PROTEOMICS DISCOVERY OF NOVEL BIOMARKERS FOR DIAGNOSIS AND STRATIFICATION OF PATIENT AT RISK FOR ATHEROSCLEROTIC CARDIOVASCULAR DISEASES

CHEOW SOK HWEE ESTHER

CHEOW SOK HWEE ESTHER

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University

in partial fulfilment of the requirement for the degree of

Doctor of Philosophy

2016

Acknowledgements

Foremost, I would like to express my deepest gratitude to my supervisor, A/Prof Sze Siu Kwan, Newman, for the opportunity to learn under his invaluable guidance, for his continuous patience and support, for sharing his technical expertise and immense knowledge over these past years. His devotion, enthusiasm and attitude towards research are admirable, and it has truly been a fruitful and wonderful research experience working in his lab.

I would like to extend my heartfelt thanks my thesis advisors, A/Prof Lin Chun Ling, Valerie and Prof Salil Kumar Bose for their precious time, insightful comments and constructive advice to the development of the project. I would like to render my appreciation to the past and present laboratory mates: Dr. Guo Tiannan, Dr. Zhu Yi, Dr. Lixin, Dr. Sun Yang , Dr. Zhu Hong Bin, Dr. Ren Yan, Dr. Hao Piliang, Dr. Park Jung Eun, Dr. Juan Wei, Dr. Aida Serra Maqueda, Dr. Bamaprasad Dutta, Dr. Sunil Shankar Advav, Sim Kae Hwan, Meng Wei, Hong Hui Qi, Qian Jingru, Zhang Qi, Belinda Chen and Gallart Palau Xavier Ramon for their helpful discussion and assistance, for their listening ear, for their constant encouragement and most importantly for their friendship. I would also like to thank Prof Sara Sandin and Dr. Andrew Wong from Nanyang Technological University for his help with the Cryo-EM images and Dr. Justin Lim from Absciex for his assistance in the analysis of the iTRAQ- labeled samples by LC-MS/MS.

I would also like to give thanks to our clinician collaborator from National University Hospital of Singapore, Dr. Vitaly Sorokin, for his support and generosity in the provision of precious clinical plasma and aortic samples and accompanying clinical information. I am grateful to Nanyang Technological University, for granting me the opportunity and financial support to make this PhD endeavor possible. Last but not least, I am deeply thankful for the unconditional love and support from my family. I would like to give special thanks to my husband, Lin Hansheng, for his love, company, understanding and encouragement throughout these years

Table of Contents

Table of Contents

Table of Contents

Table of Contents

List of Figures

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Figures

Chapter 5

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Abbreviations

%CV	Percentage Coefficient Variation
2DE	Two-Dimensional Gel Electrophoresis
ABB	Ammonium Bicarbonate
ACN	Acetonitrile
AQUA	Absolute Quantitation Of Protein
AUC	Area Under The Curve
BCA	Bicinchoninic Acid
BP	Biological Processes
CABG	Coronary Artery Bypass Grafting
CAD	Coronary Artery Disease
CD	Cluster Of Differentiation
CE	Collision Energy
C-HDL	Carbamylated-High Density Lipoprotein
CID	Collision-Induced Dissociation
CKD	Chronic Kidney Disease
C-LDL	Carbamylated-Low Density Lipoprotein
CVD	Cardiovascular Disease
DAVID	Database For Annotation, Visualization And Integrated Discovery
DTT	Dithiothreitol
ECG	Electrocardiogram
ECM	Extracellular Matrix
ELISA	Enzyme-Linked Immunosorbent Assays
EMPAI	Exponentially Modified Protein Abundance Index
ERLIC	Electrostatic Repulsion-Hydrophilic Interaction Chromatography
ESI	Electrospray Ionized
EV	Extracellular Vesicles
FA	Formic Acid
FDA	Food And Drug Administration
FDR	False Discovery Rate
GO	Gene Ontology
HAC	Acetic Acid
HPLC	High Performance Liquid Chromatography
HS	Heparan Sulfate
HUPO	The Human Proteome Organization
I/R	Ischemia-Reperfusion
IAA	Iodoacetamide
IFNɣ	Interferon-Gamma
INOS	Inducible Nitric Oxide Synthase
ITRAQ	Isobaric Tags For Relative And Absolute Quantitation

Abbreviations

LC	Liquid Chromatography
LTQ	Linear Quadrupole Ion Trap
MCP1	Chemoattractant Protein-1
MCSF	Macrophage Colony-Stimulating Factor
MF	Molecular Function
MGF	Mascot Generic File
MI	Myocardial Infarction
MMP	Matrix Metalloproteinases
MMTS	Methyl Methanethiosulfonate
MPO	Myeloperoxidase
MRM	Multiple-Reaction Monitoring
MS	Mass Spectrometry
MSMS	Tandem Mass Spectrometry
NCE	Normalized Collision Energy
NH4ACE	Ammonium Acetate
NMI	Stable Angina
PBS	Phosphate-Buffered Saline
PPM	Parts Per Million
PPP	Plasma Proteome Project
PSM	Peptide Spectrum Match
PTM	Post-Translational Modification
PUC	Prolong Ultracentrifugation
RA	Rheumatoid Arthritis
ROS	Reactive Oxygen Species
RT	Retention Time
SD	Standard Deviation
SILAC	Stable Isotope Labels With Amino Acids In Cell Culture
SMC	Smooth Muscle Cells
SPEG	Solid Phase Extraction Of N-Linked Glycoproteins
SRM	Selected Reaction Monitoring
TCEP	Tris 2-Carboxyethyl Phosphine Hydrochloride
TEAB	Triethylammonium Bicarbonate
TFA	Trifluoroacetic Acid
TNFα	Tumour Necrosis Factor Alpha
TRIS	2-Amino-2-Hydroxymethyl-Propane-1, 3-Diol
UC	Ultracentrifugation
VCAM1	Vascular Adhesion Molecule-1
WAX	Weak Anion Exchange
XIC	Extracted Ion Chromatogram

Abstract

Atherosclerotic coronary artery disease (CAD)-related deaths will continue to rise along with global population ageing. Although established conventional risk factors are in place for the prediction of cardiovascular (CV) events, however, the absence does not necessarily imply zero CAD risk. Effective screening and diagnostic methods that stratify patient at high risk for CAD, before onset of acute events, remain unsatisfactory, necessitating the search for novel measurable protein biomarkers. Thankfully, biomarker discovery has taken a leap forward with the development of mass spectrometry (MS)-based proteomics approaches, which afforded the interrogation of the plasma proteome. The incomplete understanding in the mechanisms in plaque destabilization and rupture, creates a knowledge gap in the pathophysiological underlying atherosclerosis and myocardial infraction (MI). In addition, currently there is no plasma-based screening assay that effectively differentiates atherosclerosis from MI. The development of innovative techniques that improves the detectable range of plasma proteome, is essential in advancing the field of plasma-based biomarker discovery research. My PhD research work has been divided into four parts (I - IV), the first three parts showcased the clinical utility of the plasma proteome for CAD proteomics-based biomarker discovery, and in the final part, and the successful development of a novel enrichment strategy for plasma glycoproteins and extracellular vesicles (EV) was demonstrated. Herein, the careful selection and systematic integration of several proteomics strategies have proven useful in the discovery of novel plasma biomarkers of CAD, derived from (i) undepleted plasma proteome, (ii) EV plasma proteome and (iii) carbamylated plasma proteome. Functional linkage analyses have prompted the generation of new hypotheses into the pathophysiological states that specifies either for atherosclerosis or myocardial injury. The novel plasma sub-proteome enrichment technique described here have afforded the detection of low-abundant proteins with concentration ranging from pg/mL – ng/mL in crude plasma. In essence, this thesis demonstrates the applicability and reliability of the different proteomics platforms in the discovery and development of useful biomarkers that are predictive of CAD.

Chapter 1

General Introduction

Chapter 1

1.1 Cardiovascular diseases (CVD) epidemiology

Cardiovascular disease (CVD) is the leading cause of death worldwide, with higher occurrence (approximately 80%) in low income and middle income countries (1). It has been estimated by the World Health Organization (WHO) (http://www.who.int/cardiovascular_diseases/en/, accessed on July 24, 2015) that CVD accounts for almost 31% (17.5 million) of global deaths in 2012, of which, nearly 7.4 million deaths were due to coronary artery disease (CAD), and about 6.7 million deaths were due to stroke. It has been projected that the prevalence of CVD-related mortality will continue to climb to 23.4 million by 2030 (2). The underlying basis of CAD and stroke is due to atherosclerosis, a disease process directed mainly by lifestyle factors (3). The growing incidence of CVD in the developing countries are largely attributed by urbanization, lifestyle and behavioral changes that increases cardiovascular (CV) risk factors, the lack of public prevention efforts and access to advanced healthcare treatments (3, 4).

Extensive epidemiological research have allowed the recognition of major conventional risk factors that drives the development of CVD, including smoking (5, 6), diabetes (7, 8), obesity (9, 10), dyslipidemia (11, 12) and hypertension (13, 14). These leading modifiable risk factors were established based on the following criteria: high incidence in diverse populations; significant independent impact on the risk of CVD ; and the control and treatment of these factors reduces CVD risk (15). Other non-modifiable determinants of CVD includes ageing, gender and genetic deposition (16). Although the presence of multiple conventional risk factors have been useful in the prediction and prevention of CV events, it has been shown that at least 14% of initial cardiac event occur amongst asymptomatic individuals without CVD risk factors (17). Hence, the absence of established CVD risk factors does not necessarily indicate a zero risk of developing CVD. This, as a result, have prompted the search for novel non-conventional CVD risk factors to improve risk assessment of CVD, and guide therapeutic/preventive strategies for the management of CVD.

1.2 Pathophysiology of atherosclerotic CAD, the predominant manifestation of CVD

Atherosclerosis, a systemic arterial intimal thickening process that begins in early adolescence and progresses slowly and silently for decades, is the chief cause of CAD (18). Atherosclerosis is a lipid-driven inflammatory disease, characterized by the chronic development of asymptomatic multifocal plaques (stable and unstable) from the progressive accumulation of cholesterol, lipid-filled macrophages, inflammatory cells, smooth muscle cells (SMCs) and extracellular matrix (ECM) (19, 20). Whilst stable atherosclerotic plaques are clinically silent in already narrowed artery lumen, rupture of the unstable atherosclerotic plaques and ensuing thrombosis causes serious clinical manifestations such as acute myocardial infarction (MI), angina, stroke and sudden death (21-23). It has now been well-documented that inflammation plays a central role in mediating the pathogenesis of atherosclerosis, from the initial recruitment of leukocytes to destabilization, and eventual rupture of the atherosclerotic plaques (Figure 1.1) (24, 25).

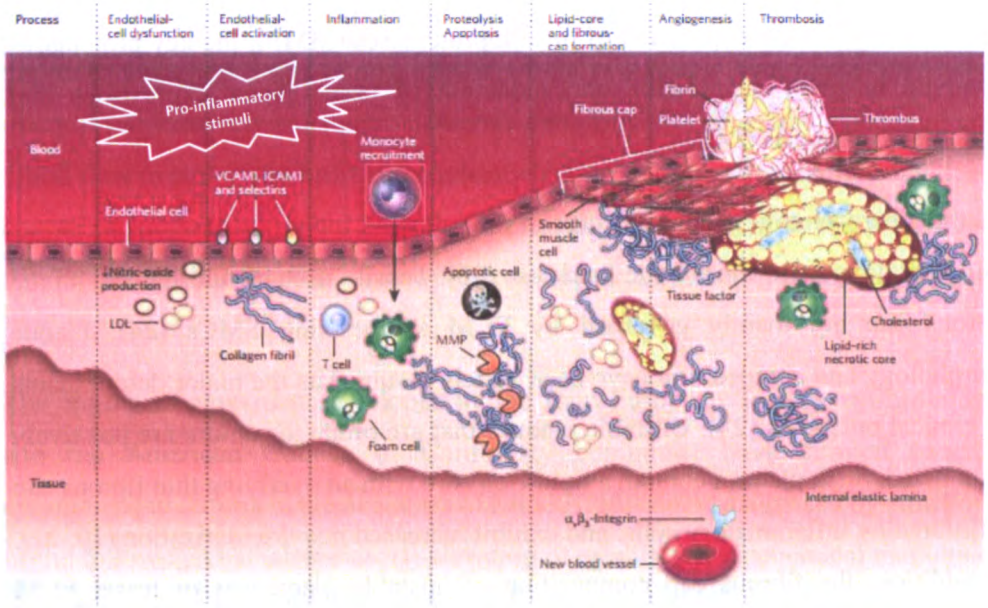

Figure 1.1. The simplified stage progression of an atherosclerotic lesion. Under chronic exposure to pro-inflammatory stimuli, developing from a normal blood vessel (far left) to a vessel with an atherosclerotic plaque and superimposed thrombus (far right). Figure adapted from Ref. (26).

Circulating monocytes do not normally adhere to the endothelium, however, constant exposure to proinflammatory stimuli such as the abovementioned modifiable risk factors, triggers endothelial expression of adhesion molecules such as P-selectin and vascular adhesion molecule-1 (VCAM1), initiating the tethering and adherence of monocytes to the dysfunctional endothelium (27). In response to modified low density lipoproteins (LDL), monocyte chemoattractant protein-1 (MCP1) is released from the vascular walls, directing the transmigration of adhered monocytes into the intima (27). Within the inflamed intima, monocytes mature into macrophages, expressing receptors for modified LDL under the influence of macrophage colony-stimulating factor (MCSF) and MCP1 (24). Macrophages then transform into foam cells after the engulfment of modified LDL, further amplifying the inflammatory state through the secretion of numerous growth factors (e.g. tumour necrosis factor alpha (TNFα) and cytokines (e.g. interleukin-6) (27). Activation of T-cells, the adaptive arm of immune response, leads to the expression of more pro-inflammatory cytokines including interferon-γ (IFNγ) and cluster of differentiation (CD) 40 ligand (CD40L, CD154), which contributes to atherogenesis (27). Ligation of CD40 triggers the expression of metalloproteinases (MMP), the key proteinases involved in ECM degradation and induction of tissue factor expressions that mediates thrombosis (24).

As the atherosclerotic plaque matures in the inflamed intima, it develops a fibrous cap overlaying a pro-thrombotic lipid core (25). Tensile strength and stability of the fibrous cap are mainly governed by ECM synthesizing SMCs (24). Plaque morphology and composition have been widely assumed as the major determinants of clinical outcomes (28). Unstable plaques that are prone to rupture are positively remodeled. Such plaques contain a lipid-rich core with an overlying thin fibrous cap infiltrated by inflammatory cells, and exhibit increased neovascularization (19, 22). In addition, the fibrous cap composition of unstable plaques is of lesser ECM (collagen, elastin and proteoglycans) and with fewer SMCs than in stable plaques (23). The thickness of the fibrous cap is determined by the delicate balance of synthesis and breakdown of ECM, which is dynamically mediated by inflammatory signals (24). Predominance of the pro-inflammatory state result in an over expression

of MMPs and inflammation-induced apoptosis of SMCs. Collectively, these leads to cap thinning, a prelude to plaque rupture and vulnerability (29). Rupture of the unstable atherosclerotic plaque releases prothrombotic lipid pool into the circulation, which induces the formation of thrombus, and the eventual blockage of blood flow through the coronary artery elicits an acute clinical event.

1.3 Clinical diagnosis and treatment of CADs

CADs are a group of clinical conditions comprising of stable angina, unstable angina, MI and sudden cardiac death (30). Anginas are defined as chest pains arising from myocardial ischemia, due to decreased coronary blood flow. Baseline electrocardiogram (ECG), is one of the initial test performed on a patient presented with chest pain (31). The clinical diagnosis of angina depends largely on the detailed assessment of CV risk factor profile, patient's history and presenting clinical symptoms (31). Stable anginas are presentation of predictable chest pains during exertional activities or stress, and are resolved with rest or sublingual administration of nitroglycerin (31). Unstable anginas are regarded as indicators of an impending MI and are presented with non ST-segment elevation on ECG tracking, unpredictable chest pains at rest, angina within two weeks after acute MI or new onset of progressively worsening patterns of angina (32). MI occurs when coronary blood flow to part of the heart is completely blocked, causing myocardial cell death due to prolonged ischemia (33). Sudden cardiac death refers to the unexpected natural death from a cardiac cause within one hour from the onset of acute symptoms (34).

The consensus agreement developed by the Joint European Society of Cardiology and the American College of Cardiology committee have defined specific recommended criteria in the diagnosis of MI (33). MI can be diagnosed by either one of the following characteristics: (i) elevation of established biochemical markers of myocardial necrosis – cardiac troponin (T or I), preferably or creatine kinase-MB fraction accompanied with at least one of the following (a) ischemic symptoms; (b) development of pathologic Q waves on the ECG; (c) ECG changes indicative of ischemia (ST segment elevation or depression); or (d) coronary artery intervention,

(ii) pathologic findings of MI via conventional imaging techniques such as cross-sectional echocardiography, radionuclide angiography and myocardial single-photon emission computed tomographic perfusion imaging (33). Of note, cardiac troponins are delayed circulatory markers of irreversible myocardial injury, hence the diagnosis of MI would require constant monitoring by serial blood sampling over a period of 6 to 12 hours after onset of acute presentation, increasing risk of complications due to delayed diagnosis (35). Furthermore, the elevation of cardiac troponins does not necessarily indicates an ischemic mechanism, as other causes (e.g. myocarditis, heart failure) of myocardial damage may also result in elevated levels of cardiac troponins.

The initial typical therapeutic treatment for unstable angina and MI includes the administration of oxygen and several medications such as aspirin (platelet inhibitior), nitroglycerin (vasodilator) and morphine (analgesic) (36). Other adjunctive treatment includes platelet inhibitors (clopidogrel, glycoprotein IIb/IIIa receptor inhibitors) (37, 38), β-adrenergic receptor blockers (carvedilol) (39, 40), thrombin inhibitors (heparins) (40, 41), anti-hypertensive agents (angiotensin-converting-enzyme inhibitors) (42, 43) and cholesterol lowering agent (statins) (44, 45). The mainstay coronary revascularization interventions for complex CAD are percutaneous coronary intervention, and coronary artery bypass grafting (CABG) surgery, with CABG faring better in patient survival and freedom from repeat revascularization (46, 47). Although these surgical procedures served as temporary effective revascularization strategies for CADs, it does not represent a remedy, as long-term prognosis are compromised by restenosis, systemic occurrence and thrombosis of unstable atherosclerotic plaques that have not been treated surgically. Therefore, patients are still at high risk for developing secondary CV events.

Early detection of CAD and timely therapeutic intervention that prevent, delay or attenuate plaque destabilization are of utmost importance in preventing a life-threatening event. However, achievements for such screening methods are far from ideal for the prevention and timely intervention of CADs. Current diagnosis and treatment of CADs are focused on the late symptomatic stages, often after

irreversible functional impairment of the myocardium, hence limiting treatment options following poor prognosis. Thereby, easily accessible, novel measurable protein signatures that reflects the varying severity of atherosclerotic manifestation for risk stratification of CAD before the onset of acute event are urgently needed.

1.4 Emerging circulatory protein biomarkers of CAD

The growing prevalence of CAD and the lack of precise biomarkers indicative of atherosclerotic progression, have fueled the active investigation into biomarker discovery in this setting. Biomarker, as defined by the National Institutes of Health Working Group, is "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathologic processes, or pharmacologic responses to a therapeutic intervention" (48). Instances of biomarkers includes recordings of blood pressure, biochemical test on biological fluids or tissues or image assessment of echocardiogram. The valuable applications of biomarkers includes their usage as diagnostic tools for the identification of disease condition; staging tools for the categorization of disease severity; prognostic tools for the evaluation and monitoring of clinical response to therapeutic intervention ; predictive tools for the determination of risk in disease development; and as surrogate endpoints for establishment of therapeutic efficacy in clinical trials (48).

The desirable properties of a clinically useful biomarker regardless of their intended use, should possess high sensitivity and specificity of an investigation, providing accurate and reproducible results of either confirmation of disease (high positive predictive value) or exclusion (negative predictive value) (49). Importantly, a biomarker should facilitate the change in disease management and improvement in patient outcome. Biomarker discovery relies on two integrated strategies: the biased approach, which capitalized on the understanding of the disease pathophysiology process to guide the design of assay for new candidate marker, and the unbiased approach involves the direct profiling of disease molecules for the characterization of candidate marker signature of the disease stage (49). The biomarker discovery pipeline consist of six essential process components: candidate discovery, qualification, verification, assay optimization, validation and commercialization (50).

Chapter 1

Biomarker discovery is moving away from the idealized single disease-specific biomarker to a panel of multi-markers that better transcend the molecular, population and epidemiological heterogeneity, so as to achieve a higher level of specificity and sensitivity required for clinical use (51). Rapid advancement in genomics and proteomics technologies have led to the proliferation in the number of candidate biomarkers of CADs, and have provided new insights into the pathophysiology of atherosclerosis. The better understanding in the pathogenesis of atherosclerosis have enabled the recognition of key processes (inflammation, oxidative stress and hemostatic disorder) involved, highlighting their capacity in prediction for atherothrombotic events. As atherosclerosis is a multifactorial disease, multi-markers approach are often used to improve CAD prediction and risk stratification, accounting for as many underlying causes as possible (52). It is widely recognized that vascular inflammation plays an essential role in all stages of atherosclerosis pathogenesis. As such, active research are currently underway to interrogate the clinical application of a variety of plasma inflammatory mediators. Epidemiological studies have showed increased CV risk in association with elevated levels of several emerging inflammatory markers, including interleukin-6 (53); interleukin-18 (54); lipoprotein-associated phospholipase A2 (55); myeloperoxidase (56); TNFα (57); intercellular adhesion molecule-1 (58); VCAM1 (59); P-selectin (60); E-selectin (59); CD40L(61); c-reactive protein (CRP) (62); fibrinogen (63) and serum amyloid A (64), in outpatient screening of apparent healthy individuals. Among the emerging biomarkers, CRP is the best characterize inflammatory markers and promising marker for CV risk.

Oxidative stress due excessive generation of reactive oxygen species (ROS) is believed to incite endothelial dysfunction, an initial step of atherosclerosis (65). Oxidative stress induced oxidation on LDL, favouring its uptake by macrophages and consequently promoting the formation of foam cells (66). Accordingly, epidemiological studies have found reduced CV risk in association with increased consumption of antioxidants (67, 68). Given the obvious involvement of oxidized LDL in progressing atherosclerosis, its usefulness as a predictor for the development of atherosclerosis have been proven in clinically healthy individuals (69, 70),

however, inconsistencies have been found in numerous studies in CAD patients (71-74). Other studies on proposed markers of oxidative stress conducted on CAD patients includes F-2 isoprostane (75) and glutathione peroxidase-1 (76). At present, there is no well-accepted biomarker of oxidative stress due to the little evidence that suggest adequacy in predicting CV events. Hemostasis, an element response to damaged vessel walls, dictates the development and outcome of atherosclerosis (77). The imbalance between the coagulation and fibrinolysis may result in a hypercoagulable state, that accelerates thrombosis, increasing risk for acute events (52). Evidence of the predictive potential of key emerging hemostatic variables with their significant association to CAD risk have been evaluated on fibrin D-dimer (48), von Willebrand factor (78), factor VIIIc (79), tissue plasminogen activator antigen (80) and plasminogen activator inhibitor-1 (81). However, the independent predictive value of hemostatic markers risk factors remains uncertain and would require further investigations.

The aforementioned players related to cardinal processes implicated in the development of atherosclerosis, have made them promising predictive markers for assessing CV risk. Despite the support from epidemiological studies, the lack of standardized commercial assays and low accuracy of available assays have hampered the clinical application of these potential biomarkers, with the exception of CRP (27). In addition, more investigations are needed before these emerging markers can find utility in primary, or secondary prevention settings of CADs. With growing occurrence in CAD-related mortality, better non-invasive biomarkers are urgently needed for CAD risk stratification. It is believed that biomarker discoveries will continue to flourish in coming years, with the rapid advances of proteomics and genomics technological platforms, the likelihood of success is optimistic.

1.5 Mass spectrometry (MS) -based proteomics in biomarker studies

Proteomics is the application of technologies to study of the function, structure and interaction of the proteome or all proteins expressed by the genome, including proteins isoforms arising from alternative splicing of transcripts and modified proteins from post-translational processing (82). The proteome represents actual

functional biomolecules within an organism, it is dynamic as it constantly changes according to biochemical and environmental interactions (83). Methods that affords the monitoring of protein expression patterns in disease-affected states, offers hopes of discovering of disease-specific biomarkers that hold clinical diagnostic and prognostic values. Enzyme-linked immunosorbent assays (ELISA) or other immunological assays, conventionally used for biomarker verification and validation exploits the binding specificity of antibodies for parallel screening of specific protein expressions (84). However, the development of reliable and sensitive antibody-based assays can be very costly and time-consuming, as it is limited by diverse physiochemical properties of proteins, cross-reactivity issues, low-throughput and subjected availability of antibodies with high specificity and high affinity for the targeted proteins (85, 86). Therefore, immunoassays once developed, are trustworthy tools for large-scale screening of validated biomarkers but less applicative for discovery and validation of new biomarkers (87).

From the mid-1970s, proteomics was pursued using the then newly developed two-dimensional gel electrophoresis (2DE) technique (88, 89). However, it was difficult to dictate the identification of separated proteins due to the absence of an efficient analytical method at that time. The emergence of mass spectrometry (MS) in the 1990s along with development of database search tools and publicly availability of entire human sequence, have made protein characterization and protein expression examination possible, marking the beginning of proteomics era (90). In practice, the technical challenges of proteomics are much greater than in genomics. Unlike in genomics, proteomics must deal with issues such as, limited sample as proteins are non-scalable, sample to sample variability, vast dynamic range in biological system poses challenges for detection of low-abundance proteins, proteome complexities attributed by a plethora of post-translational modifications (PTM), and disease and drug perturbations (91).

The progressive advances in MS technologies and improvements in sample handling techniques in the last decade, have made possible the unbiased identification and quantification of a rich diversity of systemic proteins in a high-throughput manner,

overcoming the shortcomings described in the preceding paragraph (92, 93). In shotgun or bottom-up proteomics, the typical workflow consist of the extraction of proteins from desired sample type (.e.g. cells, tissues, biofluids, extravellular vesicles (EV)), followed by proteolytically digestion into peptide mixture. Then the peptide mixture can either be fractionated by chromatography to reduce sample complexity or enriched for specific PTMs by affinity methods, depending on the aim of the study. Processed peptide fractions are analyzed by high pressure reverse-phase liquid chromatography (LC)-coupled to MS (LC-MS), where electrospray ionized (ESI) peptides are first acquired for preliminary mass analysis in MS1, and subsequently dissociated into smaller fragments usually by collision-induced dissociation (CID) in tandem MS (MS/MS or MS2) for fragment mass spectrum analysis. Following which, database search is carried out using acquired MS1 and MS2 spectra for pairing with genomic sequences for peptide and protein identification. The quality of the data can be assessed through reverse or decoy database search strategies where an estimated false discovery rate (FDR) on the entire dataset is obtained and used for biological interpretation.

