptide.

Ambiguous modification site assignment by Mascot was revised and augmented by manual interpretation of the signals arising from successive neutral losses, including the strong pyrylium and furylium signals. Identifying the glycosylated peptide using Mascot followed by

manual interpretation/validation of the product ion spectrum allowed identification of the majority of the peptide backbone ions and positive demonstration of the site of modification. Furthermore, all peptides which Mascot failed to match were considered as possible glycosylated peptides and interpreted manually. The robust methodology used in this study allowed identification of three glycosylation sites in CRP preparation exposed to 1 M glucose for 19 hours these are K13, K57 and K69, in addition to one glycosylation sites among six possible residues within the peptide FWVDGKPRVRKSLKKGTVGAE. Occupation of the same glycosylation sites was also observed after increasing the incubation time to 3 and 6 days, suggesting that these residues are favourable glycosylation sites at this glucose concentration. After 3 days of incubation with 1M glucose, K191 (also seen after 6 days) and K23 were modified, while K201 and two residues of the peptide FWVDGKPRVRKSLKKGTVGAE were observed as being modified but after 6 days of incubation. This finding suggests that glycosylation is not randomly occurring, and the free amino groups of CRP differ in their rates of glycosylation (Zhang *et al.*, 2008). All the identified glycosylation sites were lysine residues, and the common feature between the glycosylated lysines is the proximity to an acidic residue, principally situated C-terminally. For the most reactive lysine residues that have been detected after 19h of incubation (i.e. K13, K57 and K69); the distance between the  $\epsilon$ -NH<sup>+</sup> group of the glycosylated residue to the nearest carboxylate group was < 3.6 Å. Positively, the number of glycosylation sites identified *in vitro* glycosylated sample by tandem mass spectrometry is comparable to the number of condensed glucose molecules on the intact protein as determined by mass spectrometry. Among the identified glycosylation sites, K201 is involved in a salt bridge with E101 from the next protomer (Shrive *et al.*, 1996); thus its glycosylation could affect the structure of pCRP. We were unable to precisely specify which residues in FWVDGKPRVRKSLKKGTVGAE are glycosylated due to the limited sequence information in the

fragmentation spectrum, as a result of neutral losses and the localised high positive charge density in the middle of the peptide which further affects its fragmentation (Xue *et al.*, 2015). Based on the combined contribution of pKa and FAS to the lysine microenvironment, K114 has the highest probability of being modified compared to other possible glycation sites in FWVDGK<sub>14</sub>PRVRKSLKKGTVGAE. Moreover, the proximity of D112 to K114 in the primary sequence can accelerate the Amadori rearrangement by accepting the proton from the Schiff base (Rabbani *et al.*, 2016). K114 is part of putative C1q-binding site of CRP and plays a role in modulating the reactivity of CRP with C1q (Agrawal *et al.*, 2001, Agrawal & Volanakis, 1994). Thus, K114 glycation could have an impact on the CRP function. It was observed that K13 is the most reactive residue toward glycation compared to other identified glycation sites in CRP; K13 was the only modified residue in CRP exposed to 100 mM of glucose for 19 h. CRP incubated at glucose levels less than 100 mM for 19 h showed no evidence of glycation, suggesting that the extent of glycation was insufficient to be detected by tandem mass analysis.

### 5.3 Analysis of human diabetic serum samples

Serum samples were obtained from 24 participants with diabetes. These were tested for the presence of a monomeric and glycated form of CRP. All the tested samples were with high levels of pCRP (>100 mg/L) in order to maximise the chance of identifying CRP isoforms in the serum. A detectable level of mCRP was found across all the tested samples which included both type 1 and type 2 diabetes patients. To our knowledge, this is the first report for the presence of mCRP in the serum of diabetic patients, and one of the few studies of identifying a soluble circulating form of mCRP (Williams, 2017, Wang *et al.*, 2015, Moran, 2018). The result of this study is consistent with the previous findings, providing further

proof of the presence of a soluble circulating form of mCRP. A previous study by Wang *et al.* (2015) reported the presence of mCRP in the plasma of patients with myocardial infarction (Wang *et al.*, 2015); likewise, more recent work by Moran detected mCRP in the serum of patients with rheumatoid arthritis (Moran, 2018). Both studies considered a specific anti-human mCRP antibody for detection. However, this raises a question of whether the identified mCRP is circulating free within the serum or bound to microparticles (Habersberger *et al.*, 2012). In this work, mCRP was purified from serum by affinity and size exclusion chromatography before quantification by ELISA, thus confirming that the identified mCRP is a freely circulating isoform with reactivity against CRP antiserum, and molecular size of 23 kDa as further shown by intact mass analysis. The methodology used in our study to purify mCRP was adapted from Williams (2017) who identified the presence of mCRP in the serum of the patients with raised levels of CRP and unknown underlying disorders (Williams, 2017). In agreement with the Williams study (Williams, 2017), mCRP shows the ability to bind to PC, which further contributes towards the evidence supporting the ability of mCRP to bind to PC-containing cell membranes *in vivo* (Habersberger *et al.*, 2012, Eisenhardt *et al.*, 2009). The monoclonal anti-CRP antibody (clone 8) used in the ELISA experiment was able to detect both mCRP and pCRP isoforms; therefore attention must be considered when explaining the results of standard ELISA tests or the automated systems used in clinical practice for measuring serum CRP levels, as they are likely unable to distinguish between the two isoforms (Williams, 2017). The literature suggests a pro-inflammatory role for mCRP (Eisenhardt *et al.*, 2009), with others suggesting increased levels of mCRP is related to the disease progression as in myocardial infarction (Wang *et al.*, 2015), and diabetic nephropathy (Schwedler *et al.*, 2003), therefore, distinguishing between the two isoforms is important and may result in better understanding of the

patient condition. No correlation was found between mCRP and pCRP levels in patients with diabetes mellites. The lack of correlation was also reported by other studies (Moran, 2018, Schwedler *et al.*, 2003, Habersberger *et al.*, 2012), suggesting that other underlying conditions or undiagnosed disease may account for an independent increase of mCRP levels in some patients either through enhancing the conversion of pCRP to mCRP or increasing the expression of mCRP rather than pCRP. It was found that mCRP levels correlated more with the severity of the clinical condition rather than pCRP level. (Schwedler *et al.*, 2003, Wang *et al.*, 2015).

The levels of mCRP determined in this study may represent lower than the actual levels in serum, this is because there is a potential for mCRP loss through the methodology used which involved several chromatographic and concentration steps. Furthermore, the mCRP measured in this study represent the freely soluble mCRP in serum and not counting the ligand-bound mCRP on the surface of circulating microvesicles such as MPs, which preferably lost during the purification process.

Mass spectrometric analysis of pCRP failed to identify a glycated isoform of the protein in the serum samples of diabetic patients, suggesting that any *in vivo* glycation of CRP is too low abundance to be detected by the methods/instrumentation used in our study. Our *in vitro* experiments reveal that the stoichiometry of the glycation reaction is likely to be low during the period of the protein half-life in the serum. The short half-life of the protein and the relatively low quantity of pCRP in the serum compared to other protein in the serum, besides the significant loss of the protein post purification process may account for lack of detection of glycated CRP in the serum of diabetic patients.



## 5.4 Conclusion and Future Work

The research described here, using ELISA and mass spectrometry, provides experimental evidence for a monomeric form and possibility of a glycated form of human CRP. The evidence from the *in vitro* glycation experiments implies the possibility of CRP glycation in patients with poorly controlled diabetes; however, at low stoichiometry and abundance, this could limit their detection. It is less likely that this degree of glycation influences the biological functionality of CRP *in vivo*. The glycation of CRP residues was not a random process; K13 was found more reactive for glycation when the glycating agent was reduced gradually in the reaction mixture. The reactivity of CRP lysine was found to be in the following order; K13> K57, K69 and one residue in the 109-130 peptide> K191 and K23> K201 and one residue in the 109-130 peptide.

Based on the evidence of the *in vitro* glycation experiments, the failure to identify a glycated CRP in the serum of diabetic patients is the short half-life of CRP *in vivo*, which is insufficient to produce detectable levels of glycated CRP at glucose level of diabetic patients. The significant loss of protein during the purification process could be an additional reason for this failure. Additional future work on CRP glycation in the serum of diabetic patients should involve fewer purification steps and the use of multiple reaction monitoring (MRM) on triple quadrupole (QqQ) for more selective, and sensitive detection of glycated peptides in the serum. More, having a method to glycate CRP *in vitro* and hence a potential method for generating polyclonal antibodies against glycated CRP, could enable specific ELISA and other methods to be developed.

The present study provides evidence for the presence of a freely circulating form of mCRP in the serum of diabetic patients (type 1 and 2) with elevated levels of CRP, thus adding to

the growing body of literature on the existence of serum mCRP. The molecular weight of mCRP was determined in this study by mass spectrometry and found to be consistent with the mass of CRP subunit. The levels of mCRP in diabetic patients were quantified and found not correlated to pCRP levels suggesting additional factors cause an independent increase of mCRP either through enhancing the conversion of pCRP to mCRP or increasing the expression of mCRP rather than pCRP. Factors such as diabetic complications, obesity, HbA1c levels, drug treatment or other underlying conditions may affect the levels of mCRP in diabetic patients; however, this research did not incorporate ethical approval for the correlation of mCRP with these factors. Also, the small sample size in this research did not allow reliable comparison between both types of diabetes. Future research should address these limitations. It is essential to determine what are the baselines levels of mCRP in larger group of population and what factors are involved in its rise in concentration. Furthermore what do the increased levels of mCRP correspond to, this lack of understanding, and the implications of the roles of mCRP in various pro-inflammatory scenarios, necessitates similar and further studies on mCRP.

## References

- Abernethy, T.J. & Avery, O.T. (1941). The occurrence during acute infections of a protein not normally present in the blood: Distribution of the reactive protein in patients' sera and the effect of calcium on the flocculation reaction with c polysaccharide of pneumococcus. *Journal of experimental medicine*. 73 (2): 173-182.
- Ablij, H.C. & Meinders, A.E. (2002). C-reactive protein: history and revival. *European journal of internal medicine*. 13 (7): 412-422.
- Aebersold, R. & Goodlett, D.R. (2001). Mass spectrometry in proteomics. *Chemical reviews*. 101 (2): 269-296.
- Aebersold, R. & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*. 422 (6928): 198.
- Agrawal, A. (2005). CRP after 2004. *Molecular immunology*. 42 (8): 927-930.
- Agrawal, A., Shrive, A.K., Greenhough, T.J. & Volanakis, J.E. (2001). Topology and structure of the C1q-binding site on C-reactive protein. *The Journal of Immunology*. 166 (6): 3998-4004.
- Agrawal, A. & Volanakis, J.E. (1994). Probing the C1q-binding site on human C-reactive protein by site-directed mutagenesis. *The Journal of Immunology*. 152 (11): 5404-5410.
- Aguiar, F.J., Ferreira-Júnior, M., Sales, M.M., Cruz-Neto, L.M., Fonseca, L.A., Sumita, N.M., Duarte, N.J., Lichtenstein, A. & Duarte, A.J. (2013). C-reactive protein: clinical applications and proposals for rational use. *Associação Médica Brasileira (English Edition)* 59.1 (2013): 85-92.
- Aguilar, Marie-Isabel. (2004) "Reversed-phase high-performance liquid chromatography." *HPLC of Peptides and Proteins*. Springer, Totowa, NJ. 9-22.
- Aguilar, M. & Hearn, M.T. (1996). High-resolution reversed-phase high-performance liquid chromatography of peptides and proteins. In: *Anonymous Methods in enzymology*. vol. 270. Elsevier.
- American Diabetes Association (2014). Diagnosis and classification of diabetes mellitus. *Diabetes care*. 37 (Supplement 1): S90.
- Ames, J.M. (2005). Application of semiquantitative proteomics techniques to the Maillard reaction. *Annals of the New York Academy of Sciences*. 1043 (1): 225-235.
- Anderson, H.C. & McCarty, M. (1950). Determination of C-reactive protein in the blood as a measure of the activity of the disease process in acute rheumatic fever. *The American Journal of Medicine*. 8 (4): 445-455.
- Anguizola, J., Matsuda, R., Barnaby, O.S., Hoy, K.S., Wa, C., DeBolt, E., Koke, M. & Hage, D.S. (2013). Glycation of human serum albumin. *Clinica Chimica Acta*. 425: 64-76.
- Ansari, N.A. & Ali, R. (2011). Glycated lysine residues: a marker for non-enzymatic protein glycation in age-related diseases. *Disease markers*. 30 (6): 317-324.