The absolute or relative quantitative measurement of peptide and protein abundance from LC-MS/MS is performed either by stable isotopic labeling or by label-free approach. The exact differences in the amount or the concentration of proteins (.e.g. ng/ml) are determined in absolute quantification. Whereas in relative quantification, changes in protein expression are expressed as fold change of protein abundance relative to a reference protein. The profiling and identification of aberrantly expressed proteins in healthy and disease-affected subjects, favours the discovery of novel disease-specific biomarkers. Labelling approaches have been considered to be more accurate in protein abundances measurement, as it bypasses the problems of ion-suppressive effects arising from co-eluting peptides (94). In labeling strategy, the mass difference between the labeled (heavy) and unlabeled (light) peptide is recognized by MS, and quantitative ratio is determined by comparing the relative signal intensities between heavy and light peptide pair (95). Mass differentiated stable isotope tags can be introduced into peptides or proteins via four methods,

namely (i) metabolically, (ii) chemically, (iii) enzymatically or (iv) as a spiked synthetic analog (95, 96).

Stable isotope labels with amino acids in cell culture (SILAC), a relatively simple process, is a widely adopted in vivo metabolic labelling method. In SILAC, isotopically labeled essential amino acid (.e.g $^{13}C_6$-lysine, $^{13}C_6$-arginine, $^{15}N_4$-arginine) are supplied in amino acid deficient growth media, and heavy amino acids are incorporated into the newly synthesized polypeptides by the growing cells, which leads to a defined mass shift (.e.g. 4 Da mass shift for $^{15}N_4$-arginine) when compared with the light (unlabeled) counterparts (97). The application of SILAC have been proven useful in the study of protein dynamics (98, 99). The availability of several heavy labels allows for up to five cellular states for relative comparison within a single experiment (100). However, arginine when present at high levels can be converted into proline during cell division, hence complicating the quantification and result in potential inaccuracies (101). SILAC is only limited to metabolically active cells and thus cannot be applied to tissues and biological fluids. In addition, not all cell lines are well suited for growth in dialyzed serum, in the absence low molecular mass growth factors (102).

Isobaric tags for relative and absolute quantitation (iTRAQ) is a popular choice of in vitro chemical tagging approach used for MS quantification (103). The iTRAQ labels contains three element, an amine-specific reactive group for peptide labeling, a balance group that ensures that the overall label mass is indistinguishable, and a specific charged reporter group (isobaric tag) of variable mass which enables them to be differentiated upon fragmentation in MS/MS (104, 105). The attractiveness of iTRAQ-based quantitation is attributed by its compatibility to any type of sample, multiplex quantification of up four (4-plex) or eight (8-plex) samples in a single analysis and the availability of fairly well developed quantitation software (106). The relative quantitation, for example in 8-plex iTRAQ analyses, is determined by comparing the peak areas of the eight reporter ions derived at 113.1, 114.1, 115.1, 116.1, 117.1, 118.1, 119.1, and 121.1 m/z (103). However, the accuracy of the estimation of differential expression arising from variability in labeling efficiencies

(107, 108), especially in complex samples, and the compression of observed expression ratio due to inherent dynamic range limitation (109), are of major concerns of iTRAQ. Nevertheless, the successful application of this technique, have been demonstrated in a wide range of biomarker discovery studies in human pathologies, including cancer (110-112), neurodegenerative diseases (113-115) and cardiovascular diseases (116-118).

Proteolytic ^{18}O labeling is a simple and convenient labeling technique where stable isotope is introduced enzymatically (.e.g. trypsin, lys-C, glu-C) at the carboxy termini of tryptic peptides during digestion (119). Briefly, proteins digested in $H_2{}^{18}O$ is mixed proteins digested in $H_2{}^{16}O$ and analyzed by MS. Relative quantification of protein expression is achieved by comparing the signal intensities of the ^{18}O peak and ^{16}O peak (120). The occurrence of sample loss during preparation is minimal as heavy and light peptides are combined post-digestion (101). In addition, the incorporated ^{18}O atoms are resistant to back exchange, and thus are stable through LC-MS (121). This technique has been applied mostly in mechanistic and peptide sequencing studies (122, 123). It has been reported that signal interference from ^{13}C effect may lead to the misinterpretation of weak ^{18}O signals as unlabeled carboxy terminal peptides, causing variability in relative quantitation (120).

Absolute quantitation of protein (AQUA) is performed by spiking known quantities of chemically synthesized isotope-labeled peptides as internal standards into unlabeled protein digest (124). In AQUA, the absolute quantity of only one or a few targeted proteins of interest is determined, by the comparison of MS signal intensities of heavy and light targeted peptides. This approach is broadly utilize in studies aimed at large-scale verification and validation of candidate biomarkers (125, 126), or quantification of specific PTM peptide of interest (127, 128). AQUA are usually used in combination with targeted MS-based quantification, where pre-defined peptide and peptide fragments of specific proteins are isolated on selected- or multiple-reaction monitoring (SRM or MRM)-MS/MS, where precise monitoring and absolute quantitation of peptide (s) are performed with minimal interference from background ions (124). MRM-MS-based targeted approach, a multiplexed

quantitative targeted proteomic platform, affords the simultaneous monitoring of multiple peptide transitions, offering the sensitivity, specificity, reproducibility and throughput required for pre-clinical verification and validation of large number of candidate biomarkers proposed in the discovery phase, while not being limited by antibody availability (85). MRM-MS has been recognized as a rapid and cost-effective protein biomarker measurement technology (129). Because the internal standard is introduced post-digestion, any prior variability that occurs during the sample preparative steps are not corrected (124) . Unlike the aforementioned labeling techniques, where relative quantitation is performed on all proteins present in a mixture, this strategy is limited to a small number of targeted proteins.

Currently in use label-free quantification methods are hassle-free and inexpensive alternative to labeling strategies. Label-free quantification of protein abundance are estimated by two approaches, (i) spectral counting, which is based on counting the number of MS/MS spectra assigned to each protein and (ii) measurement of peptide chromatographic peak intensity where measured chromatographic peaks of peptide precursor ion of each protein are compared (130). . The peptide spectral counting approach is based on the empirical observation that the more abundant the protein, the more MS/MS spectra are acquired for peptides of the given protein (95). Differential relative quantification is achieved by the comparison in the number of acquired spectra between different set of experiments. Spectral counting of standard proteins has showed linear dynamic range of over two orders of magnitude and is highly correlated to the relative protein concentration in yeast extracts (131), therefore this method have often been used to quantify proteins over small dynamic range. As the physical property of a peptide is not measured in spectral counting, and with the assumption that the linearity response is the same for every protein, this method has remained controversial (95). In the method by peak intensity measurement, the chromatographic peaks areas of the extracted ion chromatograms (XIC), for a given peptide eluted in LC-MS/MS run, are aligned or integrated over the chromatographic time scale for the measurement of area under the curve (AUC) (132). Relative quantitative values can be obtained by comparing the measured integrated intensity of aligned peaks from identified peptides between different

experiments. It has been demonstrated that the measured integrated AUC correlates linearly with a wide range of protein abundance in complex biological samples (133-136). Label-free MS-based approaches are ideal for biomarker discovery studies as it permits the comparison of large number of samples for statistical significance (137). Tedious data processing steps that includes data filtering and normalization are essential in label-free protein quantification, for extraction of high quality and meaningful information that has been corrected for inherent bias and global variability across datasets.

With the availability of numerous labeling options to date, each with their respective weighted merits and limitations, factors that influence the choice of labeling technique depends on the complexity of sample, the number of samples, multiplexity of the experiment, availability of analytical instrument and software, quantitative objective and the cost of the experiment (101, 138). There is certainly ample space for continuous improvement in proteomics, where progressive advancement in MS technologies and dedicated development efforts with emphasis on the improvement in sensitivity, dynamic range, robustness and throughput, to meet analytical needs for the interrogation of complex biological system. MS-based proteomics research, in near future, will find clinical utility in the discovery and development of useful biomarkers that are predictive of human pathologies.

1.6 Proteomic biomarker discovery in plasma

The dynamic components present in plasma reflects ongoing physiological or pathological states, and the ease of access makes it the preferred choice for biomarker applications (139). Apart from its resident proteins, plasma contains other sub-constituents that are secreted or shed from cells, tissues and organs throughout the body (140). Unfortunately, the predominance (~99%) of high abundance classical plasma proteins, which include albumin, immunoglobulins, transferrins and complement factors, creates up to ten orders of magnitude difference between the high- and low-abundance (~1%) plasma proteins as shown in Figure 1.2 (141). This wide dynamic range poses great analytical challenges as the dynamic range of MS detection is typically up to 3 orders of magnitude (87). There is a wide belief that

clinically relevant biomarkers are expected at concentrations in the ng/ml range or lower (142).

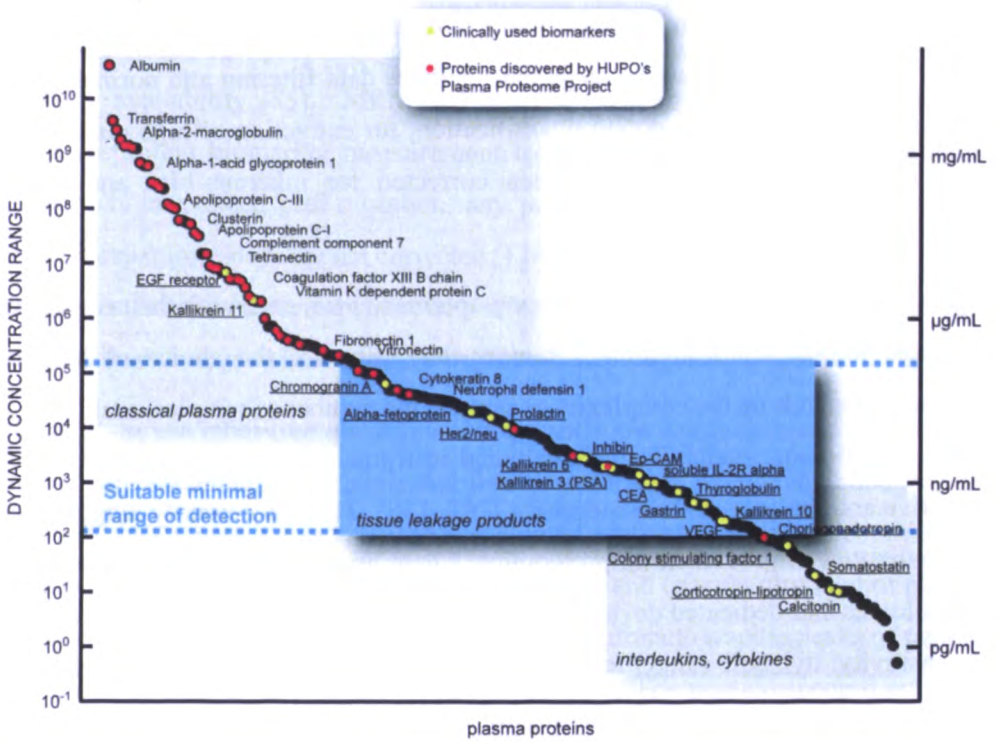

Figure 1.2. Plasma protein concentration range. Range of plasma proteins concentration and the three main protein categories as described by Anderson et al. (143). Red dots indicate proteins identified by the HUPO PPP initiative (144), and yellow dots represent currently utilized biomarkers (142). Suitable minimal range of detection for biomarker targeting in plasma is shown with dotted lines. Figure adapted from Surinova et al.(145).

Strategies aimed at simplifying the complexity of plasma includes immunoaffinity depletion of high abundant proteins (146), multi-dimensional chromatographic fractionation (147-149), and the selective enrichment for sub-proteomes (.e.g. post-translational modified proteins, microvesicles) (150-152) were developed to improve detection of low-abundance candidate proteins. The Human Proteome Organization (HUPO) initiated the pilot phase of Plasma Proteome Project (PPP) in 2002, to evaluate the feasibility of comprehensive plasma proteome analysis using diverse depletion, extensive fractionation and multidimensional MS technology platforms across the eighteen collaborating laboratories (153, 154). In 2005, HUPO PPP cataloged a publicly disclosed core dataset of 889 high-confidence serum and

plasma proteins, derived from a 3020 low-confidence dataset (144, 155, 156). Although the application of these aforementioned strategies prior to MS analyses have indeed improved the depth of plasma proteome coverage over the past decade, the desired range of detection remains far from ideal. The trade-off for in-depth analyses is constrained by limited sample throughput following extensive fractionation, or by limited proteome coverage following the specific isolation of a single subset of proteins (139).

Sub-proteomic post-translational modified proteins and extracellular vesicle (EV) proteins, both present at low abundance, are commonly pursued as alternate sources of biomarkers within the plasma, through the reduction of proteome complexity. PTMs are covalent chemical alterations that affects the protein's properties through proteolytic cleavage or addition of modifying groups to one or more specific amino acid residue(s) (157). PTMs of proteins regulate numerous biological processes, including molecular interactions, localization, activity and function, localization, stability and turnover (158). To date, over 300 different types of PTM of proteins have been reported to occur physiologically, and many of them are implicated in the pathogenesis of human diseases, including cancer, diabetes, CVD and neurological disorders (158, 159). However, the study of PTM as a sources of biomarkers in human diseases are still at its infancy.

Protein glycosylation represents the most common PTM, with 50% of human proteins estimated to be glycosylated (160, 161). Although plasma glycoproteins are present in low abundance, they play important roles in both biological and pathological processes, including cell-cell interactions, signal transductions, protein stability, developmental process, immune defenses, and changes in protein glycosylation profiles have been correlated to diseases such as cancer and neurodegenerative disorders (160-165). Indeed, most clinically used cancer protein biomarkers are glycosylated, including human epidermal growth factor 2 in breast cancer, prostate specific antigen in prostate cancer, carcinoembryonic antigen in colorectal cancer, cancer antigen 125 in ovarian cancer and α-fetoprotein in hepatocellular carcinoma (164, 166, 167). Protein glycosylation occurs commonly

in two forms, (i) N-linked, where the carbohydrate moiety is attached to an asparagine (N) residue and (ii) O-linked, where the carbohydrate moiety is attached to the hydroxyl side chain of serine (S) or threonine (T) residue. Of interest, N-linked glycosylation, typically secreted or shed from plasma membrane (168), is more prevalent and a major PTM occurring in plasma (169, 170). The study of plasma N-glycoproteome have been made possible with the application of selective enrichment techniques prior to MS analyses. Selective enrichment of glycoproteins and glycopeptides are achieved commonly by lectin affinity based method (171, 172) and hydrazide chemistry based-solid phase extraction method (173). Although there is an increasingly recognition of protein glycosylation in the modulation of CAD, CV glycoproteomics studies are lacking as compared other diseases such as cancer. Lately, three non-structural cardiac ECM glycoproteins: thrombospondin, tenascin-C, and periostin have been shown to mediate cardiac remodeling after MI (174-176). It has recently also been reported that elevated levels of glycosylated peptides found in myocardium subjected to ischemia-reperfusion injury, were associated with cardiac tissue remodeling following MI (177). In addition, recent report on the increased serum levels of glycosylated apolipoprotein A1 (APOA1) in patients with acute MI were demonstrated (178). While glycoproteomics is still an emerging field, the development of effective strategies for the selective enrichment of N-glycoproteome for MS-based analyses would provide the platform for characterization of aberrantly expressed glycosylated protein in healthy and disease-affected subjects, facilitating the discovery of novel glycosylated biomarkers for CAD.

Protein carbamylation, a dissociation product of urea is an age-related PTM mediated by cyanate, and one of the earliest PTM elucidated (179). Increased levels of protein carbamylation have been associated with decline renal function in chronic kidney disease (CKD), accelerated by raised plasma urea (180, 181). It was recently reported that increased levels of carbamylation, can be attributed by myeloperoxidase (MPO)-catalyzed oxidation of thiocyanate (derived from dietary intake and smoking), under chronic atherosclerotic inflammatory conditions (182). The application of LC-MS/MS in the quantitative assessment on a wide spectrum of

carbamylated proteins from tissue and plasma was first demonstrated by Wang et al (182). Following which, led to the development of several alternative precise measurement of carbamylation modification by LC-MS/MS (183-186). The increased levels of carbamylated-LDL (c-LDL) and carbamylated-HDL (c-HDL) in uremic patients have been shown and thought to promote the development of uremic atherosclerosis (187-191). Despite considerable carbamylation studies performed to date, knowledge of circulatory protein carbamylation in CADs, under normal renal function, remain sparse and poorly-defined. The under-explored role of carbamylated plasma proteins solely in CADs, is in part due to the extremely low production of carbamylated-derived products under basal conditions, where plasma urea concentrations are normal (182). As measured values of carbamylated proteins, in the absence of uremia are reflective of chronic inflammation burden, there is considerable interest for their usage as diagnostics biomarkers.

Membrane-bound vesicles (.e.g. exosomes, microvesicles) termed as EVs are secreted by almost all cell types into the circulation under both normal and pathological conditions (192). EVs are known to play key roles in a myriad of physiological processes including regulation of cell signaling (193, 194), cell-cell communication (195, 196), immune surveillance (197, 198), stem cell maintenance (199) and tissue repair (200). Furthermore, the association of EVs with several human pathologies including neurodegenerative diseases (199, 201), vascular diseases (202, 203), tumorigenesis (204, 205) and infectious diseases (206, 207), have presented them as highly attractive targets as disease biomarkers. Due to their low abundance in plasma, prior to LC-MS analyses, EVs are selectively isolated either by differential centrifugation followed by ultracentrifugation (UC) (208), density gradient centrifugation (209), filtration techniques (210), or by immune-affinity capture beads (211). Recent studies have demonstrated the therapeutic potential of transplanted mesenchymal stem cells (MSC)-derived exosomes in reducing myocardial injury and improving cardiac function following MI in animal models (212, 213). EVs are attractive targets with numerous potential applications in clinical utility, at present proteomics studies on circulatory EVs in MI-related

research are limited and their involvement in myocardial ischemic injury has yet to be determined.

Over the past decade, significant efforts and funding have been directed towards the discovery of novel disease biomarkers using proteomics technologies, yet the successful translation of novel proteomics discovery into regulatory approved clinical diagnostics is disheartening. To date, OVA1, is the first and only In vitro Diagnostic Multivariate Index Assay (IVDMIA) consisting of five proteomic biomarkers cleared by the US FDA (Food and Drug Administration) in 2009, for preoperative assessment of ovarian cancer risk in women diagnosed with ovarian tumor (214, 215). Considerable discussion in the research community has addressed several confounding issues that may limits the path towards clinical diagnostics, including the lack of an optimized assay with suitable analytical performance suitable for clinical deployment (216), the lack of a coordinated biomarker discovery pipeline guideline (50), the lack of sufficient evidence in preliminary validation studies to support investment for assay development and large-scale validation trials (214) and the lack of a careful and clearly defined clinical utility that balances broad applicability and practical feasibility for completing the clinical trials (214). The path from biomarker discovery to clinical diagnostic may be long and uncertain, but with patience and the rationalized strategies, time will unveil the future of protein diagnostics. Nevertheless, proteomic profiling of plasma remains the focus of recent attempts to identify new components, in hopes of discovering of disease-specific biomarkers that hold clinical diagnostic and prognostic values.

1.7 Thesis overview and research aims

Atherosclerotic CAD-related deaths will continue to increase in tandem with global population ageing, imposing major socioeconomic burden. Although established conventional risk factor are in place for the prediction and prevention of CV events, however, the absence does not necessarily imply zero CVD risk. Thus, suggesting the involvement of non-conventional risk factors in the multifactorial pathogenesis of atherosclerosis. Early detection and timely intervention are essential for the management and prevention of undesirable outcomes of CAD. Unfortunately current

diagnostic methods have been focused on the late symptomatic stages of CAD, limiting treatment options following poor prognosis. Hence, novel measurable protein signatures of CAD that effectively stratify high risk patients, before the onset of acute event, are thereby urgently needed.

As mentioned above, MS-based proteomics approaches have proven to be useful in CAD biomarker studies, not merely just in the identification of candidate markers, but also in providing new insights into the pathophysiology of atherosclerosis. Furthermore, the development of numerous effective pre-fractionation strategies and sub-proteome enrichment techniques, have enabled the interrogation of clinically attractive low abundance plasma proteins by MS. Herein, the studies presented in my thesis are fueled by three confounding reasons, (i) the lack of plasma-based screening assays that effectively discriminates underlying atherosclerosis from impending myocardial injury, (ii) the incomplete understanding of the pathological process underlying both atherosclerosis-associated stable angina and the sudden onset of MI, and (iii) the pursuit in developing innovative strategies that affords interrogation of low abundance plasma protein. In order to achieve better grasp in the abovementioned issues, my PhD research work has been divided into four parts (I - IV) (Figure 1.3), the first three parts showcased the clinical relevance and maximal utility of the plasma proteome and sub-proteomes for CAD proteomics-based biomarker discovery, and the successful development of a novel enrichment strategy for glycoproteins and EV from plasma was demonstrated in the final part.

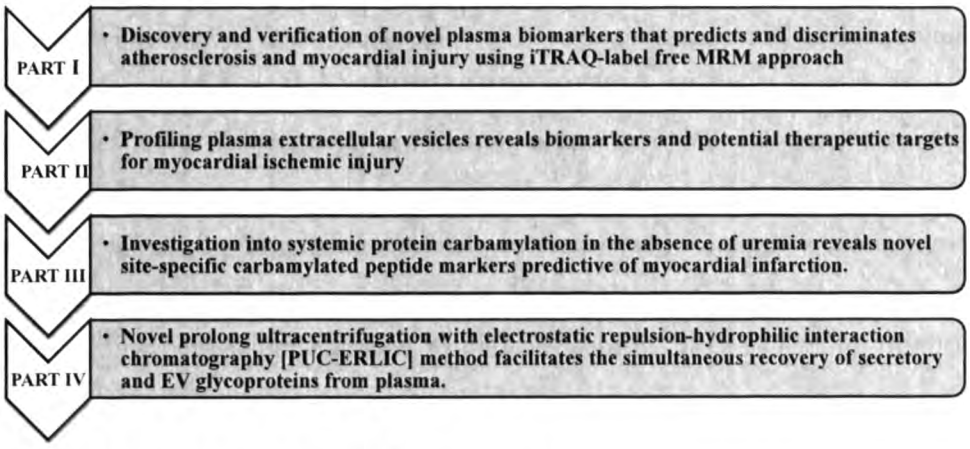

Figure 1.3. Schematic overview of PhD research work

Chapter 1

In part I, atherosclerotic-specific protein biomarkers, and myocardial injury-specific protein biomarkers were identified in undepleted pooled plasma using iTRAQ-labeled based quantitative approach. Subsequent verification of candidate markers on individual plasma was achieved with label-free MRM based approach. Functional pathological linkages analyses suggested the involvement of atherosclerotic-specific markers in endothelial dysfunction, while myocardial injury-specific markers were involved in plaque destabilization and protective mechanisms. A multi-marker diagnostic panel consisting of eight novel biomarkers that specifically predicts and discriminates the multifactorial pathophysiology of atherosclerosis and myocardial injury was proposed. In part II, myocardial injury EV protein biomarkers were identified in plasma EV proteome using label-free peak intensity quantification method. Comparative protein proteomics profiling of perturbed EV proteins suggest its participation in both myocardial damage and myocardial healing processes following MI. Validation of two complement protein on individual patient plasmas were carried out using Luminex assay.

In part III, candidate inflammatory carbamylated peptide biomarkers of myocardial injury were identified in plasma carbamylated proteome using label-free peak intensity quantification method. As a consequence of carbamylation, we proposed that the loss of immunomodulatory function in carbamylated proteins may drive the development of inflammation associated atherosclerotic conditions and impair tissue healing. In part IV, the successful development of a novel enrichment method that couples prolong ultracentrifugation (PUC) with electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) [PUC-ERLIC] for the simultaneous recovery of glycoproteins and EV from crude plasma was described. Without the use of immunodepeletion strategy, low-abundant glycoproteins with concentration ranging from pg/mL – ng/mL were detected. The identification of cardiac-specific remodeling glycoprotein in plasma of stable angina patients demonstrated the feasibility of PUC-ERLIC in identification of potential biomarkers with specific relevance to human pathologies. In order to meet the research objective in the respective studies, each part was explored with the best suited tools carefully selected from the proteomics toolbox, demonstrating the applicability and reliability of the

different proteomics platforms in the discovery and development of useful biomarkers that are predictive of CAD.

Chapter 2

Discovery and Verification of Novel Plasma Biomarkers that Predicts and Discriminates Atherosclerosis and Myocardial Injury Using iTRAQ-Label Free MRM Approach

Chapter 2

2.1 Abstract

Effective screening and diagnostic methods that stratify patients at risk for atherosclerosis or/and myocardial injury remain yet to be optimized, necessitating the need for novel non-invasive biomarkers. In this study, we performed comparative quantitative proteomics on undepleted plasmas from patients with stable angina (NMI), patients with acute myocardial infraction (MI) and control subjects without angina (Ctrl). A total of 53 high confident (FDR < 1%, $p < 0.05$) candidates with at least two ratio-fold change in expression levels were shortlisted from an initial list of 371 proteins. Among which, 27 proteins were identified as atherosclerotic-specific markers, and 26 proteins were selected as myocardial injury-specific markers in the discovery phase. In the verification phase, label-free LC-MRM-MS was employed to screen 53 candidate biomarkers in pooled plasma, peptides with poor performance were removed. Finally, 49 peptides representing eight atherosclerotic-specific markers, 15 myocardial injury-specific markers were quantified in 49 individual plasma samples. Verification analyses identified 12 out of 23 protein candidates with significant ($p < 0.05$) differential expression agreement with discovery proteomics experiments. We found two (F10, MST1) atherosclerotic- specific markers and six (ORM2, SERPINA3, CPN2, LUM, ACTG1, NAGLU) myocardial injury- specific markers be novel candidates. We speculate that increased levels of F10, SERPINA3, CPN2 and decreased levels of LUM drives the progression of atherosclerosis and/or myocardial injury. While the increase expression of MST1, ORM2, ACTG1 and decreased abundance of NAGLU promotes atheroprotective and/or tissue healing activities. Functional pathological linkages analyses have suggest the involvement of F10 and MST1 in endothelial dysfunction, SERPINA3, CPN2, LUM in plaque destabilization and ORM2, ACTG1, NAGLU in protective mechanisms, when taken together are signatures that discriminates atherosclerotic conditions from myocardial injury. We have proposed a multi-marker diagnostic panel consisting of eight novel biomarkers that specifically predicts and discriminates the multifactorial pathophysiology of atherosclerosis (F10, MST1) and myocardial injury (ORM2, SERPINA3,

CPN2, LUM, ACTG1, NAGLU). We showed that discovery and verification of novel candidate biomarkers undepleted plasma is achievable with iTRAQ-label-free MRM based approach.