- Aoki, T., Hiidome, Y., Kitahata, K., Sugimoto, Y., Ibrahim, H.R. & Kato, Y. (1999). Improvement of heat stability and emulsifying activity of ovalbumin by conjugation with glucuronic acid through the Maillard reaction. *Food Research International*. 32 (2): 129-133.
- Aragno, M. & Mastrocola, R. (2017). Dietary sugars and endogenous formation of advanced glycation endproducts: emerging mechanisms of disease. *Nutrients*. 9 (4): 385.
- Arasteh, A., Farahi, S., Habibi-Rezaei, M. & Moosavi-Movahedi, A.A. (2014). Glycated albumin: an overview of the *in vitro* models of an *in vivo* potential disease marker. *Journal of Diabetes & Metabolic Disorders*. 13 (1): 49.
- Arima, H., Kubo, M., Yonemoto, K., Doi, Y., Ninomiya, T., Tanizaki, Y., Hata, J., Matsumura, K., Iida, M. & Kiyohara, Y. (2008). High-sensitivity C-reactive protein and coronary heart disease in a general population of Japanese: the Hisayama study. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 28 (7): 1385-1391.
- Austin, G.E., Mullins, R.H. & Morin, L.G. (1987). Non-enzymic glycation of individual plasma proteins in normoglycemic and hyperglycemic patients. *Clinical chemistry*. 33 (12): 2220-2224.
- Avery, N.C. & Bailey, A.J. (2005). Enzymic and non-enzymic cross-linking mechanisms in relation to turnover of collagen: relevance to ageing and exercise. *Scandinavian Journal of Medicine & Science in Sports*. 15 (4): 231-240.
- Bai, X., Wang, Z., Huang, C., Wang, Z. & Chi, L. (2012). Investigation of non-enzymatic glycosylation of human serum albumin using ion trap-time of flight mass spectrometry. *Molecules*. 17 (8): 8782-8794.
- Barzilay, J.I., Abraham, L., Heckbert, S.R., Cushman, M., Kuller, L.H., Resnick, H.E. & Tracy, R.P. (2001). The relation of markers of inflammation to the development of glucose disorders in the elderly: the Cardiovascular Health Study. *Diabetes*. 50 (10): 2384-2389.
- Bass, J.J., Wilkinson, D.J., Rankin, D., Phillips, B.E., Szewczyk, N.J., Smith, K. & Atherton, P.J. (2017). An overview of technical considerations for Western blotting applications to physiological research. *Scandinavian Journal of Medicine & Science in Sports*. 27 (1): 4-25.
- Baynes, J.W., Thorpe, S.R. & Murtiashaw, M.H. (1984). Nonenzymatic glycosylation of lysine residues in albumin. In: *Anonymous Methods in enzymology*. vol. 106. Elsevier.
- Bereman, M.S., MacLean, B., Tomazela, D.M., Liebler, D.C. & MacCoss, M.J. (2012). The development of selected reaction monitoring methods for targeted proteomics via empirical refinement. *Proteomics*. 12 (8): 1134-1141.
- Bharadwaj, D., Stein, M., Volzer, M., Mold, C. & Du Clos, T.W. (1999). The major receptor for C-reactive protein on leukocytes is Fcγ receptor II. *Journal of Experimental Medicine*. 190 (4): 585-590.
- Biemann, K. (1992). Mass spectrometry of peptides and proteins. *Annual Review of Biochemistry*. 61 (1): 977-1010.
- Bisoendial, R., Kastelein, J. & Stroes, E. (2005). Letter to the Editor: In response to van den Berg *et al.* *Circulation research*. 97 (7): e71.

- Black, S., Kushner, I. & Samols, D. (2004). C-reactive protein. *Journal of Biological Chemistry*. 279 (47): 48487-48490.
- Blancher, C. and Jones, A., 2001. SDS-PAGE and Western blotting techniques. In *Metastasis research protocols* (pp. 145-162). Humana Press.
- Brady, L.J., Martinez, T. & Balland, A. (2007). Characterization of nonenzymatic glycation on a monoclonal antibody. *Analytical Chemistry*. 79 (24): 9403-9413.
- Brock, J.W., Hinton, D.J., Cotham, W.E., Metz, T.O., Thorpe, S.R., Baynes, J.W. & Ames, J.M. (2003). Proteomic analysis of the site specificity of glycation and carboxymethylation of ribonuclease. *Journal of proteome research*. 2 (5): 506-513.
- Brownlee, M. (2000). Negative consequences of glycation. *Metabolism*. 49 (2): 9-13.
- Btissam, R., Rajae, R., Hassane, G., Naima, B. & Mohamed, N. (2017). Protocol implementation for the detection of advanced glycation end products and screening of natural extracts involved in the prevention of diabetic complications. *Journal of Medicinal Plants*. 5 (3): 4.
- Bubb, W.A. (2003). NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts in Magnetic Resonance Part A: An Educational Journal*. 19 (1): 1-19.
- Čaparević, Z. & Kostić, N. (2007). The influence of age and the beginning of menopause on the lipid status, LDL oxidation, and CRP in healthy women. *Srpski arhiv za celokupno lekarstvo*. 135 (5-6): 280-285.
- Capuzzi, D.M. & Freeman, J.S. (2007). C-reactive protein and cardiovascular risk in metabolic syndrome and type 2 diabetes: controversy and challenge. *Clinical Diabetes*. 25 (1): 16-22.
- Casey, E.B., Zhao, H. & Abraham, E.C. (1995). Role of glycine 1 and lysine 2 in the glycation of bovine  $\gamma$ b-crystallin site-directed mutagenesis of lysine to threonine. *Journal of Biological Chemistry*. 270 (35): 20781-20786.
- Castano, L. & Eisenbarth, G.S. (1990). Type-I diabetes: a chronic autoimmune disease of human, mouse, and rat. *Annual Review of Immunology*. 8 (1): 647-679.
- Cermak, J., Key, N.S., Bach, R.R., Balla, J., Jacob, H.S. & Vercellotti, G.M. (1993). C-reactive protein induces human peripheral blood monocytes to synthesize tissue factor. *Blood*. 82 (2): 513-520.
- Chang, M., Binder, C.J., Torzewski, M. & Witztum, J.L. (2002). C-reactive protein binds to both oxidized LDL and apoptotic cells through recognition of a common ligand: phosphorylcholine of oxidized phospholipids. *Proceedings of the National Academy of Sciences*. 99 (20): 13043-13048.
- Chase, H.P., Cooper, S., Osberg, I., Stene, L.C., Barriga, K., Norris, J., Eisenbarth, G.S. & Rewers, M. (2004). Elevated C-reactive protein levels in the development of type 1 diabetes. *Diabetes*. 53 (10): 2569-2573.

- Chen, Y., Liang, L., Liu, X., Labuza, T.P. & Zhou, P. (2012). Effect of fructose and glucose on glycation of  $\beta$ -lactoglobulin in an intermediate-moisture food model system: Analysis by liquid chromatography-mass spectrometry (LC-MS) and data-independent acquisition LC-MS (LC-MSE). *Journal of Agricultural and Food Chemistry*. 60 (42): 10674-10682.
- Chernushevich, I.V., Loboda, A.V. & Thomson, B.A. (2001). An introduction to quadrupole-time-of-flight mass spectrometry. *Journal of mass spectrometry*. 36 (8): 849-865.
- Chew, K.S. (2012). What's new in Emergencies Trauma and Shock? C-reactive protein as a potential clinical biomarker for influenza infection: More questions than answers. *Journal of emergencies, trauma, and shock*. 5 (2): 115.
- Cho, S., Roman, G., Yeboah, F. & Konishi, Y. (2007). The road to advanced glycation end products: a mechanistic perspective. *Current medicinal chemistry*. 14 (15): 1653-1671.
- Choi, Y. & Lim, S. (2009). Characterization of anti-advanced glycation end product antibodies to nonenzymatically lysine-derived and arginine-derived glycated products. *Journal of Immunoassay and Immunochemistry*. 30 (4): 386-399.
- Christner, R.B. & Mortensen, R.F. (1994). Binding of human serum amyloid P-component to phosphocholine. *Archives of Biochemistry and Biophysics*. 314 (2): 337-343.
- Cimmino, G., Ragni, M., Cirillo, P., Petrillo, G., Loffredo, F., Chiariello, M., Gresele, P., Falcinelli, E. & Golino, P. (2013). C-reactive protein induces expression of matrix metalloproteinase-9: a possible link between inflammation and plaque rupture. *International journal of cardiology*. 168 (2): 981-986.
- Cirillo, P., Golino, P., Calabrò, P., Calì, G., Ragni, M., De Rosa, S., Cimmino, G., Pacileo, M., De Palma, R. & Forte, L. (2005). C-reactive protein induces tissue factor expression and promotes smooth muscle and endothelial cell proliferation. *Cardiovascular research*. 68 (1): 47-55.
- Ciubotaru, I., Potempa, L.A. & Wander, R.C. (2005). Production of modified C-reactive protein in U937-derived macrophages. *Experimental biology and medicine*. 230 (10): 762-770.
- Cloos, P.A. & Christgau, S. (2002). Non-enzymatic covalent modifications of proteins: mechanisms, physiological consequences and clinical applications. *Matrix Biology*. 21 (1): 39-52.
- Craig, R. & Beavis, R.C. (2004). TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 20 (9): 1466-1467.
- Crawford, J.R., Trial, J., Nambi, V., Hoogeveen, R.C., Taffet, G.E. & Entman, M.L. (2016). Plasma levels of endothelial microparticles bearing monomeric C-reactive protein are increased in peripheral artery disease. *Journal of cardiovascular translational research*. 9 (3): 184-193.
- Cray, C., Zaias, J. & Altman, N.H. (2009). Acute phase response in animals: a review. *Comparative medicine*. 59 (6): 517-526.
- Creasy, D.M. & Cottrell, J.S. (2002). Error-tolerant searching of uninterpreted tandem mass spectrometry data. *Proteomics*. 2 (10): 1426-1434.

- Culley, F.J., Bodman-Smith, K.B., Ferguson, M.A., Nikolaev, A.V., Shantilal, N. & Raynes, J.G. (2000). C-reactive protein binds to phosphorylated carbohydrates. *Glycobiology*. 10 (1): 59-65.
- Das, T., Mandal, C. & Mandal, C. (2003). Variations in binding characteristics of glycosylated human C-reactive proteins in different pathological conditions. *Glycoconjugate journal*. 20 (9): 537-543.
- Dasu, M.R., Devaraj, S., Du Clos, T.W. & Jialal, I. (2007). The biological effects of CRP are not attributable to endotoxin contamination: evidence from TLR4 knockdown human aortic endothelial cells. *Journal of lipid research*. 48 (3): 509-512.
- De Beer, F.C. & Pepys, M.B. (1982). Isolation of human C-reactive protein and serum amyloid P component. *Journal of immunological methods*. 50 (1): 17-31.
- De Hoffmann, E. (2000). Mass spectrometry. *Kirk-Othmer Encyclopedia of Chemical Technology*.
- Deodhar, S.D. (1989). C-reactive protein: the best laboratory indicator available for monitoring disease activity. *Cleveland Clinic journal of medicine*. 56 (2): 126-130.
- Devaraj, S., Singh, U. & Jialal, I. (2009). Human C-reactive protein and the metabolic syndrome. *Current opinion in lipidology*. 20 (3): 182.
- Devaraj, S., Xu, D.Y. & Jialal, I. (2003). C-reactive protein increases plasminogen activator inhibitor-1 expression and activity in human aortic endothelial cells: implications for the metabolic syndrome and atherothrombosis. *Circulation*. 107 (3): 398-404.
- Devaraj, S., Yun, J., Adamson, G., Galvez, J. & Jialal, I. (2009). C-reactive protein impairs the endothelial glycocalyx resulting in endothelial dysfunction. *Cardiovascular research*. 84 (3): 479-484.
- Dietrich, M. & Jialal, I. (2005). The effect of weight loss on a stable biomarker of inflammation, C-reactive protein. *Nutrition reviews*. 63 (1): 22-28.
- Du Clos, T.W. (2013). *Pentraxins: structure, function, and role in inflammation*. ISRN inflammation. 2013.
- Du Clos, T.W. & Mold, C. (2011). Pentraxins (CRP, SAP) in the process of complement activation and clearance of apoptotic bodies through Fcγ receptors. *Current opinion in organ transplantation*. 16 (1): 15.
- Duncan, B.B., Schmidt, M.I., Pankow, J.S., Ballantyne, C.M., Couper, D., Vigo, A., Hoogeveen, R., Folsom, A.R. & Heiss, G. (2003). Low-grade systemic inflammation and the development of type 2 diabetes: the atherosclerosis risk in communities study. *Diabetes*. 52 (7): 1799-1805.
- Eble, A.S., Thorpe, S.R. & Baynes, J.W. (1983). Nonenzymatic glycosylation and glucose-dependent cross-linking of protein. *Journal of Biological Chemistry*. 258 (15): 9406-9412.
- Eisenhardt, S.U., Habersberger, J., Murphy, A., Chen, Y., Woollard, K.J., Bassler, N., Qian, H., von zur Muhlen, C., Hagemeyer, C.E. & Ahrens, I. (2009). Dissociation of pentameric to

monomeric C-reactive protein on activated platelets localizes inflammation to atherosclerotic plaques. *Circulation research*. 105 (2): 128-137.