2.2 Introduction

Coronary artery atherosclerosis, the underlying cause for coronary artery disease (CAD), is the prevalent cause of cardiovascular (CV) deaths worldwide (18).The rupture of atherosclerotic plaque and ensuing thrombosis in CAD can cause serious clinical consequences such as acute myocardial infarction (MI), angina, stroke and sudden death (21-23). While established traditional risk factors are in place to assess atherosclerotic burden, at least 14% of initial cardiac event occur amongst asymptomatic individuals without CVD risk factors (17). Current imaging modalities and serological indicators used in the diagnosis and monitoring of CADs are focused on the late symptomatic stages, often after irreversible myocardial injury, thus limiting treatment options (27, 35, 217, 218). Early detection and timely therapeutic interventions that prevent, delay or attenuate plaque destabilization are of utmost importance in preventing a vascular event. Thereby, there is an urgent demand for new measurable protein signatures that reflects varying degree of atherosclerotic manifestation for risk stratification of CAD.

Atherosclerosis, a systemic disease characterized by chronic development of asymptomatic multifocal plaques from progressive accumulation of cholesterol, lipid-filled macrophages, inflammatory cells, smooth muscle cells (SMC) and extracellular matrix (ECM) (19). It has been clear that inflammation plays a central role in mediating the atherosclerosis pathogenesis, from initial recruitment of leukocytes to destabilization and eventual rupture of the atherosclerotic plaque (24, 25). As such, active biomarker research is underway to evaluate clinical utility of various plasma inflammatory mediators (e.g. interleukin-6 , c-reactive protein (CRP)) involved in the pathogenesis of atherosclerosis (25). However, achievements in this field remains inadequate for the prevention and timely intervention of this disease (115). Given that the molecular mechanism for initial inflammatory response leading to the eventual rupture of atherosclerotic plaque occurs in a stochastic fashion, it is

extremely challenging to identify the exact trigger factors that destabilizes the plaque (219).

Biomarker discovery is moving away from the idealized single disease-specific biomarker to a panel of multi-markers that better transcend the molecular, population and epidemiological heterogeneity, to achieve a higher level of specificity and sensitivity required for clinical use (51). Technological advances in mass spectrometry (MS) have enabled the unbiased identification of systemic proteomes in a high-throughput manner, hence are powerful platforms for biomarker discovery studies (92, 93). Shotgun-based quantitative proteomics approaches have made study of aberrant protein expression in control and diseased samples possible (16), generating a new pool of disease-relevant biomarkers in a single experiment. The increase application of proteomics in CV research over the years have indeed provided invaluable insights into pathophysiological events in CAD and heart failure (83). Nonetheless, the understanding in mechanism(s) and trigger(s) for plaque destabilization remains unclear, creating a significant knowledge gap in in the pathophysiological underlying atherosclerosis-associated angina and the sudden onset of acute clinical events.

The lengthy and laborious verification and validation stages of candidate biomarkers creates an undeniable bottleneck in biomarkers pipeline for clinical utility. Enzyme-linked immunosorbent assays (ELISA) are conventionally used for biomarker verification and validation, however the development of ELISA assays for multiple candidate protein targets can be very costly and time consuming (85). Multiple reaction monitoring (MRM)- MS, a multiplexed quantitative targeted proteomic platform, has been recognized as a rapid and cost-effective protein biomarker measurement technology for verification and validation (129). The capability of this emerging MRM-MS-based targeted approach that affords simultaneous monitoring of multiple peptide transitions, offers the reproducibility and throughput required for pre-clinical verification of large number of candidate biomarkers, while not being limited by antibody availability (85).

Chapter 2

In this study, we aim to identify candidate biomarkers that discriminates underlying atherosclerotic conditions from myocardial injury. In the discovery phase, isobaric tags for relative and absolute quantification (iTRAQ)-based quantitative proteomics was performed on undepleted pooled plasmas obtained from patients with stable angina (NMI), patients with acute myocardial infraction (MI) and control subjects without angina. Based on stringent selection criteria, 27 proteins were identified as atherosclerotic-specific markers, and 26 proteins identified as myocardial injury-specific markers from the discovery experiments. In the verification phase, label-free MRM quantitation approach was first employed to survey the initial list of 53 candidate biomarkers in pooled plasma. An optimized MRM-list consisting of 59 peptides representing eight atherosclerotic-specific markers, 15 myocardial injury-specific markers and three house-keeping proteins were used for label-free relative quantification in 49 individual plasma samples. MRM verification analysis identified 12 candidates with significant ($p < 0.05$) altered expression, among which, two (F10, MST1) atherosclerotic-specific markers and six (ORM2, SERPINA3, CPN2, LUM, ACTG1, NAGLU) myocardial injury-specific markers were novel. Functional pathological linkage have proposed atherosclerotic-specific protein markers involvement in endothelial dysfunction and myocardial-specific protein markers involvement in plaque destabilization and protective mechanisms. These novel candidates warrant further assessment in larger cohort verification and clinical validation studies.

2.3 Materials and Methods

2.3.1 Chemicals

All water and acetonitrile (ACN) used in this experiment were of high performance liquid chromatography (HPLC) grade (Thermo Scientific, Waltham, MA). All chemicals were purchased from Sigma-Aldrich (St Louis, MO) unless stated otherwise.

2.3.2 Human plasma samples

Thirty-five participants with different severity of CAD were recruited for this study, coronary angiogram was used to confirm CAD and cardiac troponin (cTn) test and electrocardiogram (ECG) was used to confirm MI. Fifteen patients diagnosed with unstable angina pectoris and have had recent (< 5 days) MI were grouped as MI, and 20 patients diagnosed with stable angina pectoris, without MI were grouped as NMI. General demographics and clinical characteristics of patients were documented in Table 2.1. Plasma was collected from MI patients (n=15) and NMI patients (n=20) prior to coronary artery bypass graft surgery (CABG). Control (Ctrl) plasma was obtained from 14 individuals without angina. All plasma aliquots were stored at -80°C until proteomics sample processing. Written informed consent of participation was obtained from all participants. The study was approved by the National Healthcare Group Domain Specific Review Board (NHG DSRB).

Table 2.1. Patient demographics and clinical characteristics

	Myocardial infarction (MI) (n = 15)	Stable angina (NMI) (n = 20)
Age [Mean (± SD)]	63 (± 11)	62 (± 7)
Female [n (%)]	5 (33.33)	5 (25)
Male [n (%)]	10 (66.67)	15 (75)
Chinese [n (%)]	8 (53.33)	10 (50)
Malay [n (%)]	4 (26.67)	5 (25)
Indian [n (%)]	2 (13.33)	5 (25)
Others [n (%)]	1 (6.67)	0 (0)
Hypertension [n (%)]	14 (93.33)	19 (95)
Renal impairment [n (%)]	2 (13.33)	0 (0)
Hyperlipidemia [n (%)]	15 (100)	20 (100)
Diabetes Mellitus [n (%)]	9 (30)	14 (70)
Smoking [n (%)]	5 (33.33)	10 (50)

n, number of individuals; SD, standard deviation

2.3.3 Undepleted plasma protein precipitation

In order to minimize biological variation, individual plasma from each group were equivalently pooled to final volume of 200 µL. Plasma proteins were allowed to precipitate in 80% acetone for 4 h at -20°C. Plasma proteins were pelleted by centrifugation (16, 000 x g, 10 min) and were washed twice in 80% acetone (-20°C) with centrifugation. Recovered protein pellets were quantified using bicinchoninic

acid (BCA) assay according to the manufacturer's recommendation. For each sample group, 200 µg of proteins were extracted for downstream proteomics processing.

2.3.4 In-solution tryptic digestion, peptide labeling and peptide fractionation

Extracted proteins were solubilized in lysis buffer containing 8 M urea, 50 mM triethylammonium bicarbonate (TEAB) (pH 8.0), supplemented with protease inhibitors (1:50, v/v) and phosphatase inhibitors (1:10, v/v) (Roche Diagnostics, Mannheim, DE). 200 µg of proteins approximately from each group were reduced with 5 mM tris 2-carboxyethyl phosphine hydrochloride (TCEP) for 3 h at 30°C, followed by alkylation in the dark with 10 mM methyl methanethiosulfonate (MMTS) for 1 h at room temperature. Urea concentration was diluted to less than 1 M with 50 mM TEAB, and digestion was carried out overnight at 37°C with sequencing grade modified trypsin (Promega, Madison, WI) at a 1:100 ratio (w/w, trypsin:protein). Trypsin digestion was stopped by acidification with 10% trifluoroacetic acid (TFA) to 0.1% TFA (v/v). The tryptic peptides were desalted using a Sep-Pak C18 cartridge (Waters, Milford, MA) and eluted peptides were dried in a vacuum concentrator. According to manufacturer's protocol (Applied Biosystems, Foster City, CA), the dried peptides were reconstituted in 50 mM TEAB and labeled with respective 8-plex iTRAQ isobaric tags as follows: 113_{Ctrl}, 114_{MI} and 115_{NMI}. Labeled plasma peptides were combined and dried using a vacuum concentrator. Dried iTRAQ-labeled peptides were desalted using Sep-Pak C18 cartridge (Waters, Milford, MA) and dried in a vacuum concentrator. iTRAQ-labeled dried peptides were reconstituted in 200 µL of mobile phase A (85% ACN, 0.1% acetic acid [HAc]) and fractionated using a PolyWAX LP anion-exchange column (4.6 × 200 mm, 5 µm, 300 Å, PolyLC, Columbia, MD) on a Shimadzu Prominence UFLC system (Kyoto, JP). UV spectra of the peptides were collected at 280 nm. Mobile phase A and Mobile phase B (30% ACN, 0.2% formic acid [FA]) were used with a 60 min gradient elution of 0 - 36% B over 30 min and 36 − 100% B over the next 20 min, followed by 10 min at 100% B, at a flow rate of 1 mL/min. Thirty separate fractions were collected and vacuum dried. Dried peptides in each fraction were then

reconstituted in 3% ACN, 0.1% FA for subsequent liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

2.3.5 iTRAQ-labeled LC-MS/MS

Dried labeled peptides were dissolved in 40 µL of solvent A (2% ACN, 0.1% FA) and 1 µL of sample per fraction were loaded into a trap column (0.5 mm x 200 µm) at a flow rate of 3 µL/min for 10 min and resolved on an analytical column (15 cm x 75 µm) with a linear gradients of solvent B (98% ACN, 0.1% FA) from 5%-12% in 2 min; 12% to 30% in 57 min and 30-90% in 2 min at a flow rate of 300 nL/min on the Nanoflex cHiPLC system. The nanoLC column was rinsed with 90% solvent B for 7 min and equilibrating with 95% solvent A for 13 min. For information dependent acquisition (IDA) on AB SCIEX TripleTOF® 5600 system, 250-ms survey scan (TOF-MS) and 100-ms automated MS/MS product ion scan for the top-20 ions with the highest intensity was performed with a cycle time of 2.3s. The MS/MS triggering criteria for parent ions were as follows: precursor intensity (> 125 cps), 2- 5 charge states with dynamic exclusion time of 8s and collision energy (CE) set as rolling CE script based on m/z and charged state of the precursors. Triplicate LC-MS/MS runs per fractions were performed. The peak areas of the iTRAQ reporter ions reflect the relative abundance of the corresponding proteins in the samples.

2.3.6 iTRAQ-labeled data analyses

All spectra generated from IDA acquisitions in Triple-TOF® were searched against the UniProt database (version Aug 2011, 446597sequences, 188463640 residues) using ProteinPilot™ V4.1 (AB SCIEX, Framingham, MA) software, for protein and peptide identification and iTRAQ quantification respectively. User defined parameters were configured as follows: (i) Sample Type, iTRAQ 8-plex (Peptide Labeled); (ii) Cysteine alkylation, MMTS; (iii) Digestion, Trypsin; (iv) Instrument, TripleTOF 5600; (v) Species, Human; (vi) ID Focus, Biological modifications ; (vii) Search effort, Thorough; (viii) Specific Processing, Quantitate, Bias correction,

Background correction; (ix) Results quality, Detected protein threshold [Unused Protscore (Conf)] >: 0.05 (10.0%). Peptides for quantification were automatically selected by Pro Group algorithm in ProteinPilot™ software to calculate the reporter peak area, error factor (EF) and p-value. Data was normalized by bias correction and background correction to get rid of any variations imparted due to loading error and to eliminate any background ion signal due co-elution of non-target peptides with target peptides respectively. Searched results were exported to Microsoft Excel for further comparative analysis between the replicate runs. A two-fold change cut-off was set, as such, expression ratios of $\geq$ 2 and $\leq$ 0.5 indicates up-regulated and down-regulated proteins respectively. The false discovery rate (FDR) report was generated by ProteinPilot™ for each search and the numbers of proteins reported for this study was based on global FDR < 1%. Differentially expressed proteins were extracted for further analyses.

2.3.7 Development of MRM-assay

In the initial phase of MRM assay, three to five unambiguous tryptic peptides for each candidate protein were selected. For each tryptic peptide precursor ion, three to five most intense fragment ions (transitions) were chosen for initial MRM assay. The selection of transitions and MRM conditions were aided by SRM Atlas (http://www.srmatlas.org/), a publicly accessible database containing previously acquired SRM/MRM transitions used in targeted proteomics (220). For candidate proteins that were undocumented in the SRM Atlas database, in silico digestion and fragmentations were performed using Thermo Scientific™ Pinpoint™ V1.3 software to facilitate the selection of potential transitions. MRM instrument acquisition parameters (e.g. normalized CE) predicted by Pinpoint™ V1.3 (Thermo Fisher Scientific, Cambridge, MA) software were imported directly to the instrument method setup program. Preliminary screening of unscheduled MRM-assay representing 53 candidate biomarkers were performed on unfractionated pooled plasma digest from MI, NMI and Ctrl groups. Positive peptide identification were based on the detection of at least three co-eluting transitions in all preliminary LC-MRM-MS runs, visualized using Pinpoint™ V1.3 software. Only the best peptides

with consistent transitions were selected for scheduled LC-MRM-MS on individual plasma samples.

2.3.8 LC-MRM-MS

The same set of individual patient plasma used in iTRAQ experiment for each group was individually digested, desalted and vacuum dried as described earlier, without peptide labelling and fractionation. The targeted peptides were assayed in triplicates in a TSQ Vantage triple quadrupole mass spectrometer coupled with EASY-nLC™ 1000 nanoflow UHPLC system (Thermo Scientific Inc., Bremen, Germany). The retention time (RT) of each peptide was determined from a full acquisition (unscheduled) MS/MS analysis. Subsequently, using ± 5 min error from the predicted RT, for each transition a 10 min isolation window was established for dynamic exclusion (scheduled) analyses. 1 µg of tryptic peptides were loaded each time onto Acclaim® PepMap100 trap column (75 µm x 2 cm; nanoViper C18, 3 µm, 100 Å) and resolved on an Acclaim® PepMap RSLC C18 column (75 µm x 15 cm; nanoViper C18, 2 µm, 100 Å) (Thermo Scientific, USA), at a flow rate of 300 nL/min. Mobile phase A (0.1% FA in HPLC water), and mobile phase B (0.1% FA in ACN) were used to establish the 60 min gradient; starting with 45 min of 3 – 30% B, 9 min of 30 –50% B, 1 min of 50 – 60% B and 2 min of 60% followed by re-equilibration at 3% B for 3 min. The TSQ Vantage was set to perform data acquisition in the positive ion mode. Electrospray potential of 1.5 kV and capillary temperature of 250°C were set for ionization. The selectivity for both Q1 and Q3 were set to 0.7 Da (full-width at half-maximum). The collision gas pressure of 1 mTorr of argon was set for Q2. Scan time of 10 ms and 50 ms was used for full scan MRM transitions and 10 min isolation window MRM transitions respectively.

2.3.9 MRM data analyses

All MRM raw data files generated from TSQ Vantage were processed using Pinpoint™ V1.3, the integrated peak areas of all targeted transitions were automatically detected and aligned for data visualization, relative quantification at

transition-level, peptide-level and protein-level, global normalization of peak intensities for the triplicate experiment were computed by the software. The extracted ion chromatogram of all transitions for each targeted peptide were examined manually for accurate peak integration verification. The peptide ion peak areas were used to determine the relative abundance quantification of corresponding proteins. Relative quantification values of each target protein were determined by the corresponding peptide ion peak areas of all peptide transitions.

2.3.10 Bioinformatics analyses and data annotation

Open access tools FunRich V2.1.2 (221) was used for gene ontology (GO) annotation and functional enrichment and interaction network analyses. Scatter dot plots were generated and statistically analyzed by GraphPad Prism V 6.0 (Graphpad Software, San Diego, CA). Proteolytic peptide sequences selected for MRM analyses were submitted to NCBI BLASTP (222, 223) search using default settings against a non-redundant protein sequences (nr). Parametric analyses and unpaired student's *t*-test were used to evaluate the differences between numeric variables for statistical comparisons and $p < 0.05$ was considered to be statistically significant. Data are presented as the mean ± standard deviation (SD) of the triplicate experiment.

2.4 Results and Discussion

2.4.1 iTRAQ-based quantitation workflow for plasma proteome analyses

Comparative quantitative proteomics analyses of plasma are useful for biomarker discovery. In this study, iTRAQ chemical labeling approach that affords both quantitation and multiplexing analyses in a single reagent (224) was employed in this study. In order to increase the likelihood of detecting stage-specific biomarkers that defines atherosclerotic condition from myocardial injury, preliminary patient recruitment and grouping into MI, NMI and Ctrl were carefully considered as part of the experimental design. As illustrated in Figure 2.1, for each sample group, equal concentration of pooled plasma proteins (200 µg) were proteolytically digested with trypsin. Tryptic peptides obtained from MI, NMI and Ctrl groups were labeled with

one of the isobaric tags and all labeled peptides were mixed together prior to first dimensional fractionation. Fractionated labeled peptides were analyzed in triplicates by LC-MS/MS and relative abundance of specific peptide among sample is determined by comparing the intensities of the iTRAQ reporter fragment ion in the 113-115 m/z region of the peptide product ion spectra. Differential quantitative proteomics analyses were performed using open source public tools and targeted proteins of interested were verified by LC-MRM-MS. Three technical replicates were performed in this study to increase the reliability of differences in quantitative changes in protein expression.

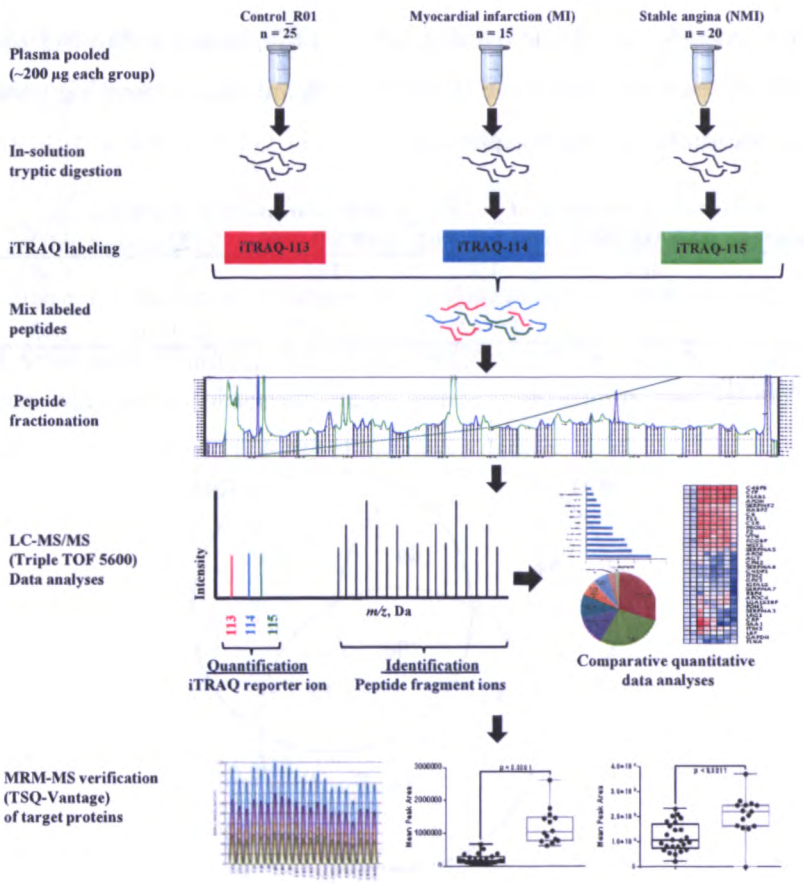

Figure 2.1. Work flow of iTRAQ and MRM-based approach. Schematic representation of the experimental design for discovery through verification quantitative proteomics on undepleted plasma using 8-plex iTRAQ and MRM-based approach.

2.4.2 Assessment of quantitative MS data quality

The run-to-run technical variation was determined in terms of percentage coefficient variation (%CV), the number of protein, peptide and spectral identified (FDR < 1%) in replicate 01 (R01), replicate 02 (R02) and replicate 03 (R03) were compared as summarized in Table 2.2. The overall %CV in repeated identification of protein, peptide and spectral observed across all triplicates were 8.36%, 2.00% and 1.43% correspondingly, which translate to minimal variation and good system repeatability. As shown in Figure 2.2, approximately 77% (356) of the total number of proteins (FDR < 1%) were quantified in common in at least two of the triplicate and nearly 62% (286) proteins were found overlapped in all triplicates LC-MS/MS runs, while 23% of all proteins were observed in single replicate, thus suggesting good protein complementation between the triplicates.

Table 2.2. Experimental variation of protein, peptide and spectral identification.

Identification	iTRAQ_R01	iTRAQ_R02	iTRAQ_R03	SD	%CV	Mean
Protein	464	450	395	36.47	8.36	436
Peptide	14848	15352	15394	303.84	2.00	15198
Spectral	159793	155425	156673	2249.86	1.43	157297

R01, Replicate 01; R02, Replicate 02; R03, Replicate 03; SD, Standard deviation; %CV, Percentage coefficient variation

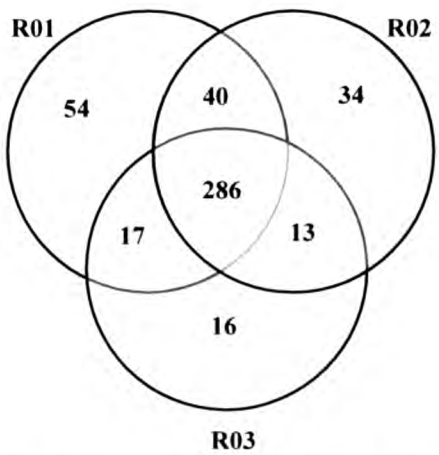

Figure 2.2. Venn diagram iTRAQ triplicates. Venn diagram of overlapping proteins identified (FDR < 1%) in three iTRAQ technical replicates. R01, replicate 1; R02, replicate 2; R03, replicate 3.

Next, to assess the reliability of iTRAQ quantitation in this study, the respective iTRAQ ratios of 114_{MI}:113_{Ctrl}, 115_{NMI}:113_{Ctrl} and 114_{MI}:115_{NMI} for each protein

(FDR < 1%) were calculated. The correlation of measured pairwise 114_{MI}:113_{Ctrl}, 115_{NMI}:113_{Ctrl} and 114_{MI}:115_{NMI} iTRAQ ratios for each protein between replicates were examined on scatter log_2 plots depicted in Figure 2.3. The generated Pearson correlation coefficients (r) and accompanied p values for each scatter log_2 plots of measured iTRAQ 114_{MI}: 113_{Ctrl} ratios between R01 and R02 (Figure 2.3A, $r = 0.7698$, $p < 0.0001$), between R01 and R03 (Figure 2.3B, $r = 0.6950$, $p < 0.0001$), and between R02 and R03 (Figure 2.3C, $r = 0.7616$, $p < 0.0001$), demonstrating consistency of measured iTRAQ ratios between the triplicates. Similarly, good reproducibility was also observed in scatter log_2 plots of measured iTRAQ 115_{NMI}:113_{Ctrl} ratios between R01 and R02 (Figure 2.3D, $r = 0.7919$, $p < 0.0001$), between R01 and R03 (Figure 2.3E, $r = 0.7314$, $p < 0.0001$), and between R02 and R03 (Figure 2.3F, $r = 0.6573$, $p < 0.0001$). Likewise, statistically significant correlation of measured iTRAQ 115_{NMI}:113_{Ctrl} ratios between R01 and R02 (Figure 2.3G, $r = 0.7904$, $p < 0.0001$), between R01 and R03 (Figure 2.3H, $r = 0.8463$, $p < 0.0001$), and between R02 and R03 (Figure 2.3I, $r = 0.8164$, $p < 0.0001$) were detected among the triplicates. The use of standardized plasma processing protocols and stable analytical system for plasma proteomics resulted in the minor system and quantitation variability observed here, hence confirming the reliability and confidence of our presented iTRAQ quantitative dataset.

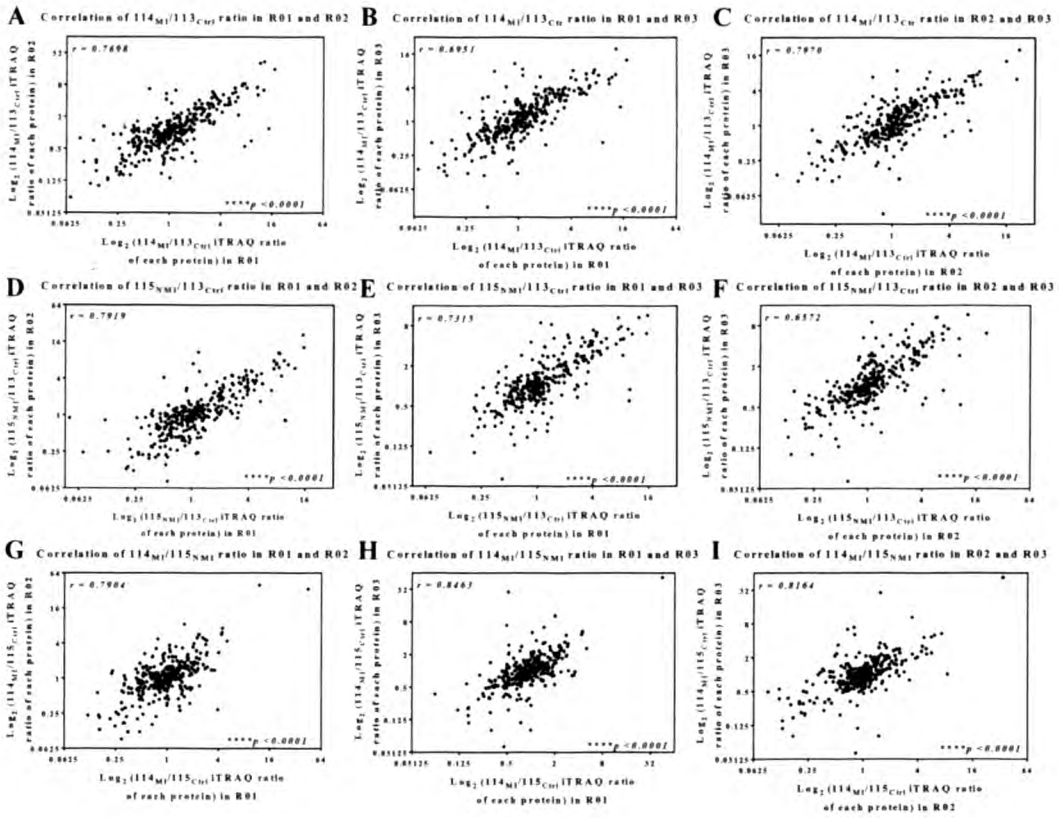

Figure 2.3. Correlation assessment between iTRAQ triplicate experiment. Correlation assessment of measured 114_{MI}:113_{Ctrl} , 115_{NMI}:113_{Ctrl} and 114_{MI}:115_{NMI} iTRAQ ratios for each protein (FDR < 1%) between triplicates. (A) Scatter $\log_2$ plots of measured 114_{MI}:113_{Ctrl} iTRAQ ratio of proteins identified in R01 and R02. (B) Scatter $\log_2$ plots of measured 114_{MI}:113_{Ctrl} iTRAQ ratio of proteins identified in R01 and R03. (C) Scatter $\log_2$ plots of 114_{MI}:113_{Ctrl} iTRAQ of proteins identified in R02 and R03. (D) Scatter $\log_2$ plots of measured 115_{NMI}:113_{Ctrl} iTRAQ ratio of proteins identified in R01 and R02. (E) Scatter $\log_2$ plots of measured 115_{NMI}:113_{Ctrl} iTRAQ ratio of proteins identified in R01 and R03. (F) Scatter $\log_2$ plots of measured 115_{NMI}:113_{Ctrl} iTRAQ ratio of proteins identified in R02 and R03. (G) Scatter $\log_2$ plots of measured 114_{MI}:115_{NMI} iTRAQ ratio of proteins identified in R01 and R02. (H) Scatter $\log_2$ plots of measured 114_{MI}:115_{NMI} iTRAQ ratio of proteins identified in R01 and R03. (I) Scatter $\log_2$ plots of 114_{MI}:115_{NMI} iTRAQ of proteins identified in R02 and R03. R01, replicate 1; R02, replicate 2; R03, replicate 3; r, Pearson correlation coefficient; ****, extremely significant.