Eisenhardt, S.U., Habersberger, J., Oliva, K., Lancaster, G.I., Ayhan, M., Woollard, K.J., Bannasch, H., Rice, G.E. & Peter, K. (2011). A proteomic analysis of C-reactive protein stimulated THP-1 monocytes. *Proteome science*. 9 (1): 1.

Eisenhardt, S.U., Habersberger, J. & Peter, K. (2009). Monomeric C-reactive protein generation on activated platelets: the missing link between inflammation and atherothrombotic risk. *Trends in cardiovascular medicine*. 19 (7): 232-237.

Eisenhardt, S.U., Thiele, J.R., Bannasch, H., Stark, G.B. & Peter, K. (2009). C-reactive protein: how conformational changes influence inflammatory properties. *Cell cycle*. 8 (23): 3885-3892.

El-Aneed, A., Cohen, A. & Banoub, J. (2009). Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews*. 44 (3): 210-230.

Eng, J.K., McCormack, A.L. & Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry*. 5 (11): 976-989.

Enguix, A., Rey, C., Concha, A., Medina, A., Coto, D. & Diéguez, M.A. (2001). Comparison of procalcitonin with C-reactive protein and serum amyloid for the early diagnosis of bacterial sepsis in critically ill neonates and children. *Intensive care medicine*. 27 (1): 211-215.

Ens, W. & Standing, K.G. (2005). Hybrid quadrupole/time-of-flight mass spectrometers for analysis of biomolecules. *Methods in enzymology*. 402: 49-78.

Farah, M.A., Bose, S., Lee, J., Jung, H. & Kim, Y. (2005). Analysis of glycated insulin by MALDI-TOF mass spectrometry. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 1725 (3): 269-282.

Faraj, M. and Salem, N., 2012. C-reactive protein. In *Blood Cell-An Overview of Studies in Hematology*. IntechOpen.

Ford, E.S. (1999). Body mass index, diabetes, and C-reactive protein among US adults. *Diabetes Care*. 22 (12): 1971-1977.

Fournet, M., Bonté, F. & Desmoulière, A. (2018). Glycation damage: A possible hub for major pathophysiological disorders and ageing. *Ageing and disease*. 9 (5): 880.

Fowler, M.J. (2008). Microvascular and macrovascular complications of diabetes. *Clinical diabetes*. 26 (2): 77-82.

Freeman, D.J., Norrie, J., Caslake, M.J., Gaw, A., Ford, I., Lowe, G.D., O'Reilly, D.S.J., Packard, C.J. & Sattar, N. (2002). C-reactive protein is an independent predictor of risk for the development of diabetes in the West of Scotland Coronary Prevention Study. *Diabetes*. 51 (5): 1596-1600.

Friedman, D.B., Hoving, S. & Westermeier, R. (2009). Isoelectric focusing and two-dimensional gel electrophoresis. In: *Anonymous Methods in enzymology*. vol. 463. Elsevier.



- Fröhlich, M., Imhof, A., Berg, G., Hutchinson, W.L., Pepys, M.B., Boeing, H., Muche, R., Brenner, H. & Koenig, W. (2000). Association between C-reactive protein and features of the metabolic syndrome: a population-based study. *Diabetes Care*. 23 (12): 1835-1839.
- Frolov, A., Hoffmann, P. & Hoffmann, R. (2006). Fragmentation behaviour of glycosylated peptides derived from d-glucose, d-fructose and d-ribose in tandem mass spectrometry. *Journal of mass spectrometry*. 41 (11): 1459-1469.
- Gabray, C. & Kushner, I. (1999). Acute phase proteins and other systemic responses to inflammation [Erratum in. *N Engl J Med*. 340 (1376): 448-454.
- Gadgil, H.S., Bondarenko, P.V., Pipes, G., Rehder, D., McAuley, A., Perico, N., Dillon, T., Ricci, M. & Treuheit, M. (2007). The LC/MS analysis of glycation of IgG molecules in sucrose containing formulations. *Journal of pharmaceutical sciences*. 96 (10): 2607-2621.
- Gadgil, H.S., Bondarenko, P.V., Treuheit, M.J. & Ren, D. (2007). Screening and sequencing of glycosylated proteins by neutral loss scan LC/MS/MS method. *Analytical Chemistry*. 79 (15): 5991-5999.
- Galván-Moroyoqui, J.M., Candia-Plata, M.C., Martínez-Soto, J.M., Soto-Guzmán, J.A., Camacho-Villa, A.Y. & López-Soto, L.F. (2016). Glycosylated ferritin induces activation and expression of tlr2 and tlr4 in human peripheral blood macrophages. *The Pharma Innovation*. 4 (1, Part A).
- Garlick, R.L. & Mazer, J.S. (1983). The principal site of nonenzymatic glycosylation of human serum albumin *in vivo*. *Journal of Biological Chemistry*. 258 (10): 6142-6146.
- Gershov, D., Kim, S., Brot, N. & Elkon, K.B. (2000). C-Reactive protein binds to apoptotic cells, protects the cells from assembly of the terminal complement components, and sustains an antiinflammatory innate immune response: implications for systemic autoimmunity. *Journal of Experimental Medicine*. 192 (9): 1353-1364.
- Ghosh, S., Pandey, N.K., Roy, A.S., Tripathy, D.R., Dinda, A.K. & Dasgupta, S. (2013). Prolonged glycation of hen egg white lysozyme generates non amyloid structures. *PLoS One*. 8 (9): e74336.
- Glish, G.L. & Burinsky, D.J. (2008). Hybrid mass spectrometers for tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry*. 19 (2): 161-172.
- Glish, G.L. & Vachet, R.W. (2003). The basics of mass spectrometry in the twenty-first century. *Nature reviews Drug discovery*. 2 (2): 140.
- Goetze, A.M., Liu, Y.D., Arroll, T., Chu, L. & Flynn, G.C. (2011). Rates and impact of human antibody glycation *in vivo*. *Glycobiology*. 22 (2): 221-234.
- Goh, S. & Cooper, M.E. (2008). The role of advanced glycation end products in the progression and complications of diabetes. *The Journal of Clinical Endocrinology & Metabolism*. 93 (4): 1143-1152.
- Goldberg, R.B. (2000). Cardiovascular disease in diabetic patients. *Medical Clinics of North America*. 84 (1): 81-93.

- Good, D.M., Wirtala, M., McAlister, G.C. & Coon, J.J. (2007). Performance characteristics of electron transfer dissociation mass spectrometry. *Molecular & Cellular Proteomics*. 6 (11): 1942-1951.
- Goodarzi, M.T., Ghahramany, S. & Mirmomeni, M.H. (2004). *In vitro* glycation of human IgG and its effect on interaction with anti-IgG. *Iranian Journal of Allergy, Asthma and Immunology*. 3 (4): 181-188.
- Grillo, M.A. & Colombatto, S. (2008). Advanced glycation end-products (AGEs): involvement in ageing and in neurodegenerative diseases. *Amino acids*. 35 (1): 29-36.
- Grundy, S.M., Brewer Jr, H.B., Cleeman, J.I., Smith Jr, S.C. & Lenfant, C. (2004). Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. *Circulation*. 109 (3): 433-438.
- Habersberger, J., Strang, F., Scheichl, A., Htun, N., Bassler, N., Merivirta, R., Diehl, P., Krippner, G., Meikle, P. & Eisenhardt, S.U. (2012). Circulating microparticles generate and transport monomeric C-reactive protein in patients with myocardial infarction. *Cardiovascular research*. 96 (1): 64-72.
- Hames, B.D. (1998). *Gel electrophoresis of proteins: a practical approach*. vol. 197. OUP Oxford.
- Hammer, M.R., John, P.N., Flynn, M.D., Bellingham, A.J. & Leslie, R. (1989). Glycated fibrinogen: a new index of short-term diabetic control. *Annals of Clinical Biochemistry*. 26 (1): 58-62.
- Handa, A. & Kuroda, N. (1999). Functional improvements in dried egg white through the Maillard reaction. *Journal of Agricultural and Food Chemistry*. 47 (5): 1845-1850.
- Hart, S.R. & Gaskell, S.J. (2005). LC-tandem MS in proteome characterization. *Trends in Analytical Chemistry*. 24 (7): 566-575.
- Heaton, R.J., Raynes, J.G. & Johnston, D.S. (1999). A study of the denaturation of human C-reactive protein in the presence of calcium ions and glycerophosphorylcholine. *Thermochimica acta*. 334 (1-2): 97-106.
- Heinecke, J.W. (2006). Lipoprotein oxidation in cardiovascular disease: chief culprit or innocent bystander? *Journal of Experimental Medicine*. 203 (4): 813-816.
- Helsens, K., Martens, L., Vandekerckhove, J. & Gevaert, K. (2007). MascotDatfile: an open-source library to fully parse and analyse MASCOT MS/MS search results. *Proteomics*. 7 (3): 364-366.
- Hex, N., Bartlett, C., Wright, D., Taylor, M. & Varley, D. (2012). Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine*. 29 (7): 855-862.
- Hilton, G.R. & Benesch, J.L. (2012). Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of The Royal Society Interface*, rsif20110823.

- Hirschfield, G.M. & Pepys, M.B. (2003). C-reactive protein and cardiovascular disease: new insights from an old molecule. *Qjm*. 96 (11): 793-807.
- Hogarth, M.B., Gallimore, J.R., Savage, P., Palmer, A.J., Starr, J.M., Bulpitt, C.J. & Pepys, M.B. (1997). Acute phase proteins, C-reactive protein and serum amyloid A protein, as prognostic markers in the elderly inpatient. *Age and Ageing*. 26 (2): 153-158.
- Holčapek, M., Jirásko, R. & Lísa, M. (2012). Recent developments in liquid chromatography-mass spectrometry and related techniques. *Journal of Chromatography A*. 1259: 3-15.
- Holmes, M.R. & Giddings, M.C. (2004). Prediction of posttranslational modifications using intact-protein mass spectrometric data. *Analytical Chemistry*. 76 (2): 276-282.
- Hua, L., Zhou, R., Thirumalai, D. & Berne, B.J. (2008). Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proceedings of the National Academy of Sciences*. 105 (44): 16928-16933.
- Huang, H.Z., Nichols, A. & Liu, D. (2009). Direct identification and quantification of aspartyl succinimide in an IgG2 mAb by RapiGest assisted digestion. *Analytical Chemistry*. 81 (4): 1686-1692.
- Huffman, F.G., Gomez, G.P. & Zarini, G.G. (2009). Metabolic syndrome and high-sensitivity C-reactive protein in Cubans. *Ethnicity & disease*. 19 (2): 115.
- Hunter, S.J., Boyd, A.C., O'Harte, F.P., McKillop, A.M., Wiggam, M.I., Mooney, M.H., McCluskey, J.T., Lindsay, J.R., Ennis, C.N. & Gamble, R. (2003). Demonstration of glycosylated insulin in human diabetic plasma and decreased biological activity assessed by euglycemic-hyperinsulinemic clamp technique in humans. *Diabetes*. 52 (2): 492-498.
- Hustoft, H.K., Malerod, H., Wilson, S.R., Reubsæet, L., Lundanes, E. & Greibrokk, T. (2012). A critical review of trypsin digestion for LC-MS based proteomics. In: Anonymous Integrative Proteomics. InTech.
- Hyogo, H. & Yamagishi, S. (2008). Advanced glycation end products (AGEs) and their involvement in liver disease. *Current pharmaceutical design*. 14 (10): 969-972.
- Imhof, A., Fröhlich, M., Loewel, H., Helbecque, N., Woodward, M., Amouyel, P., Lowe, G.D. & Koenig, W. (2003). Distributions of C-reactive protein measured by high-sensitivity assays in apparently healthy men and women from different populations in Europe. *Clinical chemistry*. 49 (4): 669-672.
- Ingle, P.V. & Patel, D.M. (2011). C-reactive protein in various disease condition—an overview. *Asian J Pharm Clin Res*. 4 (1): 9-13.
- Ito, S., Nakahari, T. & Yamamoto, D. (2011). The structural feature surrounding glycosylated lysine residues in human haemoglobin. *Biomedical Research*. 32 (3): 217-223.
- Jabs, W.J., Theissing, E., Nitschke, M., Bechtel, J.M., Duchrow, M., Mohamed, S., Jahrbeck, B., Sievers, H., Steinhoff, J. & Bartels, C. (2003). Local generation of C-reactive protein in diseased coronary artery venous bypass grafts and normal vascular tissue. *Circulation*. 108 (12): 1428-1431.