2.4.3 Identification and quantification of differentially expressed proteins

Combined triplicate LC-MS/MS database search returned with a core set of 371 quantified proteins (FDR < 1%, unused protein score > 2 correspond to 99% confidence). At 95% peptide confidence level, nearly 92% of 371 proteins were identified based on two or more peptides, 71% of proteins were identified based on five of more peptides, and only 8% of proteins were identified based on single

peptide. Generally, these results are indicative of good quality identification of quantified proteins. In order to categorize CAD stage-specific biomarkers, calculated 114_{MI}:113_{Ctrl} and 115_{NMI}:113_{Ctrl} iTRAQ ratios were used to determine potential atherosclerotic-specific biomarkers, while calculated 114_{MI}:115_{NMI} iTRAQ ratios were used to flag candidate myocardial-injury-specific biomarkers.

Differentially expressed proteins were reliably selected based on the following criteria: (a) protein ratio-fold change ≥ 2.0 (up-regulation) or ≤ 0.5 (down-regulation) and p value < 0.05, (b) matching protein regulation trend in both combined triplicate searched results and average of individual triplicate searched result and (c) protein identified based on ≥ 2 peptides with 95% confidence and present in at least two of the triplicates. According to the strict criteria, from the initial list of 371 proteins, a total of 53 proteins were found to be significantly differentially expressed ($p < 0.05$) as listed in Table 2.3. Proteins that did not meet these stringent criteria were selectively disregarded. Of the 53 shortlisted candidates in the discovery phase, 27 proteins showed parallel up-regulated (26 proteins) or down-regulated (one protein) trends in both 114_{MI}:113_{Ctrl} and 115_{NMI}:113_{Ctrl} were marked as potential atherosclerotic-specific biomarkers. Conversely, proteins with expression discrepancies between MI relative to Ctrl and NMI relative to Ctrl groups were further analyzed using 114_{MI}:115_{NMI} iTRAQ ratios, and 26 proteins exhibiting at least two-ratio fold change were chosen as indicators of myocardial injury.

Table 2.3 List of proteins with significant altered expressions.

Unused[a]	Accession[a,b]	Protein Description[a,b]	Peptides (95%)[a]	MI(114):Ctrl (113)			NMI(115):Ctrl (113)			MI (114):NMI(115)		
				2D ratio[a]	1D mean ratio [b]	P value[a]	2D ratio[a]	1D mean ratio [b]	P value[a]	2D ratio[a]	1D mean ratio [b]	P value[a]
Shortlisted biomarkers of atherosclerosis												
41.91	P02760	Protein AMBP	57	18.20	7.83	0.010	14.86	7.79	0.008	1.43	1.11	0.109
3.31	P59666	Neutrophil defensin 3	2	14.59	6.98	0.040	12.36	6.85	0.012	1.13	1.02	0.925
12.49	P20851	C4b-binding protein beta chain	10	10.57	12.78	0.047	10.47	14.01	0.037	1.14	0.97	0.677
33.65	P08519	Apolipoprotein(a)	58	10.00	8.01	< 0.0001	4.13	4.59	0.002	2.68	1.97	0.000
18.55	P00742	Coagulation factor X	14	9.38	8.72	0.020	9.73	9.70	0.012	0.98	0.91	0.489
7.28	P02776	Platelet factor 4	7	6.43	7.00	0.004	6.25	7.98	0.010	1.03	0.89	0.102
43.05	P02749	Beta-2-glycoprotein 1	69	6.43	6.67	0.000	4.53	5.26	0.002	1.46	1.36	0.025
41.22	P04004	Vitronectin	71	5.92	4.29	0.018	3.50	3.07	0.009	1.82	1.41	0.064
53.78	P00736	Complement C1r subcomponent	50	5.81	4.99	0.004	3.56	3.82	0.007	1.69	1.37	< 0.0001
18.81	P07357	Complement component C8 alpha chain	19	4.66	4.46	0.006	2.49	3.55	0.002	1.89	1.23	0.010
48.40	P43652	Afamin	47	4.53	2.81	0.004	7.45	5.08	0.001	0.55	0.50	< 0.0001
8.86	P17936	Insulin-like growth factor-binding protein 3	7	4.53	2.65	0.009	5.70	3.42	0.017	0.84	0.82	0.009
17.88	P03951	Coagulation factor XI	17	4.45	3.55	0.020	3.84	3.16	0.009	1.11	1.06	0.453
56.31	P10643	Complement component C7	67	4.41	3.25	< 0.0001	4.13	3.37	0.004	1.07	0.95	0.462
38.29	O75882	Attractin	29	4.02	3.11	0.020	3.80	3.56	0.020	1.00	0.82	0.017
11.08	P27918	Properdin	7	3.87	3.68	0.004	4.17	4.30	0.000	0.92	0.85	0.214
36.15	P05160	Coagulation factor XIII B chain	22	3.87	4.23	0.005	3.63	3.86	< 0.0001	1.11	1.06	0.625
38.56	P07225	Vitamin K-dependent protein S	38	3.84	3.47	0.000	2.99	3.01	0.002	1.29	1.17	0.391
24.93	P05156	Complement factor I	23	3.80	4.05	0.002	3.05	3.74	0.004	1.29	1.12	0.058
66.13	P03952	Plasma kallikrein	56	3.66	2.74	0.001	4.53	3.80	0.001	0.79	0.69	0.000
78.44	P00734	Prothrombin	137	3.60	4.00	0.006	2.42	3.55	0.000	1.60	1.14	0.086
20.31	Q14520	Hyaluronan-binding protein 2	22	3.16	7.62	0.019	2.54	6.56	0.003	1.24	1.16	0.390
47.18	P00748	Coagulation factor XII	43	3.05	2.79	0.017	4.66	4.62	< 0.0001	0.63	0.57	0.018
13.77	Q96IY4	Carboxypeptidase B2	15	2.83	2.61	0.002	2.31	2.53	0.010	1.22	1.04	0.634
18.12	P26927	Hepatocyte growth factor-like protein	13	2.81	4.21	0.004	2.63	4.13	0.000	1.12	1.07	0.731
9.08	Q9UGM5	Fetuin-B	11	2.09	2.74	0.002	2.27	3.38	0.019	0.95	0.86	0.444
68.16	P02766	Transthyretin	105	0.10	0.12	< 0.0001	0.15	0.22	< 0.0001	0.54	0.43	0.000
Shortlisted biomarkers of myocardial injury												
11.58	P02735	Serum amyloid A protein	10	22.91	18.26	0.003	0.05	0.13	< 0.0001	59.16	43.96	0.001
4.62	P02741	C-reactive protein	3	10.86	6.29	0.027	0.39	0.46	0.028	21.88	11.97	< 0.0001
53.50	P01011	Alpha-1-antichymotrypsin	80	3.63	3.19	0.001	0.46	0.58	0.000	7.45	5.17	0.001
25.01	P02750	Leucine-rich alpha-2-glycoprotein	32	2.91	3.23	0.006	0.46	0.62	0.011	5.75	4.86	0.001
11.89	P12814	Alpha-actinin-1	7	0.79	0.66	0.009	0.17	0.23	0.000	5.01	3.48	0.085
37.07	P63261	Actin, cytoplasmic 2	53	0.55	0.56	< 0.0001	0.12	0.17	< 0.0001	4.92	3.52	0.006
25.52	P18206	Vinculin	21	0.67	0.51	0.000	0.14	0.15	< 0.0001	4.88	4.00	0.008
7.75	P35542	Serum amyloid A-4 protein	23	1.39	1.43	< 0.0001	0.29	0.45	0.004	4.61	3.35	0.043
33.31	P02748	Complement component C9	35	3.98	4.10	0.004	0.90	1.06	0.768	4.57	3.99	0.015
38.20	P22792	Carboxypeptidase N subunit 2	37	1.63	1.41	0.001	0.39	0.54	0.015	3.91	2.85	0.024
3.44	P78417	Glutathione S-transferase omega-1	2	1.85	1.38	0.198	0.50	0.60	0.010	3.70	2.33	0.023
13.70	P19652	Alpha-1-acid glycoprotein 2	43	1.47	1.18	0.007	0.39	0.59	0.002	3.66	1.97	0.001
31.49	P02649	Apolipoprotein E	44	1.47	1.28	0.036	0.42	0.52	0.001	3.31	2.36	0.017

[a,] 2D, iTRAQ ratio of each protein obtained from combined triplicate LC-MS/MS dataset; [b,] 1D, Mean iTRAQ ratio of each protein calculated from the individual triplicate LC-MS/MS datasets; MI, Myocardial infarction; NMI, Stable angina; Ctrl, Control; 113, 114 and 115, 8-plex isobaric tag.

Table 2.3 List of proteins with significant altered expressions.(Continued)

Unused[a]	Accession[a,b]	Protein Description[a,b]	Peptides (95%)[a]	MI(114):Ctrl (113)			NMI(115):Ctrl (113)			MI (114):NMI(115)		
				2D ratio[a]	1D mean ratio [b]	P value[a]	2D ratio[a]	1D mean ratio [b]	P value[a]	2D ratio [a]	1D mean ratio [b]	P value[a]
3.83	P07195	L-lactate dehydrogenase B chain	5	1.56	1.21	0.187	0.45	0.48	0.002	3.31	2.52	0.011
54.35	P21333	Filamin-A	36	0.88	0.63	0.001	0.29	0.25	< 0.0001	3.25	2.74	0.000
41.24	Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	42	4.02	3.68	< 0.0001	1.27	1.73	0.023	3.19	2.24	0.007
17.19	P18428	Lipopolysaccharide-binding protein	17	3.16	2.56	0.002	1.16	1.03	0.720	2.81	2.40	0.001
68.17	Q9Y490	Talin-1	66	0.60	0.61	0.003	0.22	0.27	< 0.0001	2.81	2.29	0.015
10.43	P07737	Profilin-1	9	0.67	0.65	0.000	0.27	0.30	0.000	2.47	2.29	0.013
5.77	P54802	Alpha-N-acetylglucosaminidase	3	1.22	1.26	0.659	3.05	4.34	0.192	0.40	0.37	0.014
30.80	P29622	Kallistatin	27	0.49	0.46	0.000	1.24	1.29	0.099	0.38	0.34	< 0.0001
19.18	P02753	Retinol-binding protein 4	34	0.32	0.30	< 0.0001	0.81	0.87	0.103	0.37	0.33	0.001
31.67	Q96PD5	N-acetylmuramoyl-L-alanine amidase	40	0.97	1.11	0.560	2.99	3.14	0.005	0.29	0.33	< 0.0001
19.39	P51884	Lumican	15	0.20	0.21	0.000	0.64	0.76	0.009	0.28	0.26	0.000
68.37	P06396	Gelsolin	73	0.15	0.18	< 0.0001	1.10	1.27	0.155	0.13	0.14	< 0.0001
60.12	P06727	Apolipoprotein A-IV	84	0.15	0.22	< 0.0001	1.60	1.81	0.024	0.10	0.14	< 0.0001

[a.] 2D, iTRAQ ratio of each protein obtained from combined triplicate LC-MS/MS dataset; [b.] 1D, Mean iTRAQ ratio of each protein calculated from the individual triplicate LC-MS/MS datasets; MI, Myocardial infarction; NMI, Stable angina; Ctrl, Control; 113, 114 and 115, 8-plex isobaric tag.

2.4.4 Functional comparative analyses of differential plasma proteome

The list of statistically relevant differentially expressed proteins were functionally compared and classified using FunRich V2.1.2 (221) to identify any altered biological processes and pathways that discriminates underlying atherosclerotic conditions from myocardial injury. Functional enrichment analyses were ranked using the Benjamini–Hochberg method corrected p value, and categories with $p < 0.05$ were identified as key functional categories. The biological process and biological pathway distribution of deregulated atherosclerotic-specific proteins and myocardial injury-specific proteins were illustrated in Figure 2.4. Differentially expressed atherosclerotic-specific proteins were enriched in biological processes mainly related to immune response (37%), protein metabolism (33.3%) and regulation of coagulation (3.7%) (Figure 2.4A). Additionally, as depicted in Figure 2.4B, the biological pathways related to atherosclerosis were significantly involved in the clotting cascade (52.9%), common pathway (35.3%), gamma-carboxylation of protein precursor (17.6%) and cell surface interaction at the vascular wall (17.6%). The above mentioned are known processes and pathways associated with CAD.

Conversely, as shown in Figure 2.4C, myocardial injury-specific deregulated proteins were represented predominantly in categories involving cell growth and/or maintenance (37%), immune response (19.2%) and lipid transport (3.78%). The putative pathways associated with myocardial injury (Figure 2.4D) were highly enriched in platelet activation (30.8%), integrin in angiogenesis (23.1%), cell-extracellular matrix interaction (15.4%) and smooth muscle contraction (15.4%). Comparative functional enrichment predicted some subsets of biological processes and pathways that were distinctive to either atherosclerotic-specific markers or myocardial injury-specific markers, a promising display that differentiates the severity of two sub groups of CAD. These findings also indicated the appropriate choice of clinical patient and control samples used in this biomarker discovery study.

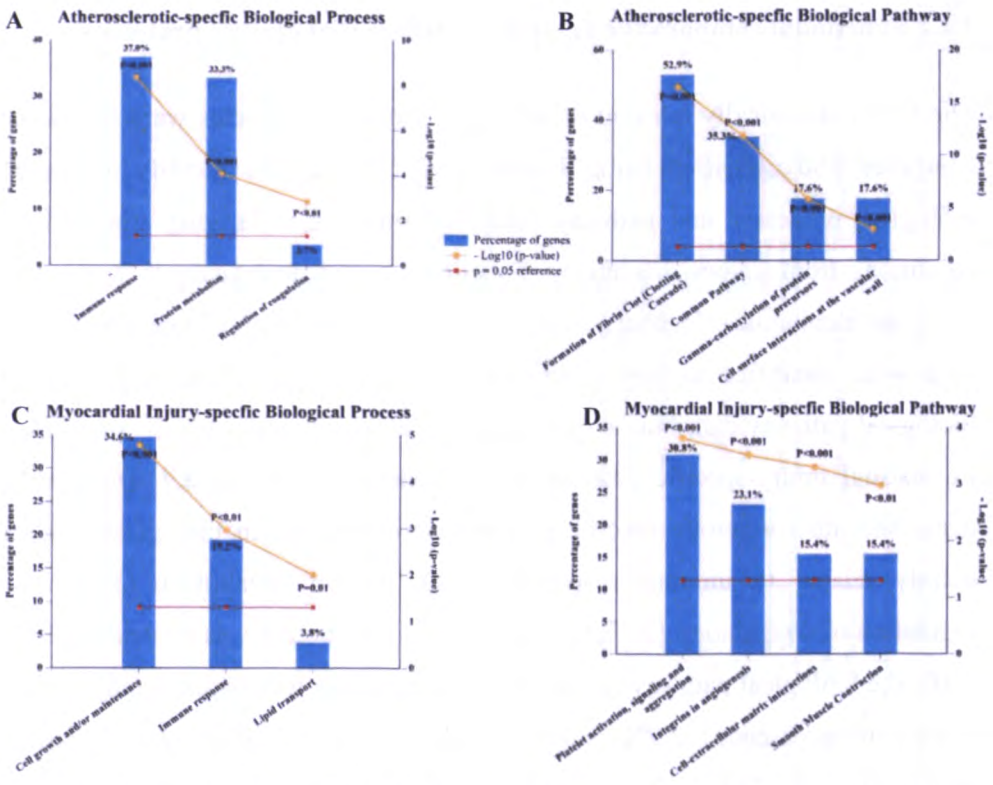

Figure 2.4. Gene ontology (GO)-based functional enrichment analyses. GO functional enrichment analyses of all statistically relevant deregulated atherosclerotic-specific proteins and myocardial injury-specific proteins. (A) Distribution of deregulated atherosclerotic-specific proteins based on their biological processes. (B) Distribution of deregulated atherosclerotic-specific proteins based on their biological pathways. (C) Distribution of deregulated myocardial injury-specific proteins based on their biological processes. (D) Distribution of deregulated myocardial injury-specific proteins based on their biological pathways. $p < 0.05$ indicates high enrichment. Predictions were generated from obtained from FunRich v2.1.2.

2.4.5 Selection of peptides for MRM assay

Moving forward from the initial discovery phase to the verification phase, label-free targeted multiplex MRM-MS approach was used in this study to verify the deregulated expression of 27 candidate biomarkers predictive of atherosclerosis, and 26 candidate biomarkers that are suggestive of myocardial injury (Table 2.3). MRM-MS assays can be coupled with label-free or stable isotope dilution (SID) to achieve relative or absolute protein quantification respectively, with the latter more frequently in used. The practicality of label-free MRM-based quantitation to be used as a first-level biomarker verification strategy for efficient, low cost screening of

large number of candidates biomarkers have been demonstrated by several studies (225-227).

In the preliminary LC-MRM-MS experiment, to verify 53 candidate biomarkers, unlabeled protein digest of pooled MI, NMI and Ctrl plasma were assayed in duplicates to evaluate the large number of targeted transitions (255 unmodified peptides, 999 transitions) derived from SRM Atlas (220) and Pinpoint™ V1.3 software. The optimal representative peptides and transitions were carefully assembled and poor performing proteins and peptides were sequentially eliminated. For each targeted peptide, positive confirmation and consistency of detection were manually checked and determined based on similarity in relative abundance, stability of peptide elution time and the identification of at least three co-eluting transition peaks, in both duplicate LC-MRM-MS runs of all pooled digest of MI, NMI and Ctrl samples. Poor quality, problematic peptides with inconsistent transitions and high background interferences, likely constrained by sample complexity which limits detection of lower abundance analyte, were removed. Based on reported experiences in the field, it is sensible to expect that only a small fraction of potential biomarkers identified with proteomics will eventually be clinically useful (225). Thus to further narrow down the list of potential markers, proteins represented with fewer than two positive unambiguous peptides in the initial MRM assay were excluded, to achieve further confidence in identification and quantification.

The refined targeted MRM list, included 49 peptides and 168 MRM transitions representing 8 potential markers of atherosclerosis and 15 candidate markers of myocardial injury. Based on our iTRAQ dataset, serum albumin (ALB), serotransferrin (TF) and alpha-2-macroglobulin (A2M) were selected as house-keeping proteins as their expressions were consistent across MI, NMI and Ctrl groups. These house-keeping proteins (10 peptides and 37 transitions) were included into the MRM-assay and quantified together with 23 targeted protein biomarkers in unfractionated digest of individual plasma samples.

2.4.6 Verification of candidate biomarkers by MRM assay

Simultaneous targeted label-free quantification of 23 candidate biomarkers, were performed on the same set of 49 (MI, n=15; NMI, n=20; Ctrl, n=14) individual plasma samples that were pooled for discovery iTRAQ experiments. The RT and normalized peak areas computed by Pinpoint V1.3 software were used for reproducibility examination. In analyzing the RT %CV based on 147 LC-MRM injection for each of the 59 assayed peptides, we found 56 peptides with CVs < 5% and 3 peptides with CVs < 10% (Table 2.4), demonstrating excellent within-run instrument reproducibility and stable HPLC system. Next, the quality of the reported peptide and protein relative abundance (peak area) within the triplicate LC-MRM-MS of Ctrl, MI and NMI samples were assessed for inconsistencies. Again as shown in Table 2.4, the overall and peptide peak areas %CV for Ctrl, MI and NMI were < 16%, < 15% and < 6% respectively, and the %CV of 26 protein peak areas in Ctrl, MI and NMI had median CVs of 4.51%, 4.71% and 1.52% correspondingly, both showing sufficient repeatability of relative protein quantification for analysis in complex plasma digest. In general, good consistencies were observed in the label-free MRM-experiments, the presented low technical and biological variations are in part due to simple plasma preparation procedure, properly maintained instrument and strict pre-selection of peptides with best response in the preliminary LC-MRM-MS survey.

In order to render the quantitative data more comparable, all individual protein peak areas were normalized to the arithmetic mean of the three house-keeping proteins. The $\log_2$ box-and-whiskers plots in Figure 2.5 A- C illustrated the overview distribution of the relative protein abundance of the house-keeping proteins (ALB, A2M, and TF) among the individual samples of Ctrl, MI and NMI groups. The median values of all three housekeeping proteins in Ctrl, MI and NMI plasmas were not far from each other, hence serving as good internal controls. In addition, there were no significance differences ($p > 0.05$) in the abundance level of ALB (Figure 2.5A), A2M (Figure 2.5B) and TF (Figure 2.5C) between MI versus Ctrl, NMI versus Ctrl and MI versus NMI groups, consistent with the iTRAQ results.

Next, we compared the relative expression levels of eight candidate markers of atherosclerosis between MI versus Ctrl and between NMI versus Ctrl groups. Based on MRM results, ATRN, TTR, AMBP and LPA showed no significant changes ($p > 0.05$) in expression levels in both MI versus Ctrl and NMI versus Ctrl groups, despite their apparent differential expression in iTRAQ findings. The expression levels of F10 ($p < 0.0001$), VTN ($p < 0.0001$), AFM ($p = 0.0121$) and MST1 ($p < 0.0001$), were found to be significantly higher in MI than in Ctrl (Figures 2.5D – 2.5F). Similarly, the expression levels of F10 ($p < 0.0001$), VTN ($p < 0.0001$), and MST1 ($p < 0.0001$), were found to be significantly higher in NMI relative to that of Ctrl. Only proteins (F10, VTN. MST1) presented with significant differences in both MI versus Ctrl and NMI versus Ctrl groups (Figure 2.5D – 2.5F) were considered as useful biomarkers for atherosclerosis.

Table 2.4. Reproducibility of MRM RT stability and integrated peak area relative quantification of targeted peptides and proteins.

Protein Accession	Gene Symbol	Sequence	Parent (Q1) m/z	Mean RT [a]	%CV[a]	Ctrl_Mean Peak Area[b]	Ctrl_ %CV[b]	MI_Mean Peak Area[b]	MI_%CV[b]	NMI_Mean Peak Area[b]	NMI_ %CV[b]
Shortlisted biomarkers of atherosclerosis											
P43652	AFM					5.43E+05	0.83	5.75E+05	4.20	8.73E+05	2.72
		DADPDTFFAK	563.76	26.00	1.97	2.44E+05	8.25	2.80E+05	5.58	4.37E+05	1.94
		FTFEYSR	475.22	22.06	2.51	3.63E+05	2.56	2.94E+05	3.30	4.36E+05	4.38
P02760	AMBP					1.32E+06	2.62	1.47E+06	7.00	2.11E+06	1.92
		AFIQLWAFDAVK	704.89	43.52	1.67	5.02E+04	1.22	2.43E+04	14.51	3.01E+04	1.07
		ETLLQDFR	511.27	27.00	2.26	1.28E+06	2.71	1.44E+06	7.21	2.07E+06	1.69
O75882	ATRN					4.85E+05	1.19	4.17E+05	2.28	4.47E+05	1.57
		CINQSICEK	519.24	30.80	3.33	1.42E+05	6.14	4.18E+04	2.90	1.04E+05	1.47
		WSVLPRPDLHHDVNR	614.32	15.96	4.63	4.14E+05	1.09	3.83E+05	2.57	3.44E+05	2.33
P00742	F10					1.06E+06	5.18	1.39E+06	1.90	1.64E+06	0.61
		MLEVPYVDR	561.29	26.39	2.70	1.94E+05	2.06	1.97E+05	1.76	2.51E+05	1.28
		SHAPEVITSSPLK	683.37	37.02	1.37	9.63E+05	1.04	1.19E+06	2.22	1.39E+06	0.72
P08519	LPA					2.73E+05	5.71	2.31E+05	8.03	2.15E+05	2.42
		GTDSCQGDSGGPLVCFEK	900.38	18.43	2.70	2.38E+05	6.56	2.12E+05	8.35	1.62E+05	0.94
		NPDAVAAPYCYTR	720.83	16.23	7.79	6.61E+04	2.12	3.46E+04	4.10	8.61E+04	3.43
P26927	MST1					1.19E+06	1.46	2.29E+06	5.18	1.85E+06	2.25
		CEIAGWGETK	547.25	20.80	3.45	4.80E+05	2.69	5.37E+05	3.90	6.07E+05	2.23
		MVCGPSGSQLVLLK	716.39	34.38	1.21	7.09E+05	0.86	1.75E+06	6.05	1.24E+06	3.52
P02766	TTR					8.88E+05	5.75	7.41E+05	6.02	7.59E+05	1.37
		GSPAINVAVHVFR	683.88	24.24	3.46	3.54E+05	3.77	1.73E+05	6.44	1.83E+05	1.75
		TSESGELHGLTTEEEFVEGIYK	819.06	35.09	1.81	6.48E+05	3.87	5.68E+05	5.90	5.97E+05	1.17
P04004	VTN					2.39E+06	4.37	4.36E+06	0.61	3.56E+06	1.29
		DVWGIEGPIDAAFTR	823.91	44.73	1.63	1.45E+06	6.99	2.78E+06	0.21	2.12E+06	0.82
		FEDGVLDPDYPR	711.83	27.62	1.82	9.44E+05	0.55	1.58E+06	1.67	1.44E+06	2.41
Shortlisted biomarkers of myocardial injury											
P63261	ACTG1					8.43E+05	5.11	1.09E+06	5.62	3.14E+05	1.92
		KDLYANTVLSGGTTMYPGIADR	1172.09	33.57	1.71	4.52E+05	3.57	6.19E+05	8.00	1.12E+05	3.60
		SYELPDGQVITIGNER	895.95	38.27	1.49	4.32E+05	2.76	4.81E+05	4.47	2.02E+05	0.99
P06727	APOA4					1.01E+06	2.53	6.73E+05	4.80	2.03E+06	1.48
		ISASAEELR	488.26	12.76	9.73	1.49E+06	8.78	5.51E+05	6.21	1.67E+06	1.80
		SELTQQLNALFQDK	817.92	39.90	1.31	2.73E+05	2.02	1.23E+05	5.28	1.90E+05	1.99
		SLAELGGHLDQQVEEFR	643.32	31.02	2.23	1.95E+05	1.07	1.82E+05	2.59	2.51E+05	0.69
P02649	APOE					2.30E+06	5.85	3.57E+06	8.43	1.80E+06	5.09
		SELEEQLTPVAEETR	865.93	24.95	1.81	9.38E+05	5.13	1.40E+06	5.37	5.79E+05	4.64
		WELALGR	422.74	26.73	2.39	1.50E+06	3.16	2.18E+06	12.49	1.22E+06	5.37
P22792	CPN2					3.69E+06	3.81	5.55E+06	5.72	2.67E+06	2.20
		LELLSLSK	451.78	27.17	2.61	2.84E+05	4.08	1.04E+05	3.65	1.55E+05	1.97
		LTVSIEAR	444.76	20.35	2.80	3.49E+06	4.01	5.45E+06	5.78	2.51E+06	2.56
P02741	CRP					2.01E+05	0.50	3.25E+06	11.87	1.82E+05	3.67
		AFVFPK	354.71	19.67	3.84	2.25E+05	7.15	2.41E+06	12.36	9.30E+04	0.90
		ESDTSYVSLK	564.77	15.05	8.55	1.08E+05	15.05	1.22E+06	8.79	1.34E+05	3.02
P21333	FLNA					1.85E+06	4.61	1.71E+06	5.26	1.93E+06	1.30
		CSGPGLER	409.70	15.93	3.25	7.50E+05	4.21	4.84E+05	3.10	7.57E+05	1.15
		ENGVYLIDVK	575.31	21.74	2.42	1.13E+06	3.54	1.24E+06	4.87	1.18E+06	1.69

m/z, mass-to-charge ratio; RT, retention time; [a] values derived from individual 147 LC-MRM runs [b] values derived from triplicate LC-MRM-MS runs.

Table 2.4. Reproducibility of MRM RT stability and integrated peak area relative quantification of targeted peptides and proteins. (Continued)

Protein Accession	Gene Symbol	Sequence	Parent (Q1) m/z	Mean RT [a]	%CV[a]	Ctrl _Mean Peak Area[b]	Ctrl_ %CV[b]	MI_Mean Peak Area[b]	MI_%CV[b]	NMI_Mean Peak Area[b]	NMI_ %CV[b]
Q06033	ITIH3					4.13E+05	5.77	7.14E+05	1.26	4.97E+05	1.05
		DYIFGNYIER	645.31	34.32	1.15	1.92E+05	5.48	3.29E+05	1.53	2.58E+05	1.47
		EVSFDVELPK	581.80	30.21	1.87	2.38E+05	6.03	3.84E+05	1.08	2.39E+05	1.74
P02750	LRG1					1.69E+06	4.77	3.40E+06	7.39	2.19E+06	0.95
		DLLLPQPDLR	590.34	32.80	1.29	1.20E+06	6.61	2.59E+06	6.86	1.76E+06	1.31
		VAAGAFQGLR	495.28	17.48	4.46	6.29E+05	4.45	8.50E+05	2.36	4.24E+05	0.72
P51884	LUM					5.46E+06	6.94	2.49E+06	13.70	7.60E+06	0.73
		FNALQYLR	512.78	29.80	2.19	2.17E+05	9.30	4.60E+04	2.56	1.21E+05	4.86
		NNQIDHIDEK	409.20	15.90	3.47	6.57E+06	5.87	2.74E+06	4.62	7.48E+06	0.88
P54802	NAGLU					4.03E+05	5.73	2.26E+05	8.44	4.95E+05	0.40
		GDTVDLAK	409.72	15.94	3.46	3.22E+05	1.18	1.72E+05	4.87	3.88E+05	0.79
		YGVSHPDAGAAWR	693.83	41.58	1.18	1.23E+05	15.15	6.08E+04	1.68	1.07E+05	2.85
P19652	ORM2					3.33E+06	5.00	7.04E+06	2.46	3.46E+06	1.61
		SDVMYTDWK	572.75	23.37	1.98	1.05E+06	6.46	1.36E+06	4.88	1.03E+06	4.03
		TEDTIFLR	497.76	22.52	2.43	1.89E+06	5.70	4.85E+06	2.50	1.86E+06	1.42
		TLMFGSYLDDEK	709.83	33.04	1.30	3.72E+05	7.12	8.25E+05	2.95	5.65E+05	1.55
P0DJI8	SAA1					8.19E+04	0.53	3.17E+06	2.46	4.54E+04	1.99
		FFGHGAEDSLADQAANEWGR	726.66	30.33	2.76	6.38E+04	5.87	3.04E+06	1.69	2.74E+04	4.63
		SFFSFLGEAFDGAR	775.87	48.87	1.31	2.34E+04	4.95	3.71E+05	10.27	2.11E+04	4.15
P35542	SAA4					4.57E+06	3.06	7.85E+06	2.87	4.06E+06	3.36
		EALQGVGDMGR	566.77	17.16	3.21	2.26E+06	4.66	3.17E+06	6.12	1.50E+06	2.69
		FRPDGLPK	465.26	12.42	4.08	2.94E+06	4.71	7.51E+06	2.14	2.56E+06	3.73
P01011	SERPINA3					1.04E+06	4.41	1.72E+06	2.10	6.88E+05	0.47
		LYGSEAFATDFQDSAAAK	946.44	31.41	2.04	5.50E+05	5.66	1.11E+06	3.25	3.93E+05	0.88
		MEEVEAMLLPETLK	816.92	38.13	1.42	4.90E+05	3.48	6.10E+05	0.00	2.95E+05	1.87
Q9Y490	TLN1					2.09E+06	6.11	1.37E+06	3.61	2.58E+06	1.79
		NCGQMSEIEAK	605.27	37.79	1.86	7.45E+05	5.16	5.71E+05	5.73	8.67E+05	3.87
		SAQPASAEPR	507.25	30.87	2.69	7.25E+05	2.37	2.80E+05	4.85	8.91E+05	1.76
		TLAESALQLLYTAK	507.96	31.04	1.63	9.32E+05	2.86	5.15E+05	2.03	8.25E+05	0.69
House-keeping proteins											
P01023	A2M					6.43E+06	5.09	6.11E+06	4.62	6.32E+06	1.27
		AIGYLNTGYQR	628.33	19.88	2.46	2.11E+06	4.51	1.58E+06	11.04	1.76E+06	1.70
		LSFYYLIMAK	624.84	40.95	1.45	3.52E+05	3.86	1.75E+05	2.00	2.79E+05	2.33
		VGFYESDVMGR	630.29	25.22	1.84	4.11E+06	4.94	4.40E+06	2.67	4.29E+06	2.37
P02768	ALB					4.19E+08	3.87	4.17E+08	3.36	4.50E+08	1.22
		AVMDDFAAFVEK	671.82	38.31	1.00	6.35E+07	9.52	4.74E+07	5.34	7.55E+07	1.86
		HPDYSVVLLLR	656.38	30.69	2.12	5.47E+06	7.11	5.36E+06	4.66	4.56E+06	1.62
		LVAASQAALGL	507.30	31.18	1.30	3.78E+07	7.67	2.24E+07	4.04	4.28E+07	2.21
		LVNEVTEFAK	575.31	21.57	2.13	3.12E+08	2.28	3.42E+08	3.71	3.26E+08	1.51
P02787	TF					2.70E+07	0.74	2.60E+07	2.40	3.09E+07	1.04
		EGYYGYTGAFR	642.29	24.91	1.77	6.89E+06	1.17	4.91E+06	2.01	6.81E+06	2.26
		SASDLTWDNLK	625.31	26.05	1.77	9.07E+06	0.50	1.02E+07	1.70	1.05E+07	2.86
		YLGEEYVK	500.75	20.15	2.40	1.10E+07	1.89	1.09E+07	4.00	1.36E+07	4.09

m/z, mass-to-charge ratio; RT, retention time; [a.] values derived from individual 147 LC-MRM runs [b.] values derived from triplicate LC-MRM-MS runs.

Finally, the relative expression levels of 15 candidate proteins of myocardial injury were examined between individual MI and NMI plasma. Among them, five proteins (Figure 2.5G – 2.5L): SERPINA3 ($p < 0.0001$), CRP ($p = 0.00691$), SAA1 ($p = 0.00116$), ORM2 ($p < 0.0001$), CPN2 ($p < 0.0001$), and ACTG1 ($p < 0.0001$) were significantly higher in MI than in NMI, and four proteins (Figure 2.5M -2.5P): APOA4 ($p < 0.0001$), LUM ($p < 0.0001$), NAGLU ($p < 0.0001$) and TLN1 ($p = 0.000769$) were seen to be significantly lower in MI than in NMI. The expression levels of 5 proteins including APOE, LRG1, FLNA, SAA4 and ITIH3 appeared insignificant ($p > 0.05$) between MI and NMI groups although they were in the discovery experiment.

The differential expression patterns of proteins with significant differences ($p < 0.05$) obtained from MRM were compared to the discovery iTRAQ results as displayed in Table 2.5. The average of all normalized peak area from each Ctrl, MI and NMI groups were used to establish the differential protein ratio-fold change between MI versus Ctrl (MI:Ctrl), NMI versus Ctrl (NMI:Ctrl) and MI versus NMI (MI:NMI). The differential expression ratios of 12 proteins exhibited similar patterns in both discovery and verification analyses, with the exception of TLN, which was filtered out as a candidate marker for myocardial injury. The inconsistencies observed in the relative protein quantitation results between iTRAQ and MRM could be attributed by several factors. As pooling strategy was used in iTRAQ experiment, the protein abundance in individual samples were averaged out for quantitation. Whereas, independent samples were analyzed for MRM and individual biological variation in protein abundance were estimated. Thus, the inevitable presence of excessive biological variation could lead to the fluctuated relative expression results detected between the iTRAQ and MRM experiments.

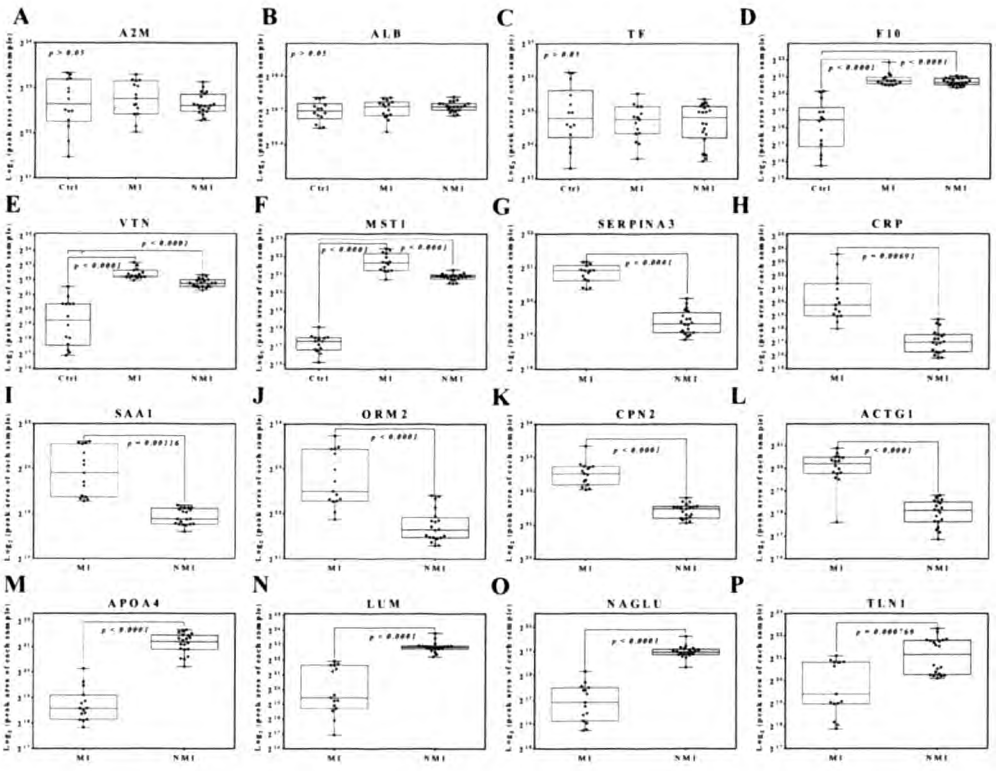

Figure 2.5. MRM-derived relative expression levels of targeted plasma proteins. Relative expression of targeted plasma proteins in 49 human plasma samples (17 patients who have had myocardial infraction [MI] and 26 stable angina [NMI] patients, 14 control subject with no prior history of angina [Ctrl]) were shown as linear box-and-whiskers plots. Line represents median, box represents 25th to 75th percentiles and whisker represent maximum and minimum ranges and dot represent individual values. (A) Expression levels ($p > 0.05$) of A2M. (B) Expression levels ($p > 0.05$) of ALB. (C) Expression levels ($p > 0.05$) of TF. (D) Expression levels ($p < 0.05$) of F10. (E) Expression levels ($p < 0.05$) of protein VTN. (F) Expression levels ($p < 0.05$) of MST1. (G) Expression levels ($p < 0.05$) of SERPINA3. (H) Expression levels ($p < 0.05$) of CRP. (I) Expression levels ($p < 0.05$) of SAA1. (J) Expression levels ($p < 0.05$) of ORM2. (K) Expression levels ($p < 0.05$) of CPN2. (L) Expression levels ($p < 0.05$) of ACTG1. (M) Expression levels ($p < 0.05$) of APOA4. (N) Expression levels ($p < 0.05$) of LUM. (O) S Expression levels ($p < 0.05$) of NAGLU. (P) Expression levels ($p < 0.05$) of TLN1.

In addition, the differences in the complexity of fractionated and unfractionated plasma analyzed for iTRAQ and MRM respectively may have influence the outcome of the quantitation. The inclusion of all contributing peptides, including splicing isoforms and post-translation modified peptides for quantitation in iTRAQ versus the used of targeted unmodified peptides for quantitation MRM may also have contributed to these variations. Overall, the MRM verification analyses revealed that 12 out of 23 protein candidates exhibited significant ($p < 0.05$, > 2 ratio fold-change)

differential expression patterns that were in good agreement with discovery proteomics. These observations indicates the capability of multiplex MRM-based assay in the verification of numerous candidates with adequate reproducibility from discovery experiments in crude plasma.

Table 2.5. Comparison of differential expression results obtained from MRM and iTRAQ experiments.

Protein Accession	Gene Symbol	Protein Description	MRM			iTRAQ		
			MI: Ctrl	NMI: Ctrl	MI: NMI	MI: Ctrl	NMI: Ctrl	MI: NMI
Shortlisted biomarkers of atherosclerosis								
P26927	MST1	Hepatocyte growth factor-like protein	22.89	12.62	1.81	2.81	2.63	1.12
P04004	VTN	Vitronectin	6.18	4.01	1.54	5.92	3.50	1.82
P00742	F10	Coagulation factor X	4.43	3.99	1.11	9.38	9.73	0.98
Shortlisted biomarkers of myocardial injury								
P0DJI8	SAA1	Serum amyloid A protein	68.35	0.79	86.57	1.39	0.29	4.61
P02741	CRP	C-reactive protein	13.94	0.94	14.78	10.86	0.39	21.88
P63261	ACTG1	Actin, cytoplasmic 2	1.73	0.39	4.41	0.55	0.12	4.92
P01011	SERPINA3	Alpha-1-antichymotrypsin	1.70	0.63	2.69	3.63	0.46	7.45
P22792	CPN2	Carboxypeptidase N subunit 2	2.10	0.96	2.18	1.63	0.39	3.91
P19652	ORM2	Alpha-1-acid glycoprotein 2	4.09	1.92	2.13	1.47	0.39	3.66
Q9Y490	TLN1	Talin-1	0.25	1.25	0.20	0.60	0.22	2.81
P54802	NAGLU	Alpha-N-acetylglucosaminidase	0.31	1.18	0.26	1.22	3.05	0.40
P06727	APOA4	Apolipoprotein A-IV	0.63	3.14	0.20	0.15	1.60	0.10
P51884	LUM	Lumican	0.25	1.25	0.20	0.20	0.64	0.28

MI, Myocardial infarctiona; NMI, Stable angina; Ctrl, Control; MRM, multiple reaction monitoring, iTRAQ, isobaric tags for relative and absolute quantification.

2.4.7 Candidate biomarkers with significant differential expression

Label-free MRM verification of the discovery findings across 49 independent plasma samples have facilitated the prioritization of 12 candidate biomarkers worthy for further investigations in larger cohort verification and clinical validation. The biological function of each protein with significant differences were discussed sequentially for their implications towards the pathophysiology of atherosclerosis and/or myocardial injury. CAD is a multifactorial disease, the identification of a panel of multi-markers that convey specific information on particular aspects of the pathophysiological process, would increase the predictive performance for risk stratification than the reliance on a single unique biomarker.

In the presence of acute MI, the relative expressions of VTN, F10 and MST1 did not differ from underlying atherosclerotic conditions, however, the expressions were

significantly up-regulated in both MI and NMI when compared with controls. Hence, highlighting these proteins as potential candidate biomarkers for the prediction of coronary atherosclerosis. Of the three candidate proteins predictive for atherosclerosis, increased plasma expression VTN have been reported to correlate with the severity of CADs (228), as such it will not be elaborated here. Elements of the coagulation system plays an important role in the maintenance of endothelial integrity, and in the initiation, progression or acceleration of atherosclerosis (229). F10 is an inactive pro-enzyme of the coagulation cascade, and is activated into F10a by tissue factor/F12a complex, which triggers thrombin formation via protease-activated receptor (PAR) signaling (230). It has been demonstrated that picomolar concentration of F12a can trigger robust signaling in keratinocyte cells expressing tissue factor and PAR2 in the presence of F10 (230). Additionally, inflammatory stimuli and vascular injury have shown to increase PAR2 expression in vascular endothelial cells (231) and vascular SMCs (232), promoting pro-coagulation characteristics. In this aspect, increased levels of F10 of ~ four ratio-fold change in both MI and NMI groups, possibly suggest its participation in the induction of PAR2 expression at sites of endothelial inflammation and/or injury during the pathogenesis of atherosclerosis, however, this speculation is at the moment is unknown.

MST1, a plasminogen-related growth factor, is secreted as an inactive precursor by hepatocytes into the circulation and cleaved into mature form by enzymes of the coagulation cascade and macrophage membrane proteases (233, 234). It has been shown that MST1 exert inhibition of inducible nitric oxide synthase (iNOS) in macrophages (235, 236). The expression of iNOS in macrophage, is induced by inflammatory mediators as part of host defense response (237). In atherosclerosis, iNOS has been proposed to be a pro-atherogenic mediator of increased oxidative stress and impaired vascular function (238). As the implication of MST1 in atherosclerosis has not been established, it is reasonable to hypothesize that the observed increased expressions of MST1 (~23 ratio-fold change in MI group, ~13 ratio-fold change in NMI group) habour atheroprotective functions by reducing the production of pro-inflammatory iNOS under atherosclerotic conditions, hence suppressing further endothelial damage. Based on the abovementioned pathological

association of F10 and MST1 to atherosclerosis, we found that endothelial dysfunction, the initial step of atherogenesis, in common to both proteins. Of particular importance was that both F10 and MST1 has not been previously reported as biomarker candidates for atherosclerosis, thus the parallel over expression of these proteins in both sub-groups of CADs relative to control, highlights their potential as novel predictors of atherosclerosis.

In attempt to identify biomarkers that discriminates myocardial injury from stable atherosclerotic conditions, we compared the expression levels of patients who have had MI against patients who were diagnosed with stable angina. Based on discovery and verification analyses, we reported nine proteins with significant altered expressions in MI as probable candidates for myocardial injury. Among these candidates, SAA1(~86 ratio-fold change) , CRP (~15 ratio-fold change), SERPINA3(~2.7 ratio-fold change), CPN2 (~2.2 ratio-fold change), ORM2 (~2.1 ratio-fold change), ACTG1 (~4.4 ratio-fold change), expressions were elevated in MI, while the expression levels of APOA4 (~0.2 ratio-fold change), LUM (~0.2 ratio-fold change) and NAGLU (~0.26 ratio-fold change) were conspicuously under expressed in MI. The inflammatory profile in patients who have had acute MI is evidently reflected by the raised levels of four acute phase proteins, CRP, SAA1, ORM2 and SERPINA3. As increased levels of CRP (239-241) and SAA1 (64, 242-244) have been extensively-documented for their pathological association with MI and evaluated for diagnostics utility, they will not be reiterated here. Although intense inflammatory response following MI is essential for cardiac healing and scar formation (245), prolonged inflammatory response has been implicated in adverse heart remodeling and heart failure (246). The exact physiological role of ORM2 is obscure despite its first description in 1950s (247). Intestinal and renal ischemia-reperfusion (I/R) functional studies on ORM2 revealed protective role against organ injury and apoptosis and inflammation respectively (248, 249). Therefore, ORM2 has been has been thought to possess anti-inflammatory (248, 250) and immunomodulatory properties (251). These experimental findings led us to hypothesize that increase expression of ORM2 in MI, may act to suppress excessive

inflammatory response, thus exerting a protective mechanisms against tissue damage associated with I/R injury.

SERPINA3, an acute phase reactant that inhibits ECM degrading neutrophil cathepsin G and mast cell chymase (252). Inhibition of mast cell chymase have shown to reduce atherosclerotic plaque progression and instability in APOE deficient mice (253). Increased levels of SERPINA3 has been linked to the pathology of Alzheimer's diseases (AD) (254, 255). It has been demonstrated that in the presence of APOE, SERPINA3 influence amyloid deposition in transgenic mouse model of AD (256, 257). With reference from these aforementioned studies, elevated SERPINA3 in MI patients could be suggestive of atherosclerotic plaque remodeling and/or destabilization, hence increasing the risk of heart attack. The large subunit of CPN, CPN2 stabilizes and protects the small active subunit CPN1 from degradation (258). CPN function to inactivate bradykinin and complement anaphylatoxin (C3a, C4a, and C5a) (259, 260). It has been suggested that CPN may be involved in the inhibition of fibrinolysis (261). The exposure of plaque components (collagen and tissue factor) to the circulation upon erosion or rupture induces thrombus formation (262). Observed up-regulation of CPN in MI may therefore could reflect an indication of clot formation over a rupture plaque, and the physiological attempt in containment of inflammation by hindering complement activation, prolonged elevation may eventually result in a clinical event. The role of CPN in anti-fibrinolysis following MI warrants further investigation to evaluate the possible beneficial or detrimental effects on vessel patency, which is at present is unknown. ACTG1 functions to maintain the cytoskeletal integrity and the repair the cytoskeleton (263). It has been shown that ACTG1 plays key roles in endothelial cell motility and neo-vessel angiogenesis (264), as well as in the regulation of epithelial gap junctions (265). Though no concrete association has been reported, the marked elevation of ACTG1 observed post-infarction may likely mediate myocardial wound healing and tissue repair.

Of the three candidate markers of myocardial injury that were under expressed, plasma APOA4, a major component of anti-atherogenic HDL, has well-assessed for

its inverse associations with risk of CADs (266-268), for this reason it will not be discussed here. LUM, a collagen-associated ECM protein, is involved in collagen fibrils assembly, tissue repair, regulation of cell proliferation, cell migration and cell adhesion (269, 270). Increased LUM expression has been reported during fibrosis wound repair process in ischemic and reperfused rat heart (271). In coronary atherosclerosis, LUM was reported to localize in the thickened intimal stroma and vascular smooth muscle cells in intima, facilitating collagen fibrillogenesis in atherosclerosis (272). In addition, deficiencies in LUM expression have resulted in cardiomyocyte hypertrophy with altered collagen architecture (273). Unfortunately, the role of LUM in plaque stability and vascular pathology remains elusive. Referencing from these studies, the stability of atherosclerotic plaque may be associated with the regulation of collagen assembly by LUM. Based on our quantitative results, we hypothesized that the decreased expression levels of LUM post-MI is indicative of plaque destabilization and rupture, hence increasing the risk of myocardial infarction.

NAGLU is a lysosomal acid hydrolase that is involved in degradation of heparan sulfate (HS) glycosaminoglycans within the lysosomes (274). Deficiencies in NAGLU causes Sanfilippo syndrome, a rare inherited disorder characterize by the abnormal accumulation of HS (275). In this study, none of the patients were diagnosed with Sanfilippo syndrome, however an acquired under expression of NAGLU was observed in the post-infarction patient group. To our knowledge, there are no available literatures associating NAGLU with CADs. In attempt to establish the linkage for the decreased NAGLU expression, we directed our focus to the matrix proteoglycan HS. In atherosclerosis, a decrease in the amount of HS in atherosclerotic vessel walls have been reported (276-278), and the cause for the loss of HS has been thought to be a general inflammatory reaction. The loss of HS have been shown to promote retention of lipoprotein to subendothelial matrix, likely to confer pro-atherogenic consequences (276). As the consequences following the loss of HS remains unclear, we hypothesized that the decreased in NAGLU could be a physiological response to prevent further degradation of already reduced HS in atherosclerotic vessel walls, limiting the progression of atherogenesis pathogenesis.

At present, no reports has examined the relationship between peripheral ORM2, SERPINA3, CPN2, LUM, ACTG1 and NAGLU expression levels with myocardial injury, to the best of our knowledge. We postulate that elevated levels of F10, SERPINA3, CPN2 and decreased levels of LUM may confer activities that accelerate the pathogenesis of atherosclerosis and/or myocardial tissue injury. Whilst the increase abundance in MST1, ORM2, ACTG1 and decreased expression of NAGLU has been thought to induce activities that confers atheroprotection and/or in the mediation of tissue healing. The functional linkage analyses of these novel candidates to myocardial injury broadly reflects two process including plaque destabilization (SERPINA3, CPN2, LUM) and protective mechanisms against further damage (ORM2, ACTG1, NAGLU). In addition, further studies of these proteins will likely uncover their roles following myocardial injury. Here, we have proposed a multi-marker diagnostic panel consisting of eight novel biomarkers that specifically reflects and discriminates the multifactorial pathophysiology of atherosclerosis (F10, MST1) and myocardial injury (ORM2, SERPINA3, CPN2, LUM, ACTG1, NAGLU). We have also identified previously established proposed markers of CAD, including VTN, CRP, SAA1 and APOA4. We are the first to report a verified signature of significantly expressed multi protein markers poised to effectively differentiate underlying atherosclerotic conditions from myocardial injury in a single study. Future investigations into the roles of these candidates would yield promising insights into the mechanisms governing atherosclerosis and plaque destabilization. These novel candidates will be further pursued and assessed in large scale verification and clinical validation studies. Hopefully, these promising biomarkers possess high predictive values which are necessary for accurate patient risk stratification, and in near future, be established as diagnostic biomarkers.