- Jahouh, F., Xu, P., Vann, W.F., Kováč, P. & Banoub, J.H. (2013). Mapping the glycation sites in the neoglycoconjugate from hexasaccharide antigen of *Vibrio cholerae*, serotype Ogawa and the recombinant tetanus toxin C-fragment carrier. *Journal of Mass Spectrometry*. 48 (10): 1083-1090.
- Jakuš, V. & Rietbrock, N. (2004). Advanced glycation end-products and the progress of diabetic vascular complications. *Physiological research*. 53 (2): 131-142.
- Janeway Jr, C.A. & Medzhitov, R. (2002). Innate immune recognition. *Annual Review of Immunology*. 20 (1): 197-216.
- Janeway, C.A., Travers, P., Walport, M. and Shlomchik, M., 1996. *Immunobiology: the immune system in health and disease* (Vol. 7, p. 26). London: Current Biology.
- Jarva, H., Jokiranta, T.S., Hellwage, J., Zipfel, P.F. & Meri, S. (1999). Regulation of complement activation by C-reactive protein: targeting the complement inhibitory activity of factor H by an interaction with short consensus repeat domains 7 and 8–11. *The Journal of Immunology*. 163 (7): 3957-3962.
- Jaye, D.L. & Waites, K.B. (1997). Clinical applications of C-reactive protein in paediatrics. *The Pediatric infectious disease journal*. 16 (8): 735-747.
- Jerić, I., Versluis, C., Horvat, Š & Heck, A.J. (2002). Tracing glycoprotein structures: electrospray ionization tandem mass spectrometric analysis of sugar-peptide adducts. *Journal of mass spectrometry*. 37 (8): 803-811.
- Ji, S., Ma, L., Bai, C., Shi, J., Li, H., Potempa, L.A., Filep, J.G., Zhao, J. & Wu, Y. (2009). Monomeric C-reactive protein activates endothelial cells via interaction with lipid raft microdomains. *The FASEB Journal*. 23 (6): 1806-1816.
- Ji, S., Wu, Y., Zhu, L., Potempa, L.A., Sheng, F., Lu, W. & Zhao, J. (2007). Cell membranes and liposomes dissociate C-reactive protein (CRP) to form a new, biologically active structural intermediate: mCRPm. *The FASEB Journal*. 21 (1): 284-294.
- Johansen, M.B., Kiemer, L. & Brunak, S. (2006). Analysis and prediction of mammalian protein glycation. *Glycobiology*. 16 (9): 844-853.
- Jolliff, C.R., Cost, K.M., Stivrins, P.C., Grossman, P.P., Nolte, C.R., Franco, S.M., Fijan, K.J., Fletcher, L.L. & Shriner, H.C. (1982). Reference intervals for serum IgG, IgA, IgM, C3, and C4 as determined by rate nephelometry. *Clinical chemistry*. 28 (1): 126-128.
- Jousilahti, P., Vartiainen, E., Salomaa, V., Harald, K., Seshasai, S.R., Kaptoge, S., Thompson, A., Di Angelantonio, E., Gao, P. & Sarwar, N. (2011). Diabetes mellitus, fasting glucose, and risk of cause-specific death.
- Juraschek, R., Dülcks, T. & Karas, M. (1999). Nanoelectrospray—more than just a minimized-flow electrospray ionization source. *Journal of the American Society for Mass Spectrometry*. 10 (4): 300-308.
- Kańska, U. & Boratyński, J. (2002). Thermal glycation of proteins by D-glucose and D-fructose. *Archivum Immunologiae et Therapiae Experimentalis*. 50 (1): 61-66.

- Kao, P.C., Shiesh, S. & Wu, T. (2006). Serum C-reactive protein as a marker for wellness assessment. *Annals of Clinical & Laboratory Science*. 36 (2): 163-169.
- Kervinen, H., Palosuo, T., Manninen, V., Tenkanen, L., Vaarala, O. & Mänttari, M. (2001). Joint effects of C-reactive protein and other risk factors on acute coronary events. *American Heart Journal*. 141 (4): 580-585.
- Khreiss, T., József, L., Potempa, L.A. & Filep, J.G. (2004). Conformational rearrangement in C-reactive protein is required for proinflammatory actions on human endothelial cells. *Circulation*. 109 (16): 2016-2022.
- Kiernan, U.A., Nedelkov, D., Tubbs, K.A., Niederkofler, E.E. & Nelson, R.W. (2004). Selected expression profiling of full-length proteins and their variants in human plasma. *Clinical Proteomics*. 1 (1): 7.
- Kilpatrick, E.S., Keevil, B.G., Jagger, C., Spooner, R.J. & Small, M. (2000). Determinants of raised C-reactive protein concentration in type 1 diabetes. *Qjm*. 93 (4): 231-236.
- Kim, M. & Pandey, A. (2012). Electron transfer dissociation mass spectrometry in proteomics. *Proteomics*. 12 (4-5): 530-542.
- King, D.E., Mainous, A.G., Buchanan, T.A. & Pearson, W.S. (2003). C-reactive protein and glycemic control in adults with diabetes. *Diabetes Care*. 26 (5): 1535-1539.
- Kinoshita, C.M., Ying, S.C., Hugli, T.E., Siegel, J.N., Potempa, L.A., Jiang, H., Houghten, R.A. & Gewurz, H. (1989). Elucidation of a protease-sensitive site involved in the binding of calcium to C-reactive protein. *Biochemistry*. 28 (25): 9840-9848.
- Koenig, R.J., Peterson, C.M., Jones, R.L., Saudek, C., Lehrman, M. & Cerami, A. (1976). Correlation of glucose regulation and haemoglobin A1c in diabetes mellitus. *New England Journal of Medicine*. 295 (8): 417-420.
- Kresl, J.J., Potempa, L.A. & Anderson, B.E. (1998). Conversion of native oligomeric to a modified monomeric form of human C-reactive protein. *The international journal of biochemistry & cell biology*. 30 (12): 1415-1426.
- Kushner, I., Rzewnicki, D. & Samols, D. (2006). What does minor elevation of C-reactive protein signify? *The American Journal of Medicine*. 119 (2): 166. e28.
- Lacinová, K., Eckhardt, A., Pataridis, S., Sedláková, P. & Mikšík, I. (2010). Glycation of proteins: their analysis and physiological aspects. *Advances in Molecular Mechanisms and Pharmacology of Diabetic Complications*, 17-38.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227 (5259): 680.
- Lagrand, W.K., Visser, C.A., Hermens, W.T., Niessen, H.W., Verheugt, F.W., Wolbink, G. & Hack, C.E. (1999). C-reactive protein as a cardiovascular risk factor: more than an epiphenomenon? *Circulation*. 100 (1): 96-102.
- Lal, S., Szwergold, B.S., Taylor, A.H., Randall, W.C., Kappler, F., Wellsknecht, K., Baynes, J.W. & Brown, T.R. (1995). Metabolism of fructose-3-phosphate in the diabetic rat lens. *Archives of Biochemistry and Biophysics*. 318 (1): 191-199.

- Lapice, E., Maione, S., Patti, L., Cipriano, P., Rivellese, A.A., Riccardi, G. & Vaccaro, O. (2009). Abdominal adiposity is associated with elevated C-reactive protein independent of BMI in healthy nonobese people. *Diabetes Care*. 32 (9): 1734-1736.
- Lapolla, A., Gerhardinger, C., Baldo, L., Fedele, D., Keane, A., Seraglia, R., Catinella, S. & Traldi, P. (1993). A study on *in vitro* glycation processes by matrix-assisted laser desorption ionization mass spectrometry. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*. 1225 (1): 33-38.
- Lapolla, A., Fedele, D., Reitano, R., Aricò, N.C., Seraglia, R., Traldi, P., Marotta, E. & Tonani, R. (2004). Enzymatic digestion and mass spectrometry in the study of advanced glycation end products/peptides. *Journal of the American Society for Mass Spectrometry*. 15 (4): 496-509.
- Lapolla, A., Molin, L. & Traldi, P. (2013). Protein glycation in diabetes as determined by mass spectrometry. *International journal of endocrinology*. 2013.
- Larsen, M.R., Trelle, M.B., Thingholm, T.E. & Jensen, O.N. (2006). Analysis of posttranslational modifications of proteins by tandem mass spectrometry: Mass Spectrometry For Proteomics Analysis. *BioTechniques*. 40 (6): 790-798.
- Lizzerini, P.E., Acampa, M., Capecchi, P.L., Fineschi, I., Selvi, E., Moscadelli, V., Zimbone, S., Gentile, D., Galeazzi, M. & Laghi-Pasini, F. (2015). Antiarrhythmic potential of anticytokine therapy in rheumatoid arthritis: tocilizumab reduces corrected QT interval by controlling systemic inflammation. *Arthritis care & research*. 67 (3): 332-339.
- Leblanc, Y., Bihoreau, N., Jube, M., Andre, M., Tellier, Z. & Chevreux, G. (2016). Glycation of polyclonal IgGs: Effect of sugar excipients during stability studies. *European Journal of Pharmaceutics and Biopharmaceutics*. 102: 185-190.
- Ledesma-Osuna, A.I., Ramos-Clamont, G. & Vázquez-Moreno, L. (2008). Characterization of bovine serum albumin glycated with glucose, galactose and lactose. *Acta Biochim Pol*. 55 (3): 491-497.
- Lee, B., Jayathilaka, G.L.P., Huang, J., Vida, L.N., Honig, G.R. & Gupta, S. (2011). Analyses of *in vitro* nonenzymatic glycation of normal and variant haemoglobins by MALDI-TOF mass spectrometry. *Journal of biomolecular techniques: JBT*. 22 (3): 90.
- Lee, S.I., Patel, M., Jones, C.M. & Narendran, P. (2015). Cardiovascular disease and type 1 diabetes: prevalence, prediction and management in an ageing population. *Therapeutic advances in chronic disease*. 6 (6): 347-374.
- Li, H., Wang, J., Wu, Y., Zhang, L., Liu, Z., Filep, J.G., Potempa, L.A., Wu, Y. & Ji, S. (2014). Topological localization of monomeric C-reactive protein determines pro-inflammatory endothelial cell responses. *Journal of Biological Chemistry*. : jbc. M114. 555318.
- Li, H., Robertson, A.D. & Jensen, J.H. (2005). Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Structure, Function, and Bioinformatics*. 61 (4): 704-721.
- Li, J. & Fang, C. (2004). C-reactive protein is not only an inflammatory marker but also a direct cause of cardiovascular diseases. *Medical Hypotheses*. 62 (4): 499-506.

- Li, R., Ren, M., Luo, M., Chen, N., Zhang, Z., Luo, B. & Wu, J. (2012). Monomeric C-reactive protein alters fibrin clot properties on endothelial cells. *Thrombosis research*. 129 (5): e256.
- Loo, J.A. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass spectrometry reviews*. 16 (1): 1-23.
- Luthra, M. & Balasubramanian, D. (1993). Nonenzymatic glycation alters protein structure and stability. A study of two eye lens crystallins. *Journal of Biological Chemistry*. 268 (24): 18119-18127.
- Lyons, T.J. & Jenkins, A.J. (1997). Glycation, oxidation, and lipoxidation in the development of the complications of diabetes: a carbonyl stress hypothesis. *Diabetes reviews (Alexandria, Va.)*. 5 (4): 365.
- Ma, X.J., Chen, H.B., Gao, J.Y., Hu, C.Q. & Li, X. (2013). Conformation affects the potential allergenicity of ovalbumin after heating and glycation. *Food Additives & Contaminants: Part A*. 30 (10): 1684-1692.
- Mahmood, T. & Yang, P. (2012). Western blot: technique, theory, and troubleshooting. *North American journal of medical sciences*. 4 (9): 429.
- Maillard, L.C. (1912). Action des acides amines sur les sucres; formation des melanoidines par voie methodique. *Comptes R.Acad.Sci.(Paris)*. 154 : 66-68.
- Manadas, B., Mendes, V.M., English, J. & Dunn, M.J. (2010). Peptide fractionation in proteomics approaches. *Expert review of proteomics*. 7 (5): 655-663.
- Mann, M. & Jensen, O.N. (2003a). Proteomic analysis of post-translational modifications. *Nature biotechnology*. 21 (3): 255.
- Marnell, L., Mold, C. & Du Clos, T.W. (2005). C-reactive protein: ligands, receptors and role in inflammation. *Clinical Immunology*. 117 (2): 104-111.
- McAlister, G.C., Phanstiel, D.H., Westphall, M.S. & Coon, J.J. (2011). Higher-energy collision-activated dissociation without a dedicated collision cell. *Molecular & Cellular Proteomics*. : mcp. O111. 009456.
- McCarthy, A.D., Cortizo, A.M., Segura, G.G., Bruzzone, L.C. & Etcheverry, S.B. (1998). Non-enzymatic glycosylation of alkaline phosphatase alters its biological. *Molecular and cellular biochemistry*. 181 (1-2): 63-69.
- McCarty, M. (1947). The occurrence during acute infections of a protein not normally present in the blood: IV. Crystallization of the C-reactive protein. *Journal of Experimental Medicine*. 85 (5): 491-498.
- McPherson, J.D., Shilton, B.H. & Walton, D.J. (1988). Role of fructose in glycation and cross-linking of proteins. *Biochemistry*. 27 (6): 1901-1907.
- Melbye, H., Halvorsen, D.S., Hartz, I., Medbø, A., Brox, J., Eggen, A.E. & Njølstad, I. (2007). Bronchial airflow limitation, smoking, body mass index, and statin use are strongly associated with the C-reactive protein level in the elderly.: The Troms Study 2001. *Respiratory medicine*. 101 (12): 2541-2549.