2.5 Conclusion

In this study, we have developed a multi-marker panel consisting of eight novel plasma proteins that predicts and discriminates atherosclerosis from myocardial injury. We have showed that systematic filtering of low confidence and poor performing proteins identified from the quantitative discovery findings through the

verification phase, have facilitated the prioritization of 12 candidate biomarkers worthy for further investigations in larger cohort verification and clinical validation. Of the 12 candidate markers, two atherosclerotic-specific markers (F10, MST1) and six myocardial injury-specific markers (ORM2, SERPINA3, CPN2, LUM, ACTG1, NAGLU) were found to be novel. Functional pathological linkages analyses have suggest the involvement of F10 and MST1 in endothelial dysfunction, SERPINA3, CPN2, LUM in plaque destabilization and ORM2, ACTG1, NAGLU in protective mechanisms, which conveys the specific signatures that potentially differentiates underlying atherosclerotic conditions from myocardial injury. This newly discovered knowledge should prompt future functional investigations into the roles of these proteins in atherosclerosis and plaque destabilization. We showed that discovery and verification of novel candidate biomarkers of atherosclerosis and myocardial injury in undepleted plasma, is achievable with iTRAQ-label-free MRM based approach.

Chapter 3

Profiling Plasma Extracellular Vesicles Reveals Biomarkers and Potential Therapeutic Targets for Myocardial Ischemic Injury

Chapter 3

3.1 Abstract

Myocardial infarction (MI), an undesirable clinical outcome of coronary artery disease (CAD) triggers an intense inflammatory reactions that is necessary for myocardial healing. However, without timely repression of inflammatory response, post-infarction remodeling and heart failure could ensue. Hence, biomarker of myocardial ischemia and therapeutic strategies that either minimizes cardiac injuries are highly sought after. The well documentation on the involvement of extracellular vesicles (EVs) in a multitude of pathophysiological processes have made them clinically attractive diagnostic biomarkers and drug vectors for therapeutic interventions. Using label-free quantitative proteomics approach, we established the differential proteome of plasma EVs from patients who have had myocardial infarction (MI) and from patient with stable angina as control with non-myocardial infarction (NMI), this categorization will facilitate the detection of myocardial injury specific markers. We presented, for the first time, the comparative proteomics profiling on 252 perturbed EV proteins with 1.2-fold change cut-off in clinical samples, and showed that EVs mediates in myocardial damage and myocardial healing processes following MI. A panel of six potential EV markers of myocardial injury reflecting post- infarction pathways of complement activation (C1Q1A, ~3.23 fold-change, $p = 0.012$; C5, ~1.27 fold-change, $p = 0.087$), lipoprotein metabolism (APOD, ~1.86-fold change, $p = 0.033$; APOC3, ~2.63-fold change, $p = 0.029$) and platelet activation (GP1BA, ~9.18-fold change, $p < 0.0001$; PPBP, ~4.72-fold change, $p = 0.027$) was proposed. The increased levels of C1Q1A, C5, APOC3, GP1BA and PPBP have been thought to aggravate post-infarction inflammation and myocardial tissue damage. While, elevated levels of APOD may possess atheroprotective activities and plaque stability. Paradoxically, all components of EV-derived fibrinogen were observed to be down- regulated, which may be a compensation mechanism to suppress coagulation post-MI. C1QA ($p = 0.005$) and C5 ($p = 0.0021$), the respective important key initiator and terminal effector of complement activation in MI response, were selected for validation in another cohort of 43 individual patients using Luminex assay. The

new insights gained from this study would prompt further investigations that drive development of novel EV-based diagnostic and therapeutic strategies that will benefit patients with CAD, improving their survival and quality of life.

3.2 Introduction

Coronary artery disease (CAD) is the principal cause of cardiovascular mortality worldwide, with increasing occurrence in low income and middle income countries (279). The prevalence of CAD will continue to climb in parallel with global population ageing (280), imposing major socioeconomic burden. Myocardial infarction (MI) caused by thrombotic occlusion of the coronary artery and angina pectoris which arises from stenosis of the coronary arteries are the two main clinical manifestations of CAD (281). Although recommended treatments are in place to alleviate initial cardiac damage, circulatory biomarkers of myocardial ischemic injury and therapeutic strategies that offers early diagnosis and improvement in myocardial tissue repair respectively, are actively sought after.

The orchestrated inflammatory cascade triggered post-MI is essential for myocardial healing, but excessive and prolonged inflammatory response could lead to adverse post-infarcted heart remodeling and subsequent pathogenesis of heart failure (245, 282). The roles and modulatory effects of complement and inflammatory mediators in tissue repair and damage are extensively studied both in experimental and clinical models of MI. Timely repression and containment of inflammation is critical for effective myocardial tissue healing (245). As such, thorough investigations of various anti-inflammatory strategies were prompted, including – inhibition of complement components or of its receptors (283-286); free radical scavengers or nonenzymatic antioxidants intervention (287-290) and depletion of infiltrating neutrophils (291-294). Though numerous experimental studies have promisingly showed the reduction of infarct size with the use of certain anti-inflammatory approach mentioned earlier, direct beneficial translation into clinical context unfortunately has been disappointing, partly owing to the complex pathophysiology and heterogeneity of human diseases. Hence, the search for novel new insights into

the pathophysiology following MI and the identification of novel target is of paramount importance.

Extracellular vesicles (EVs), including microvesicles and exosomes, are diverse population of membrane-bound entities secreted by almost all cell types into the extracellular fluids (295). The growing body of functional studies on the pathophysiological roles of EVs have evidently showed their involvement in immune response regulation (296, 297); cell-cell communication via the transfer of bioactive molecules, including proteins (199, 298), lipids (299, 300) and nucleic acid (194, 301). The influence of EVs in a multitude of physiological regulations (302, 303) and pathological disorders (201, 304, 305) made them highly desirable targets as clinical biomarkers and drug delivery vectors for therapeutic interventions. Recent studies have demonstrated the therapeutic potential of transplanted mesenchymal stem cells (MSCs)-derived exosomes in reducing myocardial injury and improving cardiac function following MI in animal models (212, 213). The ability of EVs to deliver proteins carried within to elicit a functional response, made them highly desirable vehicle for drug delivery as well as useful disease biomarkers. However, systemic differential proteomics studies on circulating EVs following MI and their plausible participation and exploitation as biomarkers for MI are under explored and less defined in depth at present.

The objectives in this study are to interrogate the deregulated plasma EV proteome to gain insights of activities modulated following MI, and to identify candidate targets that could have potential diagnostic and/or therapeutic utility. As a first step to understand the role of EVs following MI, we established the differential EVs proteomes isolated from patients with acute MI (< 5 days) and from patients with underlying stable angina for at least a month as non-MI (NMI) control, using label-free quantitative proteomics approach. Functional comparative proteomics profiling of perturbed proteins in EVs of MI with NMI were presented for the first time, our analyses showed active participation of EVs in myocardial damage and healing processes following MI. Interestingly, all components of EV-derived fibrinogen were down-regulated. Herein, we proposed a panel of six potential EV markers of

myocardial injury reflecting post-infarction pathways of complement activation (C1Q1A, ~3.23 fold-change, p = 0.012; C5, ~1.27 fold-change, p = 0.087), lipoprotein metabolism (APOD, ~1.86-fold change, p = 0.033; APOC3, ~2.63-fold change, p = 0.029) and platelet activation (GP1BA, ~9.18-fold change, p < 0.0001; PPBP, ~4.72-fold change, p = 0.027). The differential expression of C1QA (p = 0.005) and C5 (p = 0.0021), the respective important initiator and terminal effector of complement activation following MI were validated using Luminex assay. The new knowledge obtained from this study could be used to drive the development of novel EV-based diagnostic and therapeutic strategies that may provide substantial positive impact on patient care.

3.3 Materials and Methods

3.3.1 Chemicals

All water and acetonitrile (ACN) used in this experiment were of HPLC grade (Thermo Scientific, Waltham, MA). All chemicals were purchased from Sigma-Aldrich (St Louis, MO) unless stated otherwise. All buffers used in proteomic sample preparation were supplemented with protease inhibitors (1:50, v/v) and phosphatase inhibitors (1:10, v/v) (Roche Diagnostics, Mannheim, DE).

3.3.2 Patient recruitment

The study was approved by the National Healthcare Group Domain Specific Review Board (NHG DSRB). All recruited participants were informed and written consent was obtained. Thirty-five consecutive patients presented with CAD, undergoing coronary artery bypass graft surgery (CABG) were enrolled in this study. Twenty patients have stable angina pectoris (Group 1; NMI), and 15 patients have unstable angina pectoris, (Group 2; MI). Patients grouped under MI have had recent myocardial infarction (< 5 days). Patients grouped under NMI have had stable angina for period more than a month. CAD was confirmed by coronary angiogram and myocardial infarction was confirmed by ECG and cardiac troponin (cTn) test. Patient demographics and clinical history were documented and depicted in Table 2.1.

3.3.3 Plasma sampling and extracellular vesicles isolation

Blood was collected into lithium-heparin containing vacutainers from patients with (a) unstable angina, myocardial infarction (MI, n=15) and from patients with (b) stable angina (NMI, n=20) prior to CABG and immediately processed. Plasma aliquots were cryopreserved at -80°C before proteomic processing. In order to normalize biological variations, for each group, plasma from individual patients were combined in equal proportions to obtain a total volume of 5 mL. Two biological replicates were performed in this study. Enrichment of plasma extracellular vesicles was performed accordingly to a previously described method (168), with minor modifications. Briefly, pooled plasma samples were diluted to 30 mL with cold 1 x phosphate buffer saline (PBS) and differentially centrifuged at 200 x g (30 min), 2,000 x g (30 min) and 12,000 x g (60 min) to remove any intact cells and cellular debris. Plasma extracellular vesicles were pelleted by ultracentrifugation at 200,000 x g (18 h) using a Beckman L100-XP Ultracentrifuge (Beckman Coulter, Brea, CA). The extracellular vesicles-enriched pellets were washed in 1 x PBS and pelleted again at 200,000 x g (18 h) to remove any contaminants.

3.3.4 Western blotting

Extracellular vesicles proteins (50 µg) were mixed with 4 x laemmli protein sample buffer (BioRad, Hercules, CA), separated on 12% polyacrylamide gel, and transferred to a 0.45 µM nitrocellulose membrane (BioRad, Hercules, CA). Nitrocellulose membranes were blocked and probed overnight at 4°C with the following primary antibodies: anti-CD9 antibody (1:1000; sc-13118; Santa Cruz Biotechnology, Santa Cruz Biotechnology, CA) and anti-CD81 antibody (1:1000; sc-9158; Santa Cruz Biotechnology), followed by incubation with horseradish-peroxidase-conjugated secondary antibody (1:4000, sc-2371, sc-2370; Santa Cruz Biotechnology, CA). Proteins were detected by enhanced chemiluminescence assay (Millipore Coporation, Billerica, MA).

3.3.5 Cryo-electron microscopy (Cryo-EM)

Concentrated extracellular vesicles suspension (4 µL) was applied onto the holey carbon filmed (R2/2 Quantifoil) EM grids at 99% humidity. Excess sample from the grid was blotted with filter paper and rapidly plunged into liquid ethane (Vitrobot, FEI Company, Hillsboro, OR). Cryo grids were examined using a field emission gun transmission electron microscope operated at 80 kV (Arctica, FEI Company, Hillsboro, OR) and equipped with a direct electron detector (Falcon II, FEI Company, Hillsboro, OR).

3.3.6 In-solution tryptic digestion and peptide fractionation

Proteomic sample preparation followed a previously described method that minimizes experimentally-induced deamidation (306, 307), with minor modifications. Briefly, the extracellular vesicles-enriched pellet was solubilized in 8 M urea and 50 mM ammonium acetate (pH 6.0) lysis buffer. Plasma proteins concentration were determined using the bicinchoninic acid (BCA) assay as according to the manufacturer's instructions. Approximately 200 µg of proteins were reduced using 20 mM dithiothreitol (DTT) for 3 h at 30°C, and then alkylated in the dark with 55mM iodoacetamide (IAA) for 1h at room temperature. The concentration of urea was diluted to less than 1 M using 50 mM ammonium acetate buffer (pH 6.0) prior tryptic digestion to ensure adequate activity of trypsin. Proteins were enzymatically digested using sequencing-grade trypsin (Promega, Madison, WI) at a 1:100 (w/w, trypsin to protein) overnight at 37°C. The tryptic peptides were desalted using a Sep-Pak C18 cartridge (Waters, Milford, MA) and dried in a vacuum concentrator. Vacuum-dried desalted peptides were reconstituted in 200 µL mobile phase A (90% ACN containing 0.1% HAc) and fractionated using a PolyWAX LP weak anion-exchange column (4.6 × 200 mm, 5 µm, 300Å; PolyLC, Columbia, MD) on a Prominence UFLC system (Shimadzu, Kyoto, Japan). The UV spectra of the peptides were collected at 280 nm. Mobile phase A and mobile phase B (30% ACN, 0.1% FA) were used to establish the 60 min gradient; starting with 3 min of 100% A, 17 min of 0–8% B, 25 min of 8–45% B, 10 min of 45–100% B, followed by 5

min at 100% B (constant flow rate of 1 mL/min). Forty separate fractions were collected, combined into 15 pooled fractions, and then vacuum-dried. Dried peptides were reconstituted in 3% ACN, 0.1% FA for LC-MS/MS analysis.

3.3.7 LC-MS/MS

The fractionated peptides were separated and analyzed on a LC-MS/MS system equipped with an Ultimate 3000 RSLC nano-HPLC system (Dionex, Amsterdam, NL) coupled to an online LTQ-FT Ultra linear ion trap mass spectrometer (Thermo Scientific Inc., Bremen, Germany). Approximately 2 μg of peptides from each fraction were injected via the Dionex autosampler, concentrated into a Zorbax peptide trap column (Agilent Technologies, Santa Clara, CA), and subsequently separated in a capillary column (75 μm × 10 cm) packed with C18 AQ (5 μm, 300Å; Bruker Michrom Billerica, MA) at a flow rate of 300nL/min. Mobile phase A (0.1% FA in HPLC water) and mobile phase B (0.1% FA in ACN) were used to establish the 60 min gradient; starting with 1 min of 5–8% B, 44 min of 8–32% B, 7 min of 32–55% B, 1 min of 55–90% B and 2 min of 90% B, followed by re-equilibration in 5% B for 5 min. The samples were ionized using an electrospray potential of 1.5 kV in an ADVANCE™ CaptiveSpray™ Source (Bruker-Michrom). The LTQ-FT Ultra was set to perform data acquisition in the positive ion mode. A full MS scan (350–1600 m/z range) was acquired in the FT-ICR cell at a resolution of 100,000 and a maximum ion accumulation time of 1000 ms. The automatic gain control (AGC) target for FT was set at 1×10^6, and the precursor ion charge state screening was activated. The linear ion trap was used to collect peptides and measure the peptide fragments generated by collision induced dissociation (CID). The default AGC setting of full MS target 3.0×10^4, MSn 1×10^4 was used in the linear ion trap. The 10 most intense ions above a 500 count threshold were selected for MS^2 fragmentation in CID, which was performed concurrently with a maximum ion accumulation time of 200 ms. Dynamic exclusion was activated for the process, with exclusion duration of 60 s and repeat count of 1. For CID, the activation Q was set at 0.25, activation time was 30 ms, isolation width (m/z) was 2.0, and normalized collision energy was 35%.

3.3.8 Mass spectrometric data analysis

MS data analysis was performed using Thermo Scientific™ Proteome Discoverer™ (PD) 1.4 software, connected to in-house Mascot server V 2.4.1 (Matrix Science, Boston, MA). The MS/MS spectra were queried against the Uniprot Human database (Released on 11/29/2013, 88, 421 sequences, 35,070,517 residues) for protein identification. Automatic target-decoy search strategy was used in combination with Percolator to score peptide spectral matches for estimation of false discovery rate (FDR). Only peptides identified with strict spectral FDR of less than 1% (q-value < 0.01) were considered in this study. The search was constrained with maximum 2 missed trypsin cleavages; peptide precursor mass tolerances of 10 ppm; and 0.8 Da mass tolerance for fragment ions. Carbamidomethylation (+57.021 Da) of cysteine residues was fixed as static peptide modification, oxidation (+15.995 Da) of methionine residues and deamidation (+0.984 Da) of asparagine and glutamine residues were set as dynamic peptide modification. The event detector and precursor ion area detector algorithm imbedded in PD 1.4 was set at 2 ppm mass precision for calculation of each precursor ion peak area under the curve (AUC). Protein and peptide relative quantities were calculated based on the average area of three most abundant unique peptides per protein instead of all peptides per protein for greater accuracy. Searched results were exported to a tab-delimited file for further analyses. The reported ratio of MI/NMI of each protein was computed using the PD-calculated MI protein area divided by the PD-calculated NMI protein area. A 1.2-fold change cut-off was set and a fold change of 1.2 and 0.83 indicates up-regulated and down-regulated proteins respectively. Relative expression ratios of proteins of interest were extracted and p-values were determined by unpaired Student's t-test. A p-value <0.05 was used to indicate statistical significance.

3.3.9 Luminex assay

Plasma EVs from 26 NMI patients and 17 MI patients were individually isolated using ExoQuick™ (System Biosciences, Mountain View, CA) as per manufacture's protocol, with minor modification. Briefly, 250 µL of plasma was mixed with 63

µL ExoQuick™ solution and incubated overnight at 4°C. The recovered pellet containing EVs were lysed in complete lysis-M buffer (Roche Diagnostics, Mannheim, DE). Plasma EVs C5 and C1Q concentrations were measured using single-plex luminex bead assay (Luminex Corporation, Austin, TX) according to manufacturer instructions. The resulting luminex raw data, reported as fluorescent intensity were acquired with Bio-Plex Manager™ software (BioRad) to obtain concentration value

3.3.10 Bioinformatics analyses and data annotation

FunRich V2.1.2 (221), an open access standalone tool was used for plasma EV functional enrichment and interaction network analyses. Scatter dot plots were generated and statistically analyzed by GraphPad Prism V 6.0 (Graphpad Software, San Diego, CA). Parametric analyses and unpaired student's *t*-test with Welch's correction were used to evaluate the differences between numeric variables for statistical comparisons and $p < 0.05$ was considered to be statistically significant. Variances were tested with F test and differences were considered to be statistically significant when $p < 0.05$. Data are presented as the mean ± standard deviation (SD) of the duplicate experiment.

3.4 Results and Discussion

3.4.1 Human plasma EV label-free proteomics quantitation workflow

Circulatory EVs are highly sought after bioentities given their involvement and relevance in virtually every pathophysiological aspects in humans. We employed a previously described prolong ultracentrifugation (PUC) (168) approach, reported to be effective in simplifying plasma complexity, as part of EVs isolation workflow illustrated in Figure 3.1. The occurrence of co-purified protein aggregates and other contaminants in vesicles preparation isolated using UC approach is inevitable and common (308). For this reason, we term the vesicular preparation in this study as EVs. Biological duplicates were performed in this study. Briefly, harvested plasma EVs of known medical condition (MI and NMI) were enzymatically digested into

peptides with trypsin, followed by first dimensional fractionation using weak anion-exchange chromatography. Finally, fractionated EV peptides were analyzed by reverse phase LC-MS/MS and differential EVs proteomic dataset were generated based on ion intensity/AUC label-free quantification.

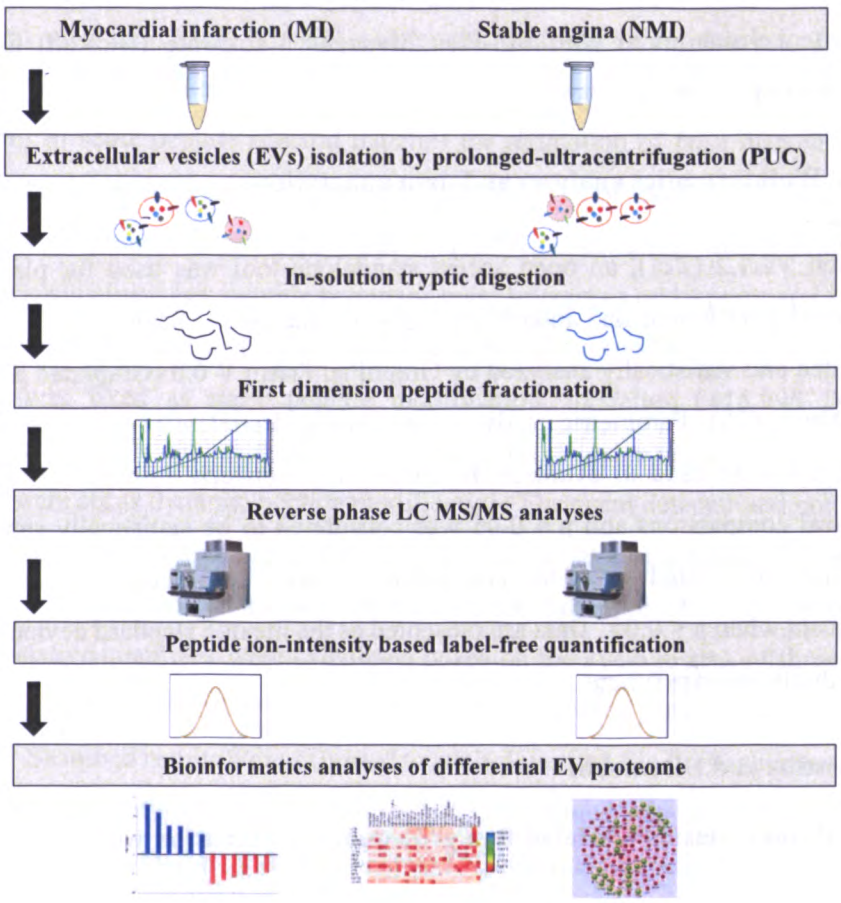

Figure 3.1. Workflow of label-free quantitative proteomics. Schematic workflow of label-free quantitative proteomics workflow of isolated extracellular vesicles (EVs) from pooled plasma obtained from patients with stable angina (NMI) and from patients with myocardial infarction (MI).

3.4.2 Verification of plasma EVs

The presence of EVs in MI and NMI plasma vesicular preparations were demonstrated by widely-used electron microscopy and immunoblotting analyses. Electron micrograph of harvested EVs (Figures 3.2A and 3.2B) showed circular-shaped membrane-surrounded structures with sizes around 50 – 100 nm in diameter. Western blot analyses of harvested EVs detected two commonly reported exosomal

tetraspanin markers CD9 and CD81(309) in the biological replicates of both MI and NMI groups as shown in Figure 3.2C. Additionally, proteomics verification of EV isolation was ascertained with the identification of exosome-specific surface markers including CD9 and ALIX/PDCD6IP. Thus, the successful isolation of EVs using PUC technique was verified and the exhibited morphology, characteristics and proteomics dataset of the harvested EVs appear to be similar to that of exosomes.

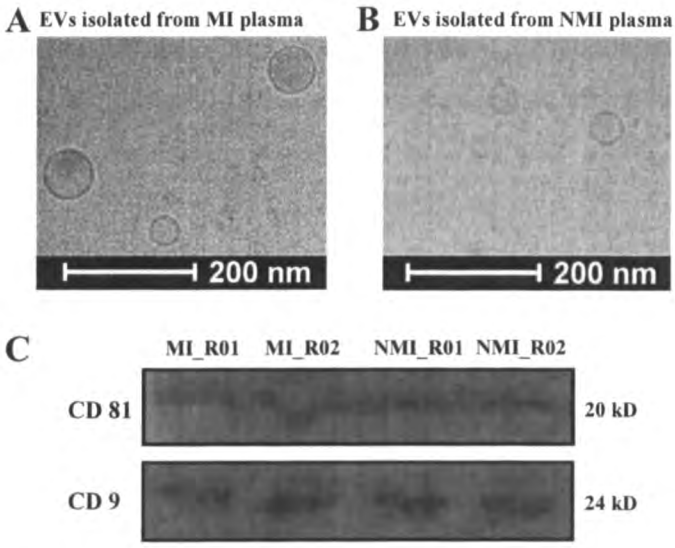

Figure 3.2. Verification of extracellular vesicles (EVs). EVs (sizes 50–100nm) isolated from human plasma. (A) Electron microscopy image of harvested EVs from pooled myocardial infarction (MI) plasma. (B) Electron microscopy image of harvested EVs from pooled stable angina (NMI) plasma. (C) Immunoblot analyses of harvested proteins against EV markers CD9 (24 kD) and CD81 (20 kD). MI_R01 and MI_R02 are biological replicates. NMI_R01 and NMI_R02 are biological replicates.

3.4.3 Plasma EVs protein identification and relative protein abundance

In total, 165223 peptide-spectrum matches (PSMs), corresponding to 4363 unique peptides, and 573 plasma EV proteins were confidently identified with strict target FDR < 1% (q-value < 0.01) across both MI and NMI groups in the combined LC-MS/MS database searched results. Of these, 302 (~53%) proteins are found common between both patient groups, with 422 proteins identified in MI group and 453 proteins identified in NMI group.

Label free peptide ion-intensity/AUC based quantification have been shown to correlate well with protein abundances in complex samples (310). In this study, reproducibility was evaluated using the LC-MS/MS datasets obtained from two independent biological replicates performed in parallel for each patient group. Using a strict FDR < 1%, the number of proteins, peptides and PSMs identification in each biological replicate run were plotted along with calculated mean and SD for repeatability assessment as depicted in Figure 3.3A. The corresponding percentage coefficient variation (%CV) in the number of proteins, peptides and PSMs identification between biological replicates of NMI were 2.35%, 7.13% and 5.95% respectively. While the %CV in the number of proteins peptides and PSMs identification between biological replicates of MI were 0.00%, 6.84% and 6.60% respectively. The low percentage of variability observed in each patient group indicates that high repeatability in protein, peptide and spectrum level was achieved within this study.

For each group, the relative quantitative values of each proteins acquired from each replicate were used for the calculation of Pearson's correlation coefficient (r). The scatter $\log_{10}$ plots of ion intensities of all of matched proteins between independent biological replicates showed significant correlation in both MI ($r = 0.8348$, $p < 0.0001$) and NMI ($r = 0.9096$, $p < 0.0001$) groups (Figure 3.3B and 3.3C), reflecting good reproducibility of the label free quantitative datasets. The use of standardized protocol for sample preparation and instrument accuracy permitted the high repeatability and reproducibility observed in this study. As can be noted from the consistency in the number of identifications and correlation analyses using extracted ion intensities, we have obtained reasonably low inconsistency and good reproducibility across the biological replicates of independent LC-MS/MS runs for both patient groups, indicating high reliability of our relative quantitative dataset.

In order to explore the changes in EV proteome in MI and NMI groups, the average area of three most abundant unique peptides per protein were used for accurate measurement of the relative protein abundance in both MI and NMI group. The differential proteome dataset was generated using the AUC-MI values divided by the

AUC-NMI values (MI/NMI), a 1.2-fold change cut-off was considered as altered expression. As a result, 125 proteins with ratios ≥ 1.2 were classified as up-regulation, whereas 127 proteins with ratios ≤ 0.83 were considered as down-regulation. This finding suggests that many biological activities are modulated following MI.

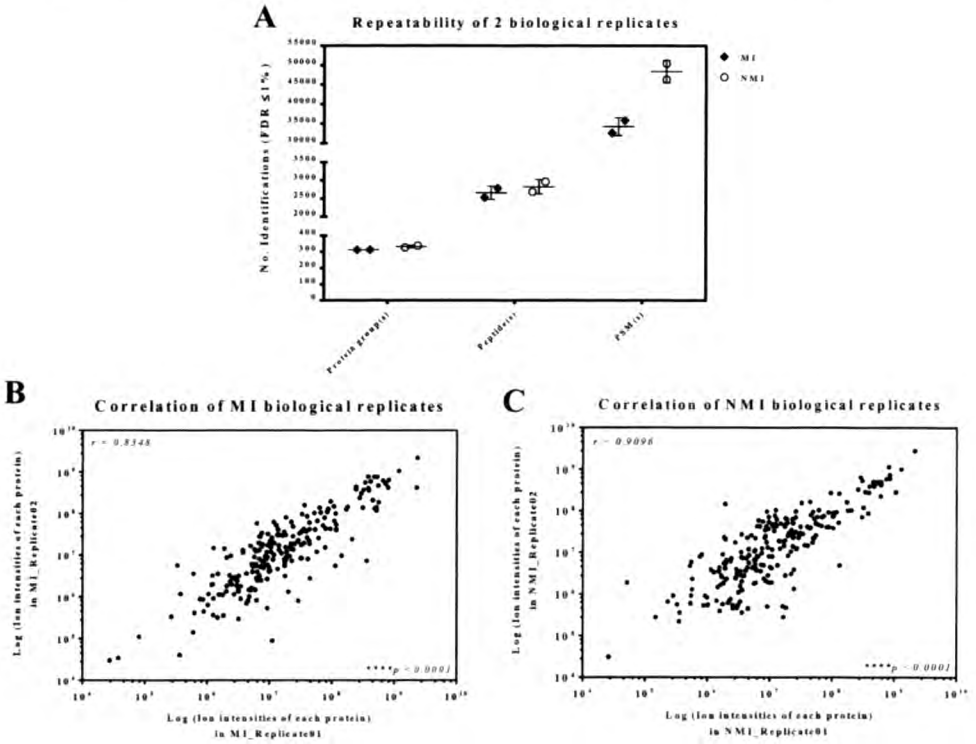

Figure 3.3. Repeatability and reproducibility assessment of replicate experiment. Quality assessment of two independent biological replicates in myocardial infarction (MI) and stable angina NMI). (A) Linear scatter plot of the number of protein group (s), peptide (s) and PSM(s) identified with FDR < 1% in each biological replicates acquired from MI and NMI groups. Dots were plotted with mean and SD. (B) Scatter plot of ion intensities of each protein identified MI replicate 01 plotted against the ion intensities of matched proteins identified in MI replicate 02 in log10 scale. (C) Scatter plot on intensities of each protein identified NMI replicate 01 plotted against the ion intensities of matched proteins identified in NMI replicate 02 in log10 scale. r, pearson correlation coefficient. ****, extremely significant.

3.4.4 Expression profiling of the plasma EV proteome in MI and NMI.

The expression of quantified EV proteins obtained from the combined search results were profiled to gain understanding of the EVs involvement in different clinical severity of CAD, notably after myocardial ischemic injury, which is at present unknown. Gene ontology (GO)-based functional enrichment and network analyses

were performed using all deregulated EV proteins. The up-regulated protein list was compared to the down-regulated protein list for each molecular function (MF), biological processes (BP) and pathway analyses using FunRich V 2.1.2. The significance of the enriched categories was ranked by the Benjamini–Hochberg adjusted p value, $p < 0.05$ indicates high enrichment. Deregulated proteins in GO-annotated categories were then subjected to biological pathway analyses in FunRich V 2.1.2 for meaningful data interpretation (Figure 3.4).

The putative MFs associated with the differentially expressed proteins after MI showed highest enrichment in complement activities (~26%) and transporter activities (~24%), with more up-regulated proteins associating with both annotated categories (Figure 3.4A). Protease inhibitor activities (~11%) and serine-type peptidase activities (~9%) showed higher representation in down-regulated proteins. The enrichment percentage of up-regulated proteins in cytoskeletal protein binding was more distinct than down-regulated proteins, while defense protein activities enrichment was observed in the opposite trend. GO-annotated immune response (~43.8%) and protein metabolism (34.5%) constitute the most prominent categories in BP classification (Figure 3.4B), with down-regulated proteins having higher representation in the latter. Cellular communication (~13.8%) and transport process (~14.9%) were overrepresented in up-regulated proteins. Interestingly, processes involving energy expenditure such as protein metabolism, cell grown and/or maintenance and energy pathways were more pronounced in down-regulated proteins following MI.

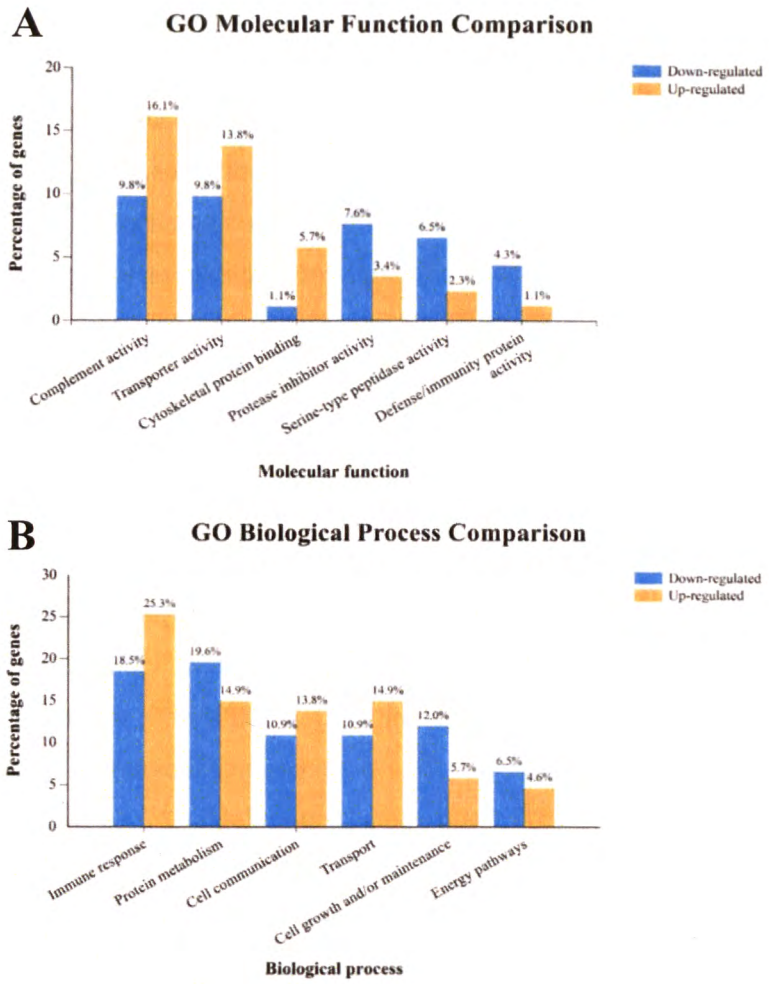

Figure 3.4. Functional enrichment analyses derived from gene ontology (GO). GO functional analyses of all deregulated extracellular vesicle (EV) proteins (A) Distribution of deregulated EV proteins based on their molecular functions. (B) Distribution of deregulated EV proteins based on their biological processes. Predictions were generated from obtained from FunRich V2.1.2.

The most perturbed pathways ($p < 0.05$) enriched in MF-complement activity, MF-serine-type peptidase activity and BP-immune response were complement cascade pathway and innate immune system pathway. EVs produced by both immune and non-immune cells are thought to have important roles in the immunoregulation (311). The complement system is the primary line of host defense and an important element of the innate immune system. The complement cascade can be activated by the classical pathway (CP), the lectin pathway (LP) and the alternative pathway (AP), with all three pathways converging at C3 activation (312). Numerous studies have

documented that complement activation is an early event in post-infarction inflammation and healing (313). Complement activation in atherosclerotic plaques mediates inflammatory damage and is associated with larger infarcts and poor clinical outcome (285). Under physiological conditions, complement cascade activation is controlled by soluble and cell-bound complement regulatory proteins (314). Active inhibitory regulation of complement activation, particularly via the AP was noted with increased levels of AP negative regulators - complement factor H (CFH, ~1.44-fold change) and complement factor I (CFI, ~1.27-fold change) and the decreased levels of AP positive regulators - complement factor B (CFB, ~0.77-fold change) and properdin (CFP, ~0.18-fold change).

In addition, LP proteins mannan-binding lectin (MBL) serine protease 1 and 2 (MASP1, MASP2) were observed to be down-regulated with fold change of ~0.369 and ~0.378 correspondingly in the post-infarct derived EVs, even with decreased levels (~0.67-fold change) of antithrombin-III (SERPINC1), the major inhibitor of MASP2 (315). Further indications of LP suppression following MI were supported by the diminish-regulation of LP activators such as ficolin-2 (~0.36-fold change) and ficolin-3 (~0.52-fold change). Recent studies have shown that the inhibition by of LP using inhibitors of MASP2 or MBL confers a protective role in myocardial ischemic injury (316, 317). Although the role and functions of MASP2 in LP activation are largely investigated, the functional aspects and potential viable option of MASP1 for therapeutic intervention in myocardial ischemia are limited. Thus, the deregulation in the above stated soluble complement regulators of AP and LP, could be an indication of the body's attempt to limit the intensity of activation, implying the involvement of EVs in complement surveillance after ischemic insult. However, the loss of complement inhibition was also observed in our dataset. Deficient levels of negative regulators of the complement system including C4b-binding protein alpha and beta chains (C4BPA and C4BPB) and clusterin (CLU) were noted.

In our differential proteomics dataset, a repertoire of complement components (C1QA, C1QB, C1QC, C1R, C2, C3, C5, C6, C8A, C8B, C8G, C9, CFH and CFI) were observed to be up-regulated in post-infarct derived EVs when compared to non-

infarct derived EVs. Of which, complement C1q subcomponent subunit A (C1QA), the initiator of CP, displayed the highest fold change of ~3.23. C-reactive protein (CRP), an antibody-independent C1q recognition molecule (318) was also observed to be up-regulated (~2.69-fold change) in our differential proteomics dataset. Apart from complement activation, positive correlation and predictive value of serum CRP with adverse outcomes of CAD have been demonstrated (319, 320). The cleavage of complement component C5 (C5) into anaphylatoxin C5a and fragment C5b triggers proinflammatory signaling and assembly of lytic membrane attack complex respectively. In our differential EV dataset, C5, the component leading to downstream terminal complement actions was up-regulated ~1.27-fold change. Increased levels of C5a have shown to contribute to the detrimental effects such as severe inflammation that leads to tissue damage (314). Complement inhibition studies (283-286) in animal models of cardiovascular disease, notably Pexelizumab (anti-C5 antibody), have yielded promising results in controlling unwanted complement activation following MI but results obtained in clinical studies were undesirable. The massive up-regulation of complement proteins highly suggest the pathogenic involvement of complement activation in the development of post-infarction inflammation. Altered levels of CP activation imitator C1QA and downstream terminal complement effector C5, may reflect the unwanted complement activation following myocardial insult.

We have showed the dilemma dual role of EVs in complement activation and suppression after myocardial injury. CP appears to be more intricately involved in the development of post-infarction inflammation, and raised levels of EV-derived complement proteins could be an indication of myocardial insult. Though several complement regulators of AP and LP are in place as demonstrated in our data set, parallel reflection of excessive complement activation overwhelms the beneficial effects of complement regulation. This overlap of erroneous complement activation and regulation is likely a consequence of myocardial insult, driving the development of post-infarct inflammation. The increased levels of EVs-derived C1QA and C5, proteins that mediates the initiation and the terminal effects of complement

activation reflects intense complement activation and inflammation associated with MI, contributing to adverse MI complications and injuries.

Prominent pathways ($p < 0.05$) enriched in MF-transporter activity and BP-transport processes were closely related to metabolism of lipids and lipoproteins pathway and chylomicron-mediated lipid transport. It is well established that high levels of circulatory low density lipoproteins (LDLs) and very low density lipoproteins (vLDLs) and low levels of high density lipoproteins (HDLs) increases one's risk for CAD (321). Depletion of the anti-atherogenic HDLs, with ~0.47-fold change in apolipoprotein A-II (APOA2) and ~0.41-fold change in apolipoprotein A-VI (APOA4) were noted in our differential dataset, an observation which is typical of acute MI (322). As anticipated with the pro-atherogenicity of LDLs in CAD pathology (323), several LDLs including apolipoprotein B (APOB), apolipoprotein C-III (APOC3), apolipoprotein E (APOE) and apolipoprotein(a) (LPA) were up-regulated. Notably, APOC3 exhibited the most pronounced fold change of ~2.63, while APOB, the major proatherogenic protein component displayed ~1.24-fold change. APOC3 inhibits the hydrolysis of triglyceride (TG)-rich particles by lipoprotein lipase and hepatic uptake, hence increasing plasma TG levels (324). Therefore, an overexpression of APOC3 levels is indicative of hypertriglyceridemia, aggravating the pathogenesis of atherosclerosis. Plasma TG levels are associated with risk of CAD, studies have demonstrated that loss of APOC3 functions confers atheroprotection and reduces the risk of clinical event (325).

Paradoxically, apoliporotein D (APOD), a subpopulation of HDL was found to be up-regulated (~1.86-fold change) after MI, as low HDL levels are usually associated with established CAD. APOD is an atypical apolipoprotein that is expressed ubiquitously in mammalian tissues, with higher level of expression in the spleen, testes and brain (326). APOD is a member of the lipocalin, a family of proteins that transports lipids and other small hydrophobic ligands for metabolism (327). Elevated levels of APOD expression have been associated with neurodegenerative diseases, cancer, psychiatric disorders and physiological aging (328). Studies have shown that APOD mediates maintenance and repair within the brain and peripheral nervous

systems (329). Most importantly, fruit flies transfected with human APOD showed increased life expectancy, associated with anti-stress and anti-oxidation activities (330). Despite being a component of HDL, APOD had received considerably less attention in cardiovascular research due to its relative low abundance (327). Consistent with our findings, APOD was also found to be highly enriched in HDL isolated from CAD subjects (331) . Although depositions of APOD have been detected in atherosclerotic plaque, whether APOD contributes to or protects against atherosclerosis remains elusive (327). Thus, increased levels of HDL APOD may act to stabilize the atherosclerotic plaque or a reflection of compensatory response to dyslipidemia after myocardial injury. Greater understanding in the mechanisms of APOD linking CAD may lead to new diagnostic and therapeutic approaches to myocardial injury. We have showed the presence of dyslipidemia in favor of atherogenesis following MI. Imbalance in the metabolism of cholesterol between LDLs and HDLs have been shown to induce endothelial dysfunction and oxidative stress (332), which could potentially drives the development of post-infarct inflammation. The increase levels of APOC3 after MI promotes the development of hyperglyceridemia, a condition that may aggravate the pathogenesis of CAD. The paradoxical discovery of raised levels of EV-derived HDL APOD after MI have raised the possibility of APOD as a post-infarction plaque stabilizer and a potential marker of myocardial injury.

The associated pathways ($p < 0.05$) mapped from MF-protease inhibitor activity, MF- cytoskeletal protein binding, BP-protein metabolism, BP-cell communication and BP-cell growth/maintenance were linked to platelet degranulation pathway, hemostasis pathway and clotting cascade. Apart from playing a role in hemostasis and thrombosis, the proinflammatory action of platelets through release of inflammatory products and formation of platelet-leukocyte complexes in atherosclerosis are well documented (333). It has been shown that platelets were amongst the first population of inflammatory cells to accumulated within the infarcted myocardium alongside with infiltrated leukocytes (334). The up-regulation (~1.41-fold change) of predominant platelet integrin alpha-IIb (ITGA2B) along with its interaction partners such as von Willebrand factor (VWF) having ~3.29-fold

change was noted. Our data also revealed an exaggerated levels VWF receptor complex components such as platelet glycoprotein Ib alpha chain (GP1BA) having a the most intense fold change of ~9.18, platelet glycoprotein Ib beta chain (GP1BB) with fold change of ~1.31and platelet glycoprotein IX (GP9) exhibiting ~1.22-fold change. The interaction of the above mention players results in platelet adhesion and aggregation (335), which are indispensable wound repair mechanisms that leads to hemostasis especially in vascular injury. However, abnormal platelet activation may result in dire consequences such as arterial thrombosis and restenosis that cause myocardial ischemia and MI. GP1BA is a major platelet receptor involved platelet adhesion and activation, processes that contributes to platelet thrombus formation (336). At the site of plaque rupture, the binding of GP1BA with VWF results in initial platelet rolling contact along exposed subendothelium collagen (336). The increased circulatory EV levels of GP1BA is a potential indicator for plaque rupture and thrombosis, increasing the risk of clinical event.

Soluble chemokines released upon platelet degranulation to trigger the recruitment of circulating neutrophils and monocytes for host defense or immune surveillance (333), were observed in high levels post-MI, with platelet factor-4 (PF4) displaying a fold change of ~2.36 and ~4.72-fold change was seen in platelet basic protein (PPBP). PPBP, possess neutrophils recruitment and activating properties (337). The up-regulation of PPBP could be potentially important in immune activation and thus a potential marker for post-MI vascular inflammation. In our differential EV data set, all components of fibrinogen were down-regulated unexpectedly, fibrinogen alpha chain (FGA) presented with the most pronounce down-regulation with ~ 0.28-fold change. Fibrinogen, an acute-phase protein are usually in raised levels in inflammatory and hypercoaguable states such as stroke and MI (338). To the best of our knowledge, there is no literature that report on diminished levels of fibrinogen in MI-related research. We reason that the reduction levels of fibrinogen post-MI could be a form of EV-mediated compensatory response to hypercoaguability after myocardial injury.

The pathway enriched in MF-defense/immunity protein activity and BP-energy pathways were not discussed here as they were not statistically significant ($p > 0.05$). CAD is a multifactorial disease; as such the identification of a panel of multimarkers relevant to biological pathways that promote and recapitulate disease progression would better increase the predictive value than the reliance on a single biomarker. Preliminary categorization of patients into MI and NMI group is expected to increase the likelihood of detecting stage-specific biomarkers that delineate atherosclerotic progression and plaque instability, hence achieving higher accuracy in patient risk stratification and to permit the development of timely therapeutic interventions of major cardiovascular complications. Capitalizing on known knowledge in MI pathophysiology, we have identified a panel of six promising EV markers of myocardial injury (Table 3.1) that reflects post-infarct pathways involvement in complement activation complement activation (C1Q1A, ~3.23 fold-change, $p = 0.012$; C5, ~1.27 fold-change, $p = 0.087$), lipoprotein metabolism (APOD, ~1.86-fold change, $p = 0.033$; APOC3, ~2.63-fold change, $p = 0.029$) and platelet activation (GP1BA, ~9.18-fold change, $p < 0.0001$; PPBP, ~4.72-fold change, $p = 0.027$). Increased levels of EV-derived C1Q1A, C5, APOC3, GP1BA and PPBP have been suggested to be involved in the aggravation of post-infarction inflammation and myocardial tissue damage. Whereas, elevated levels of APOD is likely to confer atheroprotective properties and plaque stability. The functional comparative profiling of perturbed EV proteome of MI with NMI was presented for the first time, to the best of our knowledge. Our findings suggest the association of EVs in the progression of myocardial damage and promotion of myocardial healing, with the first as main involvement.

Table 3.1. Candidate biomarkers of myocardial injury with their associated function and clinical relevance.

Accession No.	Protein description	Proteomics fold change	P value	Mean AUC (MI)	Mean AUC (NMI)	SE of difference	Clinical relevance
P02656	Apolipoprotein C-III	2.633	0.029	1.013E+08	4.037E+07	1.819E+07	Promotes development of hyperglyceridemia, atherogenic.
P05090	Apolipoprotein D	1.860	0.033	6.877E+06	3.397E+06	1.088E+06	Compensatory response to dyslipidemia after myocardial injury.
P02745	Complement C1q subcomponent subunit A	3.229	0.012	1.710E+07	5.457E+06	2.637E+06	Promotes of post-infarct inflammation, aggravate myocaradial injury
P01031	Complement C5	1.269	0.087	2.307E+07	1.713E+07	2.625E+06	Promotes of post-infarct inflammation, aggravate myocaradial injury
P07359	Platelet glycoprotein Ib alpha chain	9.181	< 0.0001	6.050E+06	5.587E+05	1.003E+05	Possible ndicator of plaque rupture and thrombosis
P02775	Platelet basic protein	4.725	0.027	5.680E+06	1.460E+06	1.240E+06	Immune activation,p ossible indicator of vascular inflammation

AUC; Area under curve; CAD, Coronary artery disease; LDL, Low-density lipoprotein, HDL, High-density lipoprotein; SE: Standard error; MI, Myocardial infarction; NMI, Stable angina. Parametric student's t test analyses was performed using ion intensities obtained from MI_biological replicate 01, MI_biological replicate 02, MI_combined replicates search, NMI_biological replicate 01, NMI_biological replicate 02 and NMI_combined replicates search. $p < 0.05$ indicates statistically significant.

3.4.5 Validation of qualitative differential expression

In effort to cross-validate the expression levels of EVs obtained from PUC, ExoQuick™ (System Biosciences, Mountain View, CA), a method for EV precipitation (339, 340) was used to isolate plasma EVs from 17 MI patients and 26 NMI patients. In order to ensure the accuracy of label-free quantitation, established Luminex single plex assay was used to validate the expression levels of two up-regulated complement proteins in 43 individual patient plasma. The proteins of interest, includes C1QA and C5 with exhibited proteomics fold change of ~3.23 and ~1.27 respectively were selected for validation based on published effect of complement proteins in post-infarction inflammation and reparative response.

Chapter 3

Overactive complement proteins that mediates the initiation (C1QA) and the terminal effects of complement activation (C5) serve as parallel indications of intense complement activation which plays important role in the pathogenesis of ischemic injury and remodelling. The median expression data were presented as linear box-and-whiskers plot (Figure 3.5). In agreement with the proteomics relative quantitation, the plasma EV C1QA level in post-MI patients was significantly higher than of NMI patients ($2.210E+^{07} \pm 1.246E^{+07}$ pg/ml in MI versus $1.189E^{+07} \pm 6.256E^{+06}$ pg/ml, p = 0.005 in NMI). Similarly, plasma EV C5 levels was evidently elevated in MI patients relative to NMI patients in Figure 4B ($1.702e+^{05} \pm 1.154e^{+05}$ pg/ml in MI versus $6.994e^{+04} \pm 8.498e^{+04}$ pg/ml, p = 0.0021 in NMI), in accordance with label-free quantified expression. Collectively, the validation by Luminex assay confirmed that EV proteins C1QA and C5 expression levels were altered following MI, with expressions levels up-regulated in MI patients, thus confirmed the confidence of the reported label-free quantitative proteomics data. In addition, the generalization of the independent quantitative dataset obtained from an alternate EV isolation method into our relative EV proteomics dataset obtained from PUC further strengthen the results of our label-free expression profiling approach.

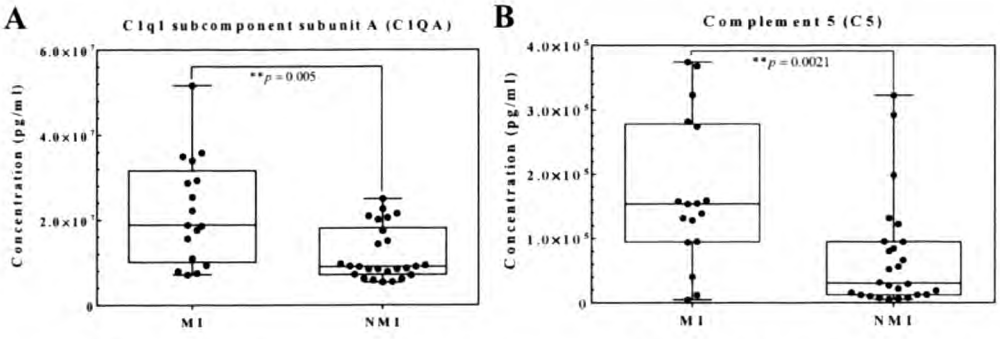

Figure 3.5. Luminex validated expression of plasma extracellular vesicular (EV). Expression levels of EV proteins in 43 human plasma samples (17 patients who have had myocardial infarction [MI] and 26 stable angina [NMI] patients) were shown as linear box-and-whiskers plots. Line represents median, box represents 25th to 75th percentiles and whisker represent maximum and minimum ranges and dot represent individual values (A) Significant (p = 0.005) differential expression data of complement C1q1 subcomponent subunit A (C1QA) measured by Luminex assay between MI and NMI patient groups. (B) Significant (p = 0.0021) differential expression data of complement 5 (C5) measured by Luminex assay between MI and NMI patient groups. **, very significant.

3.5 Conclusion

We have performed a global ion intensity label-free differential proteome analyses of plasma EVs obtained from MI patients and from NMI patient as control, this categorization will facilitate the detection of myocardial injury specific markers. We presented, for the first time, the comparative proteomics profiling on 252 perturbed EV proteins with 1.2-fold change cut-off in clinical samples, and showed the possible involvement of EVs mediation in myocardial damage and myocardial healing processes following MI. A panel of six potential EV markers of myocardial injury reflecting post-infarction pathways of complement activation (C1Q1A, C5), lipoprotein metabolism (APOD, APOC3) and platelet activation (GP1BA, PPBP) was proposed. C1QA and C5, the respective important key initiator and terminal effector of complement activation in MI response, were selected for validation in another cohort of 43 individual patients using Luminex assay, and results obtained were aligned with our quantitative proteomics dataset. Future larger studies will be important for confirmation of our data for enhancement of knowledge on the pathophysiological roles untaken by EVs following MI. Our findings have provided new insights into EV involvement after MI, the knowledge gained should encourage further studies that drives development of efficacious EV-based diagnostic and therapeutic strategies, benefiting patients with MI, improving their survival and quality of life.

Chapter 4

Investigation into Systemic Protein Carbamylation in the Absence of Uremia Reveals Novel Site-Specific Carbamylated Peptide Markers Predictive of Myocardial Infarction.