- Meltretter, J., Becker, C. & Pischetsrieder, M. (2008). Identification and site-specific relative quantification of  $\beta$ -lactoglobulin modifications in heated milk and dairy products. *Journal of Agricultural and Food Chemistry*. 56 (13): 5165-5171.
- Mennella, C., Visciano, M., Napolitano, A., Del Castillo, M.D. & Fogliano, V. (2006). Glycation of lysine-containing dipeptides. *Journal of peptide science: an official publication of the European Peptide Society*. 12 (4): 291-296.
- Messier, C., Awad, N. & Gagnon, M. (2004). The relationships between atherosclerosis, heart disease, type 2 diabetes and dementia. *Neurological research*. 26 (5): 567-572.
- Miller, A.K., Hambly, D.M., Kerwin, B.A., Treuheit, M.J. & Gadgil, H.S. (2011). Characterization of site-specific glycation during process development of a human therapeutic monoclonal antibody. *Journal of pharmaceutical sciences*. 100 (7): 2543-2550.
- Molins, B., Peña, E., de la Torre, R. & Badimon, L. (2011). Monomeric C-reactive protein is prothrombotic and dissociates from circulating pentameric C-reactive protein on adhered activated platelets underflow. *Cardiovascular research*. 92 (2): 328-337.
- Molins, B., Peña, E., Vilahur, G., Mendieta, C., Slevin, M. & Badimon, L. (2008). C-reactive protein isoforms differ in their effects on thrombus growth. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 28 (12): 2239-2246.
- Montero, I., Orbe, J., Varo, N., Beloqui, O., Monreal, J.I., Rodríguez, J.A., Díez, J., Libby, P. & Páramo, J.A. (2006). C-reactive protein induces matrix metalloproteinase-1 and-10 in human endothelial cells: implications for clinical and subclinical atherosclerosis. *Journal of the American College of Cardiology*. 47 (7): 1369-1378.
- Morais, M.P.P., Marshall, D., Flower, S.E., Caunt, C.J., James, T.D., Williams, R.J., Waterfield, N.R. & Van Den Elsen, Jean MH (2013). Analysis of protein glycation using fluorescent phenyl boronate gel electrophoresis. *Scientific reports*. 3: 1437.
- Moran, J., 2018. Investigations into the possible involvement of C-Reactive Protein (CRP) in the progression of rheumatoid arthritis, PhD thesis, Keele University.
- Morley, J.J. & Kushner, I. (1982). Serum C-reactive protein levels in disease. *Annals of the New York Academy of Sciences*. 389 (1): 406-418.
- Münch, G., Schickanz, D., Behme, A., Gerlach, M., Riederer, P., Palm, D. & Schinzel, R. (1999). Amino acid specificity of glycation and protein-AGE crosslinking reactivities determined with a dipeptide SPOT library. *Nature biotechnology*. 17 (10): 1006.
- Nanri, A., Moore, M.A. & Kono, S. (2007). Impact of C-reactive protein on disease risk and its relation to dietary factors: literature review. *Asian Pacific Journal of Cancer Prevention*. 8 (2): 167.
- Nawale, R.B., Mourya, V.K. & Bhise, S.B. (2006). Non-enzymatic glycation of proteins: a cause for complications in diabetes.
- Negre-Salvayre, A., Salvayre, R., Auge, N., Pamplona, R. & Portero-Otin, M. (2009). Hyperglycemia and glycation in diabetic complications. *Antioxidants & redox signaling*. 11 (12): 3071-3109.



- Nesvizhskii, A.I. & Aebersold, R. (2004). Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug discovery today*. 9 (4): 173-181.
- Niessen, W.M.A. (1999). *MS-MS and MSn*. Oxford: Academic Press.
- Nishimoto, N., Kishimoto, T. & Yoshizaki, K. (2000). Anti-interleukin 6 receptor antibody treatment in rheumatic disease. *Annals of the Rheumatic Diseases*. 59 (suppl 1): i27.
- Nord, F.F., Bier, M. & Terminiello, L. (1956). On the mechanism of enzyme action. LXI. The self-digestion of trypsin, calcium-trypsin and acetyltrypsin. *Archives of Biochemistry and Biophysics*. 65 (1): 120-131.
- Ohsawa, M., Okayama, A., Nakamura, M., Onoda, T., Kato, K., Itai, K., Yoshida, Y., Ogawa, A., Kawamura, K. & Hiramori, K. (2005). CRP levels are elevated in smokers but unrelated to the number of cigarettes and are decreased by long-term smoking cessation in male smokers. *Preventive medicine*. 41 (2): 651-656.
- O'Loughlin, J., Lambert, M., Karp, I., McGrath, J., Gray-Donald, K., Barnett, T.A., Delvin, E.E., Levy, E. & Paradis, G. (2008). Association between cigarette smoking and C-reactive protein in a representative, population-based sample of adolescents. *Nicotine & Tobacco Research*. 10 (3): 525-532.
- Olsen, J.V., Macek, B., Lange, O., Makarov, A., Horning, S. & Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. *Nature methods*. 4 (9): 709.
- Olufemi, S., Talwar, D. & Robb, D.A. (1987). The relative extent of glycation of haemoglobin and albumin. *Clinica chimica acta*. 163 (2): 125-136.
- Osman, R., L'Allier, P.L., Elgharib, N. & Tardif, J. (2006). Critical appraisal of C-reactive protein throughout the spectrum of cardiovascular disease. *Vascular health and risk management*. 2 (3): 221.
- Ouviña, S.M., La Greca, R.D., Zanaro, N.L., Palmer, L. & Sasseti, B. (2001). Endothelial dysfunction, nitric oxide and platelet activation in hypertensive and diabetic type II patients. *Thrombosis research*. 102 (2): 107-114.
- Owens III, A.P. & Mackman, N. (2011). Microparticles in hemostasis and thrombosis. *Circulation research*. 108 (10): 1284-1297.
- Packard, R.R. & Libby, P. (2008). Inflammation in atherosclerosis: from vascular biology to biomarker discovery and risk prediction. *Clinical chemistry*. 54 (1): 24-38.
- Parker, W.L., Stackiw, W. & Wilt, J.C. (1962). C-reactive protein in virus infection. *Canadian Medical Association journal*. 87 (15): 791.
- Pasceri, V., Chang, J., Willerson, J.T. & Yeh, E.T. (2001). Modulation of C-reactive protein-mediated monocyte chemoattractant protein-1 induction in human endothelial cells by anti-atherosclerosis drugs. *Circulation*. 103 (21): 2531-2534.
- Pasceri, V., Willerson, J.T. & Yeh, E.T. (2000). Direct proinflammatory effect of C-reactive protein on human endothelial cells. *Circulation*. 102 (18): 2165-2168.

- Pepys, M.B. & Hirschfield, G.M. (2003). C-reactive protein: a critical update. *The Journal of clinical investigation*. 111 (12): 1805-1812.
- Pepys, M.B., Hirschfield, G.M., Tennent, G.A., Gallimore, J.R., Kahan, M.C., Bellotti, V., Hawkins, P.N., Myers, R.M., Smith, M.D. & Polara, A. (2006). Targeting C-reactive protein for the treatment of cardiovascular disease. *Nature*. 440 (7088): 1217.
- Perkins, D.N., Pappin, D.J., Creasy, D.M. & Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis: An International Journal*. 20 (18): 3551-3567.
- Pickup, J.C. & Crook, M.A. (1998). Is type II diabetes mellitus a disease of the innate immune system? *Diabetologia*. 41 (10): 1241-1248.
- Potempa, L.A., Maldonado, B.A., Laurent, P., Zemel, E.S. & Gewurz, H. (1983). Antigenic, electrophoretic and binding alterations of human C-reactive protein modified selectively in the absence of calcium. *Molecular immunology*. 20 (11): 1165-1175.
- Potempa, L.A., Siegel, J.N., Fedel, B.A., Potempa, R.T. & Gewurz, H. (1987). Expression, detection and assay of a neoantigen (Neo-CRP) associated with a free, human C-reactive protein subunit. *Molecular immunology*. 24 (5): 531-541.
- Póvoa, P. (2002). C-reactive protein: a valuable marker of sepsis. *Intensive care medicine*. 28 (3): 235-243.
- Pradhan, A.D., Manson, J.E., Rifai, N., Buring, J.E. & Ridker, P.M. (2001). C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *Jama*. 286 (3): 327-334.
- Priego Capote, F. & Sanchez, J. (2009). Strategies for proteomic analysis of non-enzymatically glycosylated proteins. *Mass spectrometry reviews*. 28 (1): 135-146.
- Priego-Capote, F., Ramírez-Boo, M., Finamore, F., Gluck, F. & Sanchez, J. (2014). Quantitative analysis of glycosylated proteins. *Journal of proteome research*. 13 (2): 336-347.
- Quan, L. & Liu, M. (2013). CID, ETD and HCD fragmentation to study protein post-translational modifications. *Mod Chem Appl*. 1 (1): 1-5.
- Rabbani, N., Ashour, A. & Thornalley, P.J. (2016). Mass spectrometric determination of early and advanced glycation in biology. *Glycoconjugate journal*. 33 (4): 553-568.
- Rabbani, N. & Thornalley, P.J. (2012). No title. *Glycation research in amino acids: a place to call home*.
- Rabilloud, T., Vaezzadeh, A.R., Potier, N., Lelong, C., Leize-Wagner, E. & Chevallet, M. (2009). Power and limitations of electrophoretic separations in proteomics strategies. *Mass spectrometry reviews*. 28 (5): 816-843.
- Rahbar, S., Natarajan, R., Yerneni, K., Scott, S., Gonzales, N. & Nadler, J.L. (2000). Evidence that pioglitazone, metformin and pentoxifylline are inhibitors of glycation. *Clinica chimica acta*. 301 (1-2): 65-77.
- Ramadan, M.A., Shrive, A.K., Holden, D., Myles, D.A., Volanakis, J.E., DeLucas, L.J. & Greenhough, T.J. (2002). The three-dimensional structure of calcium-depleted human C-

reactive protein from perfectly twinned crystals. *Acta Crystallographica Section D*. 58 (6-2): 992-1001.

Rathore, D., Faustino, A., Schiel, J., Pang, E., Boyne, M. & Rogstad, S. (2018). The role of mass spectrometry in the characterization of biologic protein products. *Expert review of proteomics*. 15 (5): 431-449.

Reeves, G. (2007). C-reactive protein. *Australian Prescriber*. 30 (3): 74-76.

Ridker, P.M. (1998). C-reactive protein and risks of future myocardial infarction and thrombotic stroke.

Ridker, P.M. (2003). Rosuvastatin in the primary prevention of cardiovascular disease among patients with low levels of low-density lipoprotein cholesterol and elevated high-sensitivity C-reactive protein: rationale and design of the JUPITER trial. *Circulation*. 108 (19): 2292-2297.

Ridker, P.M., Bassuk, S.S. & Toth, P.P. (2003). C-reactive protein and risk of cardiovascular disease: evidence and clinical application. *Current atherosclerosis reports*. 5 (5): 341-349.

Ridker, P.M., Pare, G., Parker, A., Zee, R.Y., Danik, J.S., Buring, J.E., Kwiatkowski, D., Cook, N.R., Miletich, J.P. & Chasman, D.I. (2008). Loci related to metabolic-syndrome pathways including LEPR, HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study. *The American Journal of Human Genetics*. 82 (5): 1185-1192.

Ridker, P.M., Rifai, N., Rose, L., Buring, J.E. & Cook, N.R. (2002). Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *New England journal of medicine*. 347 (20): 1557-1565.

Ridker, P.M., Stampfer, M.J. & Rifai, N. (2001). Novel risk factors for systemic atherosclerosis: a comparison of C-reactive protein, fibrinogen, homocysteine, lipoprotein (a), and standard cholesterol screening as predictors of peripheral arterial disease. *Jama*. 285 (19): 2481-2485.

Ridker, P.M., Wilson, P.W. & Grundy, S.M. (2004). Should C-reactive protein be added to metabolic syndrome and to assessment of global cardiovascular risk? *Circulation*. 109 (23): 2818-2825.

Rifai, N. & Ridker, P.M. (2003). Population distributions of C-reactive protein in apparently healthy men and women in the United States: implication for clinical interpretation. *Clinical chemistry*. 49 (4): 666-669.

Roantree, R.J. & RANTZ, L.A. (1955). Clinical experience with the C-reactive protein test. *AMA archives of internal medicine*. 96 (5): 674-682.

Rodríguez-Hernández, H., Simental-Mendía, L.E., Rodríguez-Ramírez, G. & Reyes-Romero, M.A. (2013). Obesity and inflammation: epidemiology, risk factors, and markers of inflammation. *International journal of endocrinology*. 2013.

Rondeau, P. & Bourdon, E. (2011). The glycation of albumin: structural and functional impacts. *Biochimie*. 93 (4): 645-658.

- Rosales, C. & Uribe-Querol, E. (2013). Fc receptors: cell activators of antibody functions. *Advances in Bioscience and Biotechnology*. 4 (04): 21.
- Ross, R. (1993). The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature*. 362 (6423): 801.
- Rossky, P.J. (2008). Protein denaturation by urea: slash and bond. *Proceedings of the National Academy of Sciences*. 105 (44): 16825-16826.
- Rumley, A., Emberson, J.R., Wannamethee, S.G., Lennon, L., Whincup, P.H. & Lowe, G. (2006). Effects of older age on fibrin D-dimer, C-reactive protein, and other hemostatic and inflammatory variables in men aged 60–79 years. *Journal of thrombosis and haemostasis*. 4 (5): 982-987.
- Sacks, D.B. & McDonald, J.M. (1996). The pathogenesis of type II diabetes mellitus: a polygenic disease. *American Journal of Clinical Pathology*. 105 (2): 149-156.
- Salazar, J., Martínez, M.S., Chávez-Castillo, M., Núñez, V., Añez, R., Torres, Y., Toledo, A., Chacín, M., Silva, C. & Pacheco, E. (2014). C-reactive protein: an in-depth look into structure, function, and regulation. *International scholarly research notices*. 2014.
- Salzano, A.M. & Crescenzi, M. (2005). Mass spectrometry for protein identification and the study of post-translational modifications. *Ann Ist Super Sanita*. 41 (4): 443-450.
- Sarwar, N., Gao, P. & Seshasai, S.R. (2010). Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease [published correction appears in *Lancet*. 2010; 376 (9745): 958]. *Lancet*. 375 (9733): 2215-2222.
- Schalkwijk, C.G., Poland, D., Van Dijk, W., Kok, A., Emeis, J.J., Dräger, A.M., Doni, A., Van Hinsbergh, V. & Stehouwer, C. (1999). Plasma concentration of C-reactive protein is increased in type I diabetic patients without clinical macroangiopathy and correlates with markers of endothelial dysfunction: evidence for chronic inflammation. *Diabetologia*. 42 (3): 351-357.
- Schleicher, E. & Wieland, O.H. (1986). Kinetic analysis of glycation as a tool for assessing the half-life of proteins. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 884 (1): 199-205.
- Schwedler, S.B., Guderian, F., Dämmrich, J., Potempa, L.A. & Wanner, C. (2003). Tubular staining of modified C-reactive protein in diabetic chronic kidney disease. *Nephrology Dialysis Transplantation*. 18 (11): 2300-2307.
- Selvin, E., Paynter, N.P. & Erlinger, T.P. (2007). The effect of weight loss on C-reactive protein: a systematic review. *Archives of Internal Medicine*. 167 (1): 31-39.
- Sensi, M., Bruno, M.R., Valente, L., Cioccia, G.P., Chianelli, M. & Pozzilli, P. (1990). Retinol-binding protein: a short half-life determinant of protein non-enzymatic glycation in diabetes. *Diabetes research (Edinburgh, Scotland)*. 13 (4): 195-198.
- Seo, S., Karboune, S., L'Hocine, L. & Yaylayan, V. (2013). Characterization of glycated lysozyme with galactose, galactooligosaccharides and galactan: Effect of glycation on structural and functional properties of conjugates. *LWT-Food Science and Technology*. 53 (1): 44-53.

- Shapiro, R., McManus, M.J., Zalut, C. & Bunn, H.F. (1980). Sites of nonenzymatic glycosylation of human haemoglobin A. *Journal of Biological Chemistry*. 255 (7): 3120-3127.
- Sheng, B., Larsen, L.B., Le, T.T. & Zhao, D. (2018). Digestibility of Bovine Serum Albumin and Peptidomics of the Digests: Effect of Glycation Derived from  $\alpha$ -Dicarbonyl Compounds. *Molecules*. 23 (4): 712.
- Shrive, A.K., Gheetham, G.M., Holden, D., Myles, D.A., Turnell, W.G., Volanakis, J.E., Pepys, M.B., Bloomer, A.C. & Greenhough, T.J. (1996). Three dimensional structure of human C-reactive protein. *Nature Structural and Molecular Biology*. 3 (4): 346.
- Silva, A.M., Coimbra, J.T., Castro, M.M., Oliveira, A., Bras, N.F., Fernandes, P.A., Ramos, M.J. and Rangel, M., 2018. Determining the glycation site specificity of human holo-transferrin. *Journal of inorganic biochemistry*, 186, pp.95-102.
- Silva, A.M., Vitorino, R., Domingues, M.R.M., Spickett, C.M. & Domingues, P. (2013). Post-translational modifications and mass spectrometry detection. *Free radical biology and medicine*. 65: 925-941.
- Simpson, D.C., Ahn, S., Pasa-Tolic, L., Bogdanov, B., Mottaz, H.M., Vilkov, A.N., Anderson, G.A., Lipton, M.S. & Smith, R.D. (2006). Using size exclusion chromatography-RPLC and RPLC-CIEF as two-dimensional separation strategies for protein profiling. *Electrophoresis*. 27 (13): 2722-2733.
- Singh, U., Devaraj, S. & Jialal, I. (2005). C-reactive protein decreases tissue plasminogen activator activity in human aortic endothelial cells: evidence that C-reactive protein is a procoagulant. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 25 (10): 2216-2221.
- Singh, V.P., Bali, A., Singh, N. & Jaggi, A.S. (2014). Advanced glycation end products and diabetic complications. *The Korean Journal of Physiology & Pharmacology*. 18 (1): 1-14.
- Sjöberg, A.P., Trouw, L.A., McGrath, F.D., Hack, C.E. & Blom, A.M. (2006). Regulation of complement activation by C-reactive protein: targeting of the inhibitory activity of C4b-binding protein. *The Journal of Immunology*. 176 (12): 7612-7620.
- Sleno, L. & Volmer, D.A. (2004). Ion activation methods for tandem mass spectrometry. *Journal of mass spectrometry*. 39 (10): 1091-1112.
- Smith, P.R. & Thornalley, P.J. (1992). Influence of pH and phosphate ions on the kinetics of enolisation and degradation of fructosamines. Studies with the model fructosamine, N epsilon-1-deoxy-D-fructos-1-yl-hippuryl-lysine. *Biochemistry international*. 28 (3): 429-439.
- Sparvero, L.J., Asafu-Adjei, D., Kang, R., Tang, D., Amin, N., Im, J., Rutledge, R., Lin, B., Amoscato, A.A. & Zeh, H.J. (2009). RAGE (Receptor for Advanced Glycation Endproducts), RAGE ligands, and their role in cancer and inflammation. *Journal of translational medicine*. 7 (1): 17.
- Sell, D.R., Biemel, K.M., Reihl, O., Lederer, M.O., Strauch, C.M. & Monnier, V.M. (2005). Glucosepane is a major protein cross-link of the senescent human extracellular matrix relationship with diabetes. *Journal of Biological Chemistry*. 280 (13): 12310-12315.
- Spiller, S., Li, Y., Blüher, M., Welch, L. & Hoffmann, R. (2017). Glycated lysine-141 in haptoglobin improves the diagnostic accuracy for type 2 diabetes mellitus in combination

- with glycated hemoglobin HbA 1c and fasting plasma glucose. *Clinical proteomics*. 14 (1): 10.
- Spiro, R.G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*. 12 (4): 56R.
- Stastna, M. & Van Eyk, J.E. (2012). Analysis of protein isoforms: can we do it better? *Proteomics*. 12 (19-20): 2937-2948.
- Stern, D.M., Du Yan, S., Yan, S.F. & Schmidt, A.M. (2002). Receptor for advanced glycation endproducts (RAGE) and the complications of diabetes. *Ageing research reviews*. 1 (1): 1-15.
- Stumpe, M.C. & Grubmüller, H. (2007). Interaction of urea with amino acids: implications for urea-induced protein denaturation. *Journal of the American Chemical Society*. 129 (51): 16126-16131.
- Stumpf, C. & Hilgers, K.F. (2009). C-reactive protein: more than just a marker of inflammation? *Journal of hypertension*. 27 (9): 1748-1749.
- Sun, S., Zhou, J., Yang, W. & Zhang, H. (2014). Inhibition of protein carbamylation in urea solution using ammonium-containing buffers. *Analytical Biochemistry*. 446: 76-81.
- Sun, Y., Hayakawa, S., Ogawa, M. & Izumori, K. (2005). Evaluation of the site-specific protein glycation and antioxidant capacity of rare sugar– protein/peptide conjugates. *Journal of Agricultural and Food Chemistry*. 53 (26): 10205-10212.
- Swamy-Mruthinti, S. & Schey, K.L. (1997). Mass spectroscopic identification of *in vitro* glycated sites of MIP. *Current eye research*. 16 (9): 936-941.
- Swaney, D.L., Wenger, C.D. & Coon, J.J. (2010). Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *Journal of proteome research*. 9 (3): 1323-1329.
- Szalai, A.J. (2002). The biological functions of C-reactive protein. *Vascular pharmacology*. 39 (3): 105-107.
- Taylor, K.E., Giddings, J.C. & Van Den Berg, Carmen W (2005). C-reactive protein–induced *in vitro* endothelial cell activation is an artefact caused by azide and lipopolysaccharide. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 25 (6): 1225-1230.
- Taylor, K.E. & Van Den Berg, Carmen W (2007). Structural and functional comparison of native pentameric, denatured monomeric and biotinylated C-reactive protein. *Immunology*. 120 (3): 404-411.
- Tendler, Y., Aviram, M., Volkova, N., & Kaplan, M., (2018). Human Serum CRP – Rich Exosomes Fraction Increases the Activity of Macrophage Paraoxonase 2 (Pon2). *Annals of Atherosclerosis Research*, 1(1): 1-3.
- Ter Haar, R., Schols, H.A. & Gruppen, H. (2011). Effect of saccharide structure and size on the degree of substitution and product dispersity of  $\alpha$ -lactalbumin glycated via the Maillard reaction. *Journal of Agricultural and Food Chemistry*. 59 (17): 9378-9385.

Thiede, B., Höhenwarter, W., Krah, A., Mattow, J., Schmid, M., Schmidt, F. & Jungblut, P.R. (2005). Peptide mass fingerprinting. *Methods*. 35 (3): 237-247.

Thiele, J.R., Zeller, J., Bannasch, H., Stark, G.B., Peter, K. & Eisenhardt, S.U. (2015). Targeting C-reactive protein in inflammatory disease by preventing conformational changes. *Mediators of inflammation*. 2015.

Thompson, D., Pepys, M.B. & Wood, S.P. (1999). The physiological structure of human C-reactive protein and its complex with phosphocholine. *Structure*. 7 (2): 169-177.

Thornalley, P.J. (1999). Clinical significance of glycation. *Clin.Lab*. 45: 263-273.

Thornalley, P.J., Argirova, M., Ahmed, N., Mann, V.M., Argirov, O. & Dawnay, A. (2000). Mass spectrometric monitoring of albumin in uremia. *Kidney international*. 58 (5): 2228-2234.

Thornalley, P.J. (2006). Advanced glycation end products in renal failure. *Journal of renal nutrition*. 16 (3): 178-184.