Chapter 4

4.1 Abstract

Protein carbamylation accompanied with loss of functionality in proteins, result in inappropriate cellular response and deleterious effects. Current knowledge in systemic carbamylation associated with coronary artery disease (CAD), under normal renal function, remain sparse and poorly-defined. Herein, we report the first differential relative quantitative assessment on characterized carbamylated plasma peptides/proteins that likely reflects in vivo inflammatory driven carbamylation in plasma obtained from non-uremic patients who have had myocardial infarction (MI) and from apparently healthy subjects. The reliability of our quantitative dataset was demonstrated by the low analytical variability (CV < 7%) and good protein complementation among the triplicate LC-MS/MS experiment. The occurrence of total plasma protein carbamylation in the absence of uremia was approximately 4% each in both MI and Ctrl groups. A total of 30 high confident (FDR ≤ 1%, $p < 0.05$) carbamylated peptides, assigned to 15 proteins were significantly elevated (≤ 1.5 ratio-fold change, $p < 0.05$) in the MI group, likely a consequence of inflammation-associated atherosclerosis. Among which, 7 novel carbamylated proteins (c-HP, c-TF, c-ORM1, c-SERPING1, c-C3, c-A2M and c-ITIH4) were identified. Of substantial interest, the prominent up-regulation of c-HP (K-LPECEADDGC PK*PPEIAHGYVEH SVR-Y, ~ 6.43-fold change, $p = 0.00011$), c-TF (R-SMGGK*EDLIWELLNQ*AQEH FGKDK-S, ~ 7.38-fold change, $p = 0.00033$) and c-ORM1 (K-TYM*L AFDVNDEK*N *WGLSVYAD KPETTK-E, ~ 4.46-fold change, $p < 0.0001$) have made them the best candidate inflammatory peptide markers of MI. As a consequence of carbamylation, we speculate that the loss of immunomodulatory function in HP, TF, and ORM1 favours the pathogenesis of atherosclerosis. The new findings in this study would direct future molecular investigations aimed at unravelling the roles and functions of atherosclerosis relevant carbamylated peptides of c-HP, c-ORM1 and c-TF. The diagnostic values and possible utility of these novel carbamylated peptides in the treatment and management of atherosclerotic CAD, warrants further examinations in a valid outcome study.

Chapter 4

4.2 Introduction

Systemic atherosclerosis is an inevitable age-dependent pathology (341). Alongside with global longevity boom, the prevalence of coronary artery disease (CAD) will continue to rise, escalating cardiovascular (CV)-related morbidity and mortality (342). It has been proposed that age-associated decline in CV health comes about, in part, from the gradual accumulation of non-enzymatic post-translational modification (PTM) of long-lived proteins (343). The deleterious side effects of carbamylation in humans was first demonstrated by the oral administration of sodium cyanate for treatment of sickle cell diseases in the 1970s (344).

Non-enzymatic protein carbamylation is one of the earliest PTM elucidated, primarily reported as a chemical product generated by the spontaneous reaction of urea-derived cyanate and protein (179). Isocyanic acid, the reactive form of cyanate, can be generated either by spontaneous urea decomposition under physiological condition or by myeloperoxidase (MPO)-catalyzed oxidation of thiocyanate (derived from dietary intake and smoking) during inflammation (345, 346). Isocyanic acid preferentially reacts irreversibly with the lower pK_a Ɛ-amino group of lysine residue, forming homocitrulline, a marker used to assess carbamylation in biological systems (182, 347). Cyanate are usually present in low concentrations in plasma and elevated levels accompanied with protein carbamylation has been evidently implicated in chronic inflammatory conditions including chronic kidney disease (CKD) (180, 181), rheumatoid arthritis (RA) (348, 349) and atherosclerosis (187, 350). Owing to the abundance of circulating proteins (351) and long half-lives of extracellular matrix (ECM) proteins (180), they are common targets for carbamylation.

Chronic inflammation, has been established as a key player in all stages of atherosclerosis. MPO, a known risk factor of CAD, is a heme protein expressed predominantly by inflammatory cells including neutrophils, monocytes and certain tissue macrophages, such as those found in human atheroma (352). Recent findings showed that carbamylated low density lipoprotein (c-LDL) confers atherogenic

activities including increased LDL binding to macrophage scavenger receptor promoting foam-cell formation, enhanced endothelial cell death and smooth muscle cell proliferation (182, 187). Furthermore, functional assessment of carbamylated high density lipoprotein (c-HDL) demonstrated pro-inflammatory and pro-oxidative activities (353). MPO catalyzed carbamylated proteins therefore potentially constitute a molecular substratum of atherosclerosis, offering a plethora of ideal diagnostics biomarkers predictive for atherosclerosis. The extent of protein carbamylation in plasma can be accessed either by quantitating a specific carbamylated protein (e.g. c-LDL) or by the overall assessment of carbamylated plasma protein via colormetric assay (189). The measurement of protein-bound homocitruline using high performance liquid chromatography (HPLC) coupled with tandem mass spectrometry (MS/MS) have been recognized for its high analytical specificity and sensitivity for the assessment of a wide spectrum of carbamylated proteins (182). The specificity and potential diagnostic values of PTM biomarkers ascribed to specific disease can be hampered by the concomitant association of PTM in various disease pathogenesis, and the co-occurrence of different unanticipated PTM found within the same peptide. For these reasons, the use of well structurally defined PTM such as specific modified peptides is therefore advisable (345). MS-based peptide profiling in biological fluids have shown diagnostic promise in cancer and cardiovascular research (354-357). In contrast to the growing research that supports the pro-atherogenic effects of dysfunctional carbamylated lipoproteins, the current understanding in the functional molecular implication of systemic protein dysfunction due to carbamylation and its potential as molecular substratum of atherosclerosis has been elusive.

In this study, we hypothesized that high abundance circulatory carbamylated proteins are progressively accumulated and magnified in the course of chronic atherosclerosis, elevated levels of site-specific carbamylation contributes to the specific pathogenesis of atherosclerosis and therefore predicts for high risk patients. Using label-free quantitative proteomics approach, we assessed the differential carbamylated plasma proteome of patients who have had myocardial infarction (MI) and from control subjects (Ctrl) with no prior history of angina. Under normal renal function,

approximately 4% of total plasma proteome were found to be carbamylated. A total of 30 high confident (FDR ≤ 1%, $p < 0.05$) carbamylated peptides assigned to 15 proteins were found to be significantly elevated ($p < 0.05$) in MI group, with ≥ 1.5-ratio fold change. Of the 7 novel carbamylated proteins identified, carbamylated peptides of c-HP (K-LPECEADDGCPK*PPEIAHGYVEHSVR-Y, ~ 6.43-fold change, $p = 0.00011$), c-TF (R-SMGGK*EDLIW ELLNQ*AQEHFGKDK-S, ~ 7.38-fold change, $p = 0.00033$) and c-ORM1 (K-TYM*LAFDVND EK*N*WGLSVYADKPETTK-E, ~ 4.46-fold change, $p < 0.0001$) exhibited prominent up-regulation, hence represent the most promising candidate inflammatory peptide markers of MI. We speculated that the loss of HP, TF and ORM1 function as a consequence of carbamylation would aggravating inflammation associated atherosclerotic conditions.

4.3 Materials and Methods

4.3.1 Chemicals

All chemicals were purchased from Sigma-Aldrich (St Louis, MO) unless stated otherwise. All water and acetonitrile (ACN) used in this experiment were of HPLC grade (Thermo Scientific, Waltham, MA).

4.3.2 Human plasma sampling

Written informed consent of participation Written informed consent of participation was obtained from all study participants. The study was approved by the National Healthcare Group Domain Specific Review Board (NHG-DSRB). Recruited patients were categorized according to their severity of CAD. Thirteen patients presenting with have unstable angina pectoris and have had myocardial infarction (< 5 days) were grouped under MI. Blood obtained from MI patients (n=13) undergoing coronary artery bypass graft (CABG) surgery were immediately processed for plasma. Control plasma was obtained from 23 healthy individuals with no prior history of angina. All plasma aliquots were cryopreserved at -80°C until proteomics processing.

4.3.3 Plasma protein extraction

For each group, non-depleted plasma from individual patients were equivalently pooled to a final volume of 200 µL to normalize biological variation. Plasma proteins were recovered by centrifugation at 16, 000 x g (10 min), and washed in pre-chilled 80% acetone twice to remove any residual contaminants, if any. Protein contents were quantified in a 96-well plate using bicinchoninic acid (BCA) assay according to the manufacturer's protocol. 200 µg of proteins from each group were used for subsequent proteomics processing.

4.3.4 In-gel tryptic digestion

Extracted proteins (200 µg) were reconstituted in 4 x laemmli protein sample buffer (BioRad, Hercules, CA) supplemented with 50 mM dithiothreitol (DTT) and heated at 95°C (5 min). The denatured protein mixture were then resolved on a 12% polyacrylamide gel by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Twelve coomassie brillant blue G-250 stained protein bands were excised, diced and de-stained completely in 25mM ammonium bicarbonate (ABB). Proteins were reduced with 10mM DTT for 1h at 56°C, and were then alkylated in the dark with 55mM iodoacetamide (IAA) at room temperature (RT). Proteins were enzymatically digested overnight at 37°C using sequencing-grade trypsin (Promega, Madison, *WI)* at a 1:100 ratio (w/w, trypsin:protein). Tryptic peptides were extracted with buffer containing 50% ACN, 5% acetic acid (HAc) and extracted peptides were dried in a vacuum concentrator. The tryptic peptides were desalted using a Sep-Pak C18 cartridge (Waters, Milford, MA*)* according to the manufacturer's protocol and the eluted peptides were dried in a vacuum concentrator. Dried peptides were then reconstituted in 3% ACN, 0.1% formic acid (FA) for subsequent LC-MS/MS analysis

4.3.5 LC-MS/MS

Label-free dried peptides were injected and analyzed in a LC-MS/MS system comprised of a Dionex Ultimate 3000 RSLC nano-HPLC system (Dionex, Amsterdam, NL), coupled to an online Q-Exactive™ (QE) hybrid quadrupole orbitrap mass spectrometer (Thermo Scientific, Hudson, NH). Approximately 1 µg of dried peptides in each fraction were injected onto a Acclaim® PepMap100 trap column (75 µm × 2 cm) packed with C18 (3 µm, 100 Å) (Thermo Scientific, Hudson, NH, USA) via the auto-sampler of the Dionex system with a flow rate of 1 µL/min. The peptides in each injection were resolved in an EASY-spray column (75 µm × 15 cm) packed with C18 (3 µm, 100 Å; Bruker-Michrom, Auburn, CA) at a flow rate of 300 nl/min. Mobile phase A (0.1% FA) and B (90% ACN,0.1% FA) were used to established the 60-min gradient; starting with 1 min of 3 - 6% B,46 min of 6 -35 % B, 6 min of 35 - 60% B, 1 min of 60 - 70 % B and 1 min of 70% B followed by re-equilibration at 3% B for 5 min. The samples were ionized in an EASY-Spray™ Source (Thermo Scientific, Hudson, NH) with a nano- electrospray potential of 1.5 kV. The QE MS instrument was set to perform data acquisition in the positive ion mode. A full MS scan (350–1600 m/z range) was acquired at a resolution of 70,000 in the orbitrap when the maximum automatic gain control (AGC) target of 3×10^6 or a maximum ion accumulation time of 100 ms was reached. The 10 most intense peptide ions with dynamic exclusion duration of 30s and charge states of 2 -5 were sequentially fragmented in HCD collision cell and isolated to a maximum AGC target of 2×10^5 or a maximum ion accumulation time of 120 ms, MS/MS scan was acquired at a resolution of 35,000 in the orbitrap. For HCD, the normalized collision energy (NCE) was 28%, isolation width (m/z) was 2.0 and under fill ratio was defined as 0.5%. Triplicate LC-MS/MS runs per fraction were performed.

4.3.6 Data analyses

Standard de-isotoping method as found in Thermo Scientific Proteome Discoverer™ (PD) 1.4.1.14 software was used to extract and convert de-istope tandem MS spectra from the raw data of QE™ into Mascot Generic Format (MGF) files. Protein identification was performed by querying against the extracted Uniprot Human database (Released on 20131129; 176946 sequences, 70141034 residues), by means of an in-house Mascot server (version 2.4.1, Matrix Science, Boston, MA). Target-decoy search strategy was employed for the estimation of false positive identification. The search was limited to a maximum of 2 missed trypsin cleavages; # 13 C of 2; peptide precursor mass tolerances of 5.1 ppm; and 0.02Da mass tolerance fragment ions. Carbamidomethylation (+57.021 Da) of cysteine residues was set as a fixed modification, while oxidation (+15.995 Da) of methionine residues, deamidation (+0.984 Da) of asparagine and glutamine residues, acetylation of lysine residues (+42.0106) and carbamylation (+43.0058 Da) of lysine residues were set as variable modifications. Mascot results were exported to csv file format for further processing in Microsoft Excel. Peptides identified were sorted from the smallest to largest Mascot peptide expect value, which is a measure of random match probability. Peptide false discovery rate [FDR = 2*(decoy hits/total hits)*100%] was generated using an in-house written program and the cut-off used in this study was set at FDR $\leq$ 1. Mass precision of 3 ppm was set for the calculation of each precursor ion peak area under the curve (AUC) from the extracted ion chromatogram (XIC). The XIC of all each carbamylated peptide was manually verified for accurate peak integration. Carbamylated peptide relative quantities were calculated based on the average area obtained from the triplicate experiment. The relative expression ratios of carbamylated peptides were calculated using the peptide area of MI divided by the peptide area of Ctrl (MI:Ctrl), and p-values were determined by unpaired Student's t-test. A p-value < 0.05 was used to indicate statistical significance. In to minimize outliers in our dataset, only peptides with individual ion scores above homology or identity scores ($p < 0.05$) were considered.

4.3.7 Bioinformatics analyses and data annotation

Scatter plots were generated and statistically analyzed by GraphPad Prism V 6.0 (Graphpad Software, San Diego, CA). Parametric analyses and unpaired student's *t*-test were used to evaluate the differences between numeric variables for statistical comparisons and $p < 0.05$ was considered significant unless specified otherwise. Data are presented as the mean ± standard deviation (SD) of the triplicate experiment.

4.4 Results and Discussion

4.4.1 Human plasma urea-free proteomics workflow

The accurate determination of global carbamylated profile of human plasma, proteomics sample preparation is a major factor that needs to be well-thought-out. Urea, a chaotropic agent, is frequently used to solubilize proteins as part of proteomics sample processing workflow. In aqueous solution, the decomposition of urea to isocyanic acid is facilitated by high temperature (≥ 37°C) and alkaline conditions (pH 7-9), isocyanic acid in excess reacts with side chains of lysine and arginine to form carbamylated residues, mimicking in vivo carbamylation (358). Studies have shown that artificial carbamylation of proteins not occur during electrophoresis in the presence of urea, yet can be induced by sample preparation (359). In efforts to minimize artifactual carbamylation, the use of urea in sample protein preparation in this study was avoided. Briefly, as depicted in Figure 4.1, acetone precipitated plasma proteins were solubilize in SDS containing lysis buffer and fractionated by SDS-PAGE. In the absence of urea, gel-trapped proteins were reduced, alkylated and subsequently enzymatically digested into peptides with trypsin. The tryptic peptides were desalted and eluates were dried under vacuum. Finally, dried peptides were analyzed in triplicates by LC-MS/MS and data generated were analyzed using bioinformatics tools.

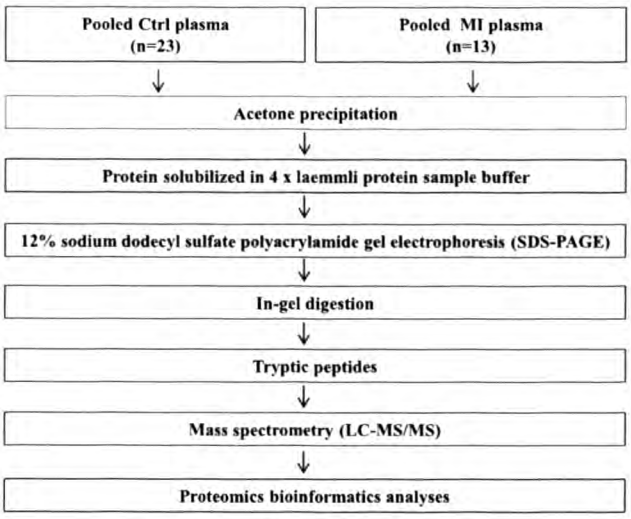

Figure 4.1. Work flow of in-gel digestion for carbamylated proteomics. Urea-free in-gel digestion proteomics workflow of pooled plasma obtained from control subjects (Ctrl) and patients who have had myocardial infarction (MI).

4.4.2 Quality assessment of LC-MS/MS data sets

In order to ensure that findings from this study are meaningful, variability resulting from analytical system were assessed using datasets (FDR ≤ 1%) obtained from the triplicates LC-MS/MS runs for each sample group. As displayed in Figure 4.2A, the datasets were plotted along the mean ± SD of triplicate runs, and high repeatability in proteins, peptides and spectrum level were observed. The calculated %CV in the repeated identification of proteins, peptides and PSMs between Ctrl triplicates were 4.31%, 2.45% and 2.82% respectively. While corresponding %CV of 6.41%, 0.99% and 2.70% were calculated using the number of proteins peptides and PSMs identified from MI triplicates. Low variability was demonstrated in our dataset with an overall %CV of < 7% for both patient groups. Furthermore, an overlapped of 84% (524 proteins) proteins identification were shared between at least two Ctrl replicates (Figure 4.2B), and nearly 73% (476 proteins) of proteins identification were common across all triplicates. Similarly in Figure 4.2C, out of 476 proteins identified from MI triplicate analysis, approximately 86% (408 proteins) were detected in at least two MI replicates, with 368 (~77%) proteins common to all triplicates. In general, both Venn diagrams reflects good protein complementation between the three LC-

MS/MS runs. The evaluation of repeatability and variability with each sample dataset showed high consistency and good reproducibility across the triplicate analyses. In this study, good stability of the analytical system have resulted in low between runs variations, thus contributed to the reliable and confident proteomics dataset for carbamylation studies. The global plasma proteome of Ctrl and MI groups were obtained from the collective mascot search result of the triplicate LC-MS/MS experiments.

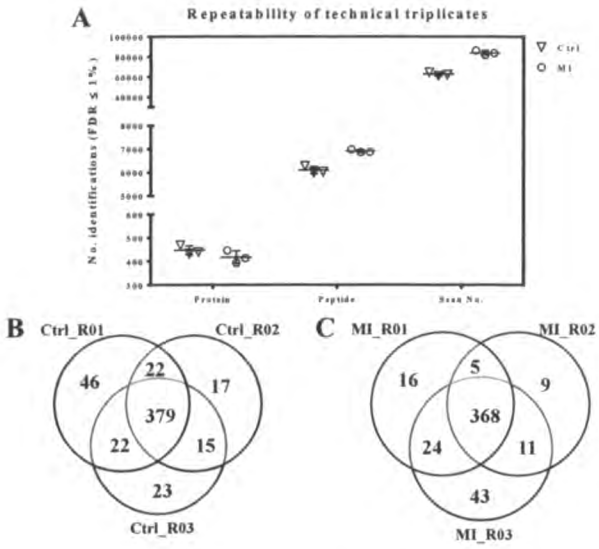

Figure 4.2. Data quality assessment of triplicate LC-MS/MS runs. Repeatability and reproducibility assessment of the technical triplicate LC-MS/MS runs (A) Linear scatter plot of (FDR ≤ 1%) proteins, peptides and spectral matches identified (FDR ≤ 1%) in each technical triplicates acquired from MI and Ctrl groups. Data were plotted along with calculated mean and standard deviation (SD) for repeatability assessment. (B) Venn diagrams depicts the distribution of identified proteins (FDR ≤ 1%) among technical triplicates of Ctrl group. (C) Venn diagrams depicts the distribution of identified proteins (FDR ≤1%) among technical triplicates of MI group. Ctrl, control; MI, myocardial infarction; R01, replicate 1; R02, replicate 2; R03, replicate 3.

4.4.3 Carbamylated plasma protein detection

At peptide FDR ≤ 1%, a total of 7508 unique peptides assigned to 473 proteins were identified in the Ctrl group, whereas 8446 unique peptides corresponding to 437 proteins were confidently detected in MI group. Lysine carbamylation (+43.00582 Da) and lysine acetylation (+42.0106 Da) have a mass difference of only 0.995 Da, this mass shift unfortunately cannot be differentiated in MS/MS (360), resulting in

the ambiguous assignment and identification of carbamylated and acetylated peptides. In attempt to improve the accuracy of monoisotopic peak assignment of precursor ion masses, prior to database search, all MS/MS spectra data were deisotope to eliminate computational miss-assignment of C13 isotopic peak of the acetylated peptide as carbamylated peptide and vice versa (361, 362). Furthermore, database search was carried out using a narrow precursor ion mass tolerance window of 5 ppm and a strict peptide FDR $\leq$ 1% was implemented to the proteomics dataset before analyses to facilitate high fidelity peptide identification (363, 364).

In addition, only peptides hits with mascot peptide ion score > 41 and/or mascot expect value < 0.05, which indicates identity or extensive homology ($p < 0.05$) were considered for further analyses. The fragmentation MS/MS spectra of all unique carbamylated plasma peptides from Ctrl and MI groups were manually evaluated. Following these criteria, 48 unique carbamylated peptides and 35 carbamylation sites, across 21 proteins were detected in Ctrl plasma, while 55 unique peptides with 41 carbamylation sites, corresponding to 20 proteins were confidently assigned as carbamylated in MI plasma group. Overall, carbamylated plasma proteins have accounted to approximately 4% of total plasma proteome in each Ctrl and MI group. We detect no obvious difference in the occurrence of total plasma protein carbamylation between MI and Ctrl groups. The under-explored role of carbamylated plasma proteins solely in CADs, is in part due to the extremely low production of carbamylated-derived products under basal conditions, where plasma urea concentrations are normal (182). To our knowledge, this is the first description of the unbiased estimation of total in vivo plasma protein carbamylation, in apparent healthy individuals and patients who have had MI and in the absence of uremia.

4.4 Label-free quantitative comparison of carbamylated peptides

Current knowledge in systemic differential proteomics on circulatory carbamylation in CADs, under normal renal function, remain sparse. In attempt to explore the differential changes of carbamylation in MI, the relative quantities of all 79 unique carbamylated peptides obtained from both MI and Ctrl group were assessed by measuring the integrated peak areas from the XIC of each peptide. Of note, all carbamylated peptides reported in this study were incompletely digested, as lysine carbamylation changes the protein charge state, causing protein conformational alterations, hindering proteolytic digestion by trypsin (365). Thereby, the miss cleaved peptide containing additional basic amino acids (K) might be protonated multiple times by nanospray, resulting in the detection of the same peptide with different charge states (e.g. z = 3+, 4+, 5+). In this study, the different charge species and/or the co-occurrence of different PTM found on the same carbamylated peptide, if any, were treated as separate entries and included for investigation of consistency in quantitative results.

The differential expression ratio of carbamylated peptides in MI and Ctrl (MI:Ctrl), were established by dividing the MI mean peak area values with the Ctrl mean peak area values obtained from the triplicate experiments. Label-free quantitative comparison of MI and Ctrl carbamylated peptides revealed a total of 47 significantly ($p < 0.05$) regulated peptides. In order to increase the reliability of the quantitative dataset, only carbamylated peptides with mean peak area CV < 15% and a 1.5-ratio fold change were considered as altered expression. Based on these filtering conditions, we report a final list of 30 unique carbamylated peptides assigned to 15 high abundant plasma proteins that showed significant up-regulated expression in the MI group (Table 4.1). Interestingly, all carbamylated peptides were observed to be magnified in MI when compared with Ctrl. Consistent trend in regulation were observed on same carbamylated peptides that contain different co-occurring PTMs (.e.g. deamidation, oxidation), indicating that the changes detected here is not likely due to the unanticipated PTM found within the same peptide. Similarly consistent

regulation pattern were seen in carbamylated peptides ionized in different charge states, demonstrating good intensity consistency. The confidence in our differential expression findings were further supported with the identification of 8 proteins (APOB, APOA1, ALB and 5 IGs), that have been previously demonstrated to be heavily carbamylated in atherosclerotic subjects/models with renal disease, and/or under inflammatory conditions. The differential changes in the relative abundance of carbamylated peptides of APOB, APOA1, ALB and IGG, as calculated by XIC, were depicted in Figure 4.

Chapter 4

Table 4.1. List of significantly regulated carbamylated peptides and proteins in human plasma.

Protein Accession	Protein description	Peptide sequence	Charge state	K-Carmylation Residue Position	Mean Peak Area (Ctrl)	%CV (Ctrl)	Mean Peak Area (MI)	%CV (MI)	MI:Ctrl	P value
P01023	Alpha-2-macroglobulin	K-YDVENCLAN*K*VDLSFSPSQSLPASHAHLR-V	4	K567	1.46E+07	9.26	3.79E+07	2.91	2.59	< 0.0001
		R-AVDQ*SVLLMK*PDAELSASSVYNLLPEK-D	3	K608	9.98E+07	1.02	1.59E+08	2.67	1.59	< 0.0001
		R-HNVYINGITYTPVSSTN*EK*DMYSFLEDM*GLK-A	4	K664	3.05E+07	3.48	5.41E+07	6.28	1.77	0.00033
P02787	Serotransferrin	R-K*PVDEYK*DCHLAQVPSHTVVAR-S	5	K252	1.16E+07	6.70	2.15E+07	2.83	1.85	< 0.0001
		R-SMGGK*EDLIWELLNQ*AQEHFGKDK-S	4	K278	8.12E+06	5.95	5.99E+07	13.07	7.38	0.00033
		R-SMGGK*EDLIWELLNQAQ*EHFGKDK-S	4	K278	8.30E+06	4.04	6.01E+07	13.45	7.25	0.00038
		R-SM*GGK*EDLIWELLNQAQEHFGKDK-S	4	K278	1.59E+08	0.63	2.90E+08	7.68	1.82	0.00052
		R-SMGGK*EDLIWELLNQAQEHFGKDK-S	4	K278	1.69E+07	1.49	3.27E+07	4.68	1.93	< 0.0001
P02768	Serum albumin	K-K*VPQVSTPTLVEVSR-N	2	K438	1.62E+08	1.29	3.14E+08	3.74	1.94	< 0.0001
		K-LVNEVTEFAK*TCVADESAENCDK-S	3	K75	4.38E+08	5.53	7.51E+08	8.78	1.72	0.00151
		K-TCVADESAENCDK*SLHTLFGDKLCTVATLR-E	4	K88	3.00E+08	1.86	9.02E+08	9.48	3.00	0.00026
		K-VFDEFKPLVEEPQN*LIK*Q*NCELFEQLGEYKFQNALLVR-Y	3	K413	6.19E+06	5.80	1.02E+07	6.12	1.65	0.00065
		K-VFDEFKPLVEEPQN*LIK*QNCELFEQLGEYKFQNALLVR-Y	4	K413	6.42E+07	4.58	1.32E+08	9.47	2.06	0.00079
		K-VFDEFKPLVEEPQNLIK*QN*CELFEQLGEYKFQNALLVR-Y	5	K413	1.74E+07	9.25	3.37E+07	5.21	1.94	0.00029
		K-VFDEFKPLVEEPQNLIK*QNCELFEQ*LGEYKFQNALLVR-Y	4	K413	1.14E+08	6.00	2.47E+08	7.94	2.15	0.00038
		K-VFDEFKPLVEEPQNLIK*QNCELFEQ*LGEYKFQNALLVR-Y	5	K413	2.31E+07	2.61	3.93E+07	3.48	1.70	< 0.0001
		K-VFDEFKPLVEEPQNLIK*QNCELFEQLGEYK-F	4	K413	1.19E+08	4.71	1.94E+08	1.36	1.63	< 0.0001
		K-VFDEFKPLVEEPQNLIK*QNCELFEQLGEYKFQNALLVR-Y	4	K413	3.59E+07	6.23	1.20E+08	4.91	3.34	< 0.0001
		K-VFDEFKPLVEEPQNLIK*QNCELFEQLGEYKFQNALLVR