Thornalley, P.J., Langborg, A. & Minhas, H.S. (1999). Formation of glyoxal, methylglyoxal and 3-deoxyglucosone in the glycation of proteins by glucose. *Biochemical Journal*. 344 (1): 109-116.

Thorpe, S.R. & Baynes, J.W. (2003). Maillard reaction products in tissue proteins: new products and new perspectives. *Amino acids*. 25 (3-4): 275-281.

Tillett, W.S. & Francis, T. (1930). Serological reactions in pneumonia with a non-protein somatic fraction of pneumococcus. *Journal of Experimental Medicine*. 52 (4): 561-571.

Timpson, N.J., Lawlor, D.A., Harbord, R.M., Gaunt, T.R., Day, I.N., Palmer, L.J., Hattersley, A.T., Ebrahim, S., Lowe, G.D. & Rumley, A. (2005). C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. *The Lancet*. 366 (9501): 1954-1959.

Tramontano, A.F., Lyubarova, R., Tsiakos, J., Palaia, T., Deleon, J.R. & Ragolia, L. (2010). Circulating endothelial microparticles in diabetes mellitus. *Mediators of inflammation*. 2010.

Tsiatsiani, L. & Heck, A.J. (2015). Proteomics beyond trypsin. *The FEBS journal*. 282 (14): 2612-2626.

Udvarnoki, K., Cervenak, L., Uray, K., Hudecz, F., Kacs Kovics, I., Spallek, R., Singh, M., Füst, G. & Prohászka, Z. (2007). Antibodies against C-reactive protein cross-react with 60-kilodalton heat shock proteins. *Clinical and Vaccine Immunology*. 14 (4): 335-341.

Ulrich, P. & Cerami, A. (2001). Protein glycation, diabetes, and ageing. *Recent progress in hormone research*. 56 (1): 1-22.

Urdal, P., Borch, S.M., Landaas, S., Krutnes, M.B., Gogstad, G.O. & Hjortdahl, P. (1992). Rapid immunometric measurement of C-reactive protein in whole blood. *Clinical chemistry*. 38 (4): 580-584.

Van der Zee, P.M, Biró, E., Trouw, L.A., Ko, Y., de Winter, R.J., Hack, C.E., Sturk, A. & Nieuwland, R. (2010). C-reactive protein in myocardial infarction binds to circulating

- microparticles but is not associated with complement activation. *Clinical Immunology*. 135 (3): 490-495.
- Vanzi, F., Madan, B. & Sharp, K. (1998). Effect of the protein denaturants urea and guanidinium on water structure: A structural and thermodynamic study. *Journal of the American Chemical Society*. 120 (41): 10748-10753.
- Venugopal, S.K., Devaraj, S. & Jialal, I. (2003). C-reactive protein decreases prostacyclin release from human aortic endothelial cells. *Circulation*. 108 (14): 1676-1678.
- Venugopal, S.K., Devaraj, S., Yuhanna, I., Shaul, P. & Jialal, I. (2002). Demonstration that C-reactive protein decreases eNOS expression and bioactivity in human aortic endothelial cells. *Circulation*. 106 (12): 1439-1441.
- Verma, S., Badiwala, M.V., Weisel, R.D., Li, S., Wang, C., Fedak, P.W., Li, R. & Mickle, D.A. (2003). C-reactive protein activates the nuclear factor- $\kappa$ B signal transduction pathway in saphenous vein endothelial cells: implications for atherosclerosis and restenosis. *The Journal of thoracic and cardiovascular surgery*. 126 (6): 1886-1891.
- Verma, S., Li, S., Badiwala, M.V., Weisel, R.D., Fedak, P.W., Li, R., Dhillon, B. & Mickle, D.A. (2002). Endothelin antagonism and interleukin-6 inhibition attenuate the proatherogenic effects of C-reactive protein. *Circulation*. 105 (16): 1890-1896.
- Verma, S., Szmítko, P.E. & Yeh, E.T. (2004). C-reactive protein: structure affects function. *Circulation*. 109 (16): 1914-1917.
- Verma, S., Wang, C., Li, S., Dumont, A.S., Fedak, P.W., Badiwala, M.V., Dhillon, B., Weisel, R.D., Li, R. & Mickle, D.A. (2002). A self-fulfilling prophecy: C-reactive protein attenuates nitric oxide production and inhibits angiogenesis. *Circulation*. 106 (8): 913-919.
- Vermeire, S., Van Assche, G. & Rutgeerts, P. (2004). C-reactive protein as a marker for inflammatory bowel disease. *Inflammatory bowel diseases*. 10 (5): 661-665.
- Verrastro, I., Pasha, S., Jensen, K.T., Pitt, A.R. & Spickett, C.M. (2015). Mass spectrometry-based methods for identifying oxidized proteins in disease: advances and challenges. *Biomolecules*. 5 (2): 378-411.
- Vidyasagar, S., Abdul Razak, U.K., Prashanth, C.K., Muralidhar Varma, D. & Bairy, K.L. (2013). Highly sensitive C-reactive protein in metabolic syndrome. *JACM*. 14 (3-4): 230-234.
- Vigushin, D.M., Pepys, M.B. & Hawkins, P.N. (1993). Metabolic and scintigraphic studies of radioiodinated human C-reactive protein in health and disease. *The Journal of clinical investigation*. 91 (4): 1351-1357.
- Visser, M., Bouter, L.M., McQuillan, G.M., Wener, M.H. & Harris, T.B. (1999). Elevated C-reactive protein levels in overweight and obese adults. *Jama*. 282 (22): 2131-2135.
- Vlassara, H. (1994). Pathogenic effects of advanced glycosylation: biochemical, biologic, and clinical implications for diabetes and ageing. *Lab Invest*. 70: 138-151.
- Volanakis, J.E. (2001). Human C-reactive protein: expression, structure, and function. *Molecular immunology*. 38 (2-3): 189-197.



- Volanakis, J.E., Clements, W.L. & Schrohenloher, R.E. (1978). C-reactive protein: purification by affinity chromatography and physicochemical characterization. *Journal of immunological methods*. 23 (3-4): 285-295.
- Volanakis, J.E. & Kaplan, M.H. (1971). Specificity of C-reactive protein for choline phosphate residues of pneumococcal C-polysaccharide. *Proceedings of the Society for Experimental Biology and Medicine*. 136 (2): 612-614.
- Volanakis, J.E. & Wirtz, K.W. (1979). Interaction of C-reactive protein with artificial phosphatidylcholine bilayers. *Nature*. 281 (5727): 155.
- Wa, C., Cerny, R.L., Clarke, W.A. & Hage, D.S. (2007). Characterization of glycation adducts on human serum albumin by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clinica Chimica Acta*. 385 (1-2): 48-60.
- Wang, H., Wu, Y., Chen, Y. & Sui, S. (2002). Polymorphism of structural forms of C-reactive protein. *International journal of molecular medicine*. 9 (6): 665-671.
- Wang, H., Tu, Z., Liu, G., Liu, C., Huang, X. & Xiao, H. (2013). Comparison of glycation in conventionally and microwave-heated ovalbumin by high-resolution mass spectrometry. *Food Chemistry*. 141 (2): 985-991.
- Wang, J., Tang, B., Liu, X., Wu, X., Wang, H., Xu, D. & Guo, Y. (2015). Increased monomeric CRP levels in acute myocardial infarction: a possible new and specific biomarker for diagnosis and severity assessment of disease. *Atherosclerosis*. 239 (2): 343-349.
- Wang, M.S., Messersmith, R.E. & Reed, S.M. (2012). Membrane curvature recognition by C-reactive protein using lipoprotein mimics. *Soft Matter*. 8 (30): 7909-7918.
- Wang, S., Wang, T., Wu, C. & Chen, S. (2014). In-depth comparative characterization of haemoglobin glycation in normal and diabetic blood by LC-MSMS. *Journal of the American Society for Mass Spectrometry*. 25 (5): 758-766.
- Wang, Y. & Ho, C. (2012). Flavour chemistry of methylglyoxal and glyoxal. *Chemical Society Reviews*. 41 (11): 4140-4149.
- Wannamethee, S.G., Lowe, G.D., Shaper, A.G., Rumley, A., Lennon, L. & Whincup, P.H. (2005). Associations between cigarette smoking, pipe/cigar smoking, and smoking cessation, and haemostatic and inflammatory markers for cardiovascular disease. *European heart journal*. 26 (17): 1765-1773.
- Watkins, N.G., Thorpe, S.R. & Baynes, J.W. (1985). Glycation of amino groups in protein. Studies on the specificity of modification of RNase by glucose. *Journal of Biological Chemistry*. 260 (19): 10629-10636.
- Welsh, K.J., Kirkman, M.S. & Sacks, D.B. (2016). Role of glycated proteins in the diagnosis and management of diabetes: research gaps and future directions. *Diabetes Care*. 39 (8): 1299-1306.
- Wijetunge, D. & Perera, H. (2014). A novel *in vitro* method to identify protein glycation inhibitors. *Asia Journal of Medical Sciences*. 5 (3): 15-21.

- Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D. & Wishart, D.S. (2003). VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic acids research*. 31 (13): 3316-3319.
- Williams, R.D. (2017). Characterisation of molecular variations in human C-reactive protein. *Characterisation of molecular variations in human C-reactive protein*.
- World Health Organization (2016). *Global report on diabetes*.
- Wu, V. & Steward, L.A. (1991). Purification of glycated haemoglobin free of haemoglobin A1c and its use to produce monoclonal antibodies specific for deoxyfructosyllysine residues in glycohemoglobin. *Biochemical and biophysical research communications*. 176 (1): 207-212.
- Wu, Y., Potempa, L.A., El Kebir, D. & Filep, J.G. (2015). C-reactive protein and inflammation: conformational changes affect function. *Biological chemistry*. 396 (11): 1181-1197.
- Xie, F., Smith, R.D. & Shen, Y. (2012). Advanced proteomic liquid chromatography. *Journal of Chromatography A*. 1261: 78-90.
- Xue, G., Liu, Z., Wang, L. & Zu, L. (2015). The role of basic residues in the fragmentation process of the lysine-rich cell-penetrating peptide TP10. *Journal of Mass Spectrometry*. 50 (1): 220-227.
- Yamada, S., Gotoh, T., Nakashima, Y., Kayaba, K., Ishikawa, S., Nago, N., Nakamura, Y., Itoh, Y. & Kajii, E. (2001). Distribution of serum C-reactive protein and its association with atherosclerotic risk factors in a Japanese population: Jichi Medical School Cohort Study. *American Journal of Epidemiology*. 153 (12): 1183-1190.
- Yasojima, K., Schwab, C., McGeer, E.G. & McGeer, P.L. (2001). Generation of C-reactive protein and complement components in atherosclerotic plaques. *The American journal of pathology*. 158 (3): 1039-1051.
- Yaylayan, V.A., Huyghues-Despointes, A. & Feather, M.S. (1994). Chemistry of Amadori rearrangement products: analysis, synthesis, kinetics, reactions, and spectroscopic properties. *Critical Reviews in Food Science & Nutrition*. 34 (4): 321-369.
- Yeboah, F.K. & Yaylayan, V.A. (2001). Analysis of glycated proteins by mass spectrometric techniques: qualitative and quantitative aspects. *Food/Nahrung*. 45 (3): 164-171.
- Yoshiuchi, K., Matsuhisa, M., Katakami, N., Nakatani, Y., Sakamoto, K., Matsuoka, T., Umayahara, Y., Kosugi, K., Kaneto, H. & Yamasaki, Y. (2008). Glycated albumin is a better indicator for glucose excursion than glycated haemoglobin in type 1 and type 2 diabetes. *Endocrine journal*, 0804280122.
- Zhang, H. & Ge, Y. (2011). Comprehensive analysis of protein modifications by top-down mass spectrometry. *Circulation: Genomic and Precision Medicine*. 4 (6): 711.
- Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A. & Ma, B. (2012). PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Molecular & Cellular Proteomics*. 11 (4): M111. 010587.

- Zhang, L., Eugeni, E.E., Parthun, M.R. & Freitas, M.A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma*. 112 (2): 77-86.
- Zhang, Q., Ames, J.M., Smith, R.D., Baynes, J.W. & Metz, T.O. (2008). A perspective on the Maillard reaction and the analysis of protein glycation by mass spectrometry: probing the pathogenesis of chronic disease. *Journal of proteome research*. 8 (2): 754-769.
- Zhang, Q., Frolov, A., Tang, N., Hoffmann, R., van de Goor, T., Metz, T.O. & Smith, R.D. (2007). Application of electron transfer dissociation mass spectrometry in analyses of non-enzymatically glycosylated peptides. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*. 21 (5): 661-666.
- Zhang, Q., Petyuk, V.A., Schepmoes, A.A., Orton, D.J., Monroe, M.E., Yang, F., Smith, R.D. & Metz, T.O. (2008). Analysis of non-enzymatically glycosylated peptides: neutral-loss-triggered MS3 versus multi-stage activation tandem mass spectrometry. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*. 22 (19): 3027-3034.
- Zhang, Q., Tang, N., Brock, J.W., Mottaz, H.M., Ames, J.M., Baynes, J.W., Smith, R.D. & Metz, T.O. (2007). Enrichment and analysis of nonenzymatically glycosylated peptides: boronate affinity chromatography coupled with electron-transfer dissociation mass spectrometry. *Journal of proteome research*. 6 (6): 2323-2330.
- Zhang, Q., Tang, N., Schepmoes, A.A., Phillips, L.S., Smith, R.D. & Metz, T.O. (2008). Proteomic profiling of nonenzymatically glycosylated proteins in human plasma and erythrocyte membranes. *Journal of proteome research*. 7 (5): 2025-2032.
- Zhang, Y., Fonslow, B.R., Shan, B., Baek, M. & Yates III, J.R. (2013). Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews*. 113 (4): 2343-2394.
- Zhang, Y., Cocklin, R.R., Bidasee, K.R. & Wang, M. (2003). Rapid determination of advanced glycation end products of proteins using MALDI-TOF-MS and PERL script peptide searching algorithm. *Journal of biomolecular techniques: JBT*. 14 (3): 224.
- Zhang, Z., Yang, Y., Hill, M.A. & Wu, J. (2012). Does C-reactive protein contribute to atherothrombosis via oxidant-mediated release of pro-thrombotic factors and activation of platelets? *Frontiers in physiology*. 3: 433.
- Zieske, A.W., Tracy, R.P., McMahan, C.A., Herderick, E.E., Homma, S., Malcom, G.T., McGill Jr, H.C. & Strong, J.P. (2005). Elevated serum C-reactive protein levels and advanced atherosclerosis in youth. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 25 (6): 1237-1243.
- Zou, Q., Habermann-Rottinghaus, S.M. & Murphy, K.P. (1998). Urea effects on protein stability: hydrogen bonding and the hydrophobic effect. *Proteins: Structure, Function, and Bioinformatics*. 31 (2): 107-111

## Appendix 1



**Health Research Authority**

Mr Sameer Mahmood  
School of Life Sciences  
Huxley Building  
Keele University, Keele, Staffordshire  
ST5 5BG

Email: [hra.approval@nhs.net](mailto:hra.approval@nhs.net)

25 April 2017


Dear Mr Mahmood,

**Letter of HRA Approval**

<b>Study title:</b>	<b>Are modified forms of C-reactive protein present in the serum of diabetic patients with elevated CRP levels?</b>
<b>IRAS project ID:</b>	<b>211529</b>
<b>Protocol number:</b>	<b>RG-0091-16-FNS</b>
<b>REC reference:</b>	<b>17/WM/0150</b>
<b>Sponsor</b>	<b>Keele University</b>

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

## Appendix 2

University Hospitals of North Midlands   
NHS Trust

University Hospitals of North Midlands

The Royal Stoke University Hospital

Newcastle Rd,

Stoke-on-Trent,

Staffordshire

ST4 6QG

01782 715444

## PATIENT INFORMATION SHEET

**Study Title:** < **Modifications to C-reactive protein in Diabetes Mellitus** >.

### Invitation

You are being invited to consider taking part in the research study <**Modifications to C-reactive protein in Diabetes Mellitus**>. This project is being undertaken by <Dr A Shrive, Mr S. Mahmood, Dr S. Hart, Prof T. Fryer and Prof T Greenhough>.

Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involve for you. One of the research team will go through the information sheet with you and will answer any questions you may have. We expect this should take no more than 15 minutes. If you are happy to participate we will ask you to sign a written consent form to acknowledge this.

### Aims of the Research

The high blood sugar of patients with diabetes can cause sugar molecules to attach to proteins in a process called glycation. This happens to the protein, haemoglobin, which helps doctors assess how well controlled the diabetes is. We are interested in C-reactive protein (CRP). CRP is a protein which the body produces to help respond to infection and inflammation. Glycation of CRP is a neglected area of research and we aim to learn more about the role and consequences of glycation of CRP. The study will also aim to look in blood samples for the existence of a specific version of CRP known as monomeric CRP within the blood of patients who have already been identified as having high levels of CRP during routine blood tests. There is evidence suggest that this form of CRP may cause increased inflammation and the current blood tests for CRP do not include this version of the protein in their measurements. We therefore want to analyse blood samples for the existence of these possible modified forms of this protein.

### Why have I been invited?

You have been invited to take part in this study because you have diabetes and your routine blood test has shown raised CRP levels. We aim to recruit 40 people. We would like to use the blood sample that is left over the following completion of your routine blood tests so the sample taken will be the sample normally taken for your routine blood tests and there will be no need to collect any extra blood from you.

Version No: 2.0  
Date: 24 April 2017



**Do I have to take part?**

You are free to decide whether you wish to take part or not. If you do decide to take part you will be asked to sign consent form and you should keep this information sheet for future reference. You are free to withdraw from this study at any time and without giving reasons, your sample and consent form will be destroyed immediately. Your medical care will not be affected in any way by this study.

**What will I have to do if I take part?**

You will not need to do anything other than sign the consent form. If you decide to take part you will:

- i. Agree to read through, sign and date our participant consent forms (2x copies)
- ii. Agree for us to use any leftover blood taken from your routine blood test procedure (that would normally be discarded), in order to test for other versions of CRP.

You will not be required to undergo any procedure that is not already part of your routine clinical care at the hospital. This will complete your role within the study.

**What are the benefits (if any) of taking part?**

We don't expect any direct benefit to you from taking part in this study. However, in the long run, the results obtained from the study will help us to determine whether different versions of CRP are present within the blood of people with diabetes and if any of these forms have any relationship with diabetes and its complications.

**What are the risks (if any) of taking part?**

Since you will not be required to undergo any procedure that is not already part of your routine clinical care at the hospital, we do not expect any or risks over and above that of normal clinical practice.

Your name will be included in the consent form, but all other study information and data will be kept anonymised, confidential, and secure. You will be never identified by name in any communications or publications arising from this research.

**How will information about me be used?**

If there is any blood remaining after your routine blood tests it will be stored and sent to structural biology research group at Keele University. It will then be used by the Research Team to test experimental new blood tests for detecting different versions of CRP. We will not use these samples to test for anything that does not directly relate to this project. Once the study is completed, blood samples will be transferred back to the hospital and disposed off using the normal procedures. Your sample will remain anonymised and you will never be identified in any communications or publications from any research. Information generated from this study will be included within the thesis write up by Mr Sameer Mahmood, PhD student at Keele University, and any associate research publications.

### Who will have access to information about me?

Individuals who have taken part will not be identifiable in any way. All samples will be anonymised by a research team member at the University Hospital of North Midlands (UHNM). Each sample will be given a specific sample code at UHNM and then transferred to Keele University for analysis. Research team members at Keele University will therefore be unable to identify any patient samples. The reported results will be stored on a password protected computer database or in locked filing cabinets which only the research members have access to.

### Who is funding and organising the research?

The project is funded by Keele University and under the supervision of Dr Annette Shrive. Her co-investigators are Mr. Sameer Mahmood, Dr Sarah Hart, Prof T Greenhough (Keele University) and Prof T. Fryer (University Hospitals of North Midlands) and the laboratory test will be conducted at Keele University.

### What if there is a problem?

If you have a concern about any aspect of this study, you may wish to speak to the researchers who will do their best to answer your questions (contact details on the next page).

If you remain unhappy about the research and/or wish to raise a complaint about any aspect of the way that you have been approached or treated during the course of the study please write to Dr Clark Crawford who is the University's contact for complaints regarding research at the following address:-

Dr Clark Crawford  
Head of Research and Clinical Governance  
Directorate of Engagement and Partnerships  
Innovation Centre 2 Building  
Keele University  
ST5 5NH  
E-mail: [research.governance@keele.ac.uk](mailto:research.governance@keele.ac.uk)  
Tel: 01782 733371

### Contact for further information

Patients can contact the Patient Advise and Liaison service (PALS) by telephone on 01782 676450/676455 or via e-mail to [patient.advice@uhns.nhs.uk](mailto:patient.advice@uhns.nhs.uk) for an independent contact point where they can seek general advice about taking part in research. In addition, for further information about this study, patient can also contact the researcher listed below.

Mr. Sameer M Mahmood  
Postgraduate researcher  
Lecturer  
Huxley building, Keele university  
Keele  
ST5 5BG  
[s.m.mahmood@keele.ac.uk](mailto:s.m.mahmood@keele.ac.uk)  
Tel: 01782 734441

Dr. Annette Shrive  
Structural Biology Researcher and Senior  
Huxley building, Keele university  
Keele  
ST5 5BG  
[a.k.shrive@keele.ac.uk](mailto:a.k.shrive@keele.ac.uk)  
Tel: 01782 733419

Thank you for your time reading this information sheet and for considering taking part in this study, please keep this information sheet for your records.

## Appendix 3

University Hospitals of North Midlands  
The Royal Stoke University Hospital  
Newcastle Rd,  
Stoke-on-Trent,  
Staffordshire  
ST4 6QG  
01782 715444

### PATIENT CONSENT FORM

**Title of Project: Modifications to C-reactive protein in Diabetes Mellitus**

**Name and contact details of Principal Investigator:** Prof Anthony A Fryer,  
Professor of Clinical Biochemistry,  
Department of Clinical Biochemistry

University Hospital of North Midlands  
Tel: 01782 674245  
[Anthony.Fryer@uhns.nhs.uk](mailto:Anthony.Fryer@uhns.nhs.uk)

**Study Centre: Royal Stoke University Hospital / Keele University**

**Study Number: CRP1**

**Patient ID Number for this study: A001**

**Please initial box if you agree with the statement**

1. I confirm that I have read and understood the information sheet dated 24/Jun/2016, version (1), for the above study and have had the opportunity to ask questions
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected.
3. I understand that sections of my medical notes may be looked at by responsible individuals from regulatory authorities, the NHS trust or from the sponsor (Keele University), where it is relevant to my taking part in this research. I give permission for these individual to have access to my data.
4. I understand that once the study has been completed, my serum sample will be transferred back to the Royal Stoke University Hospital and disposed off using the normal procedures.
5. I agree to take part in the above study.



\_\_\_\_\_  
Name of participant

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name of person taking consent

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

*[Original to be filed in site file, copy to be given to patient; copy to be filed in patient's medical records]*

## Appendix 4

Concentrations of CRP in the initial samples provided by UHNM and in the purified samples of mCRP and pCRP, the volumes of purified mCRP & pCRP samples are also provided.

Sample ID	Type of diabetes	Concentration of CRP provided by UHNM (mg/L)	Volume of Purified pCRP Samples (mL)	Concentration of pCRP in purified samples (mg/L)	Volume of purified mCRP Samples (mL)	Concentration of mCRP in purified samples (mg/L)
CRP01	Type II	100.4	0.5	26.96	0.5	0.065
CRP02	Type II	113.3	0.6	18.92	0.8	0.074
CRP03	Type II	152.2	0.7	57.75	0.9	0.100
CRP04	Type II	121.55	0.48	12.30	0.8	0.180
CRP05	Type II	141.1	0.78	4.09	0.6	0.082
CRP06	Type II	157.2	0.48	26.60	0.45	0.044
CRP07	Type II	131.3	0.48	38.07	0.5	0.056
CRP08	Type II	224.8	0.4	42.38	0.45	0.043
CRP09	Type II	164.2	0.47	39.06	0.45	0.045
CRP10	Type II	100.2	0.48	16.61	0.5	0.047
CRP11	Type II	169.5	1	23.77	0.9	0.076
CRP12	Type II	104.3	1	11.24	0.6	0.066
CRP13	Type II	101.4	0.8	12.95	0.95	0.091
CRP14	Type I	120.1	0.9	41.15	1.1	0.101
CRP15	Type II	227.5	0.9	47.02	0.9	0.110
CRP16	Type II	134.2	0.9	15.52	0.85	0.088
CRP17	Type II	152	0.7	46.76	0.8	0.219
CRP18	Type II	118	0.85	28.75	0.75	0.126
CRP19	Type II	162.1	1	37.87	0.9	0.105
CRP20	Type I	118	0.6	31.46	0.6	0.132
CRP21	Type II	248.2	0.7	48.43	0.8	0.174
CRP22	Type II	185.1	0.7	34.60	0.7	0.078
CRP23	Type I	143	0.9	32.34	0.7	0.067
CRP24	Type I	127.2	0.9	26.07	0.45	0.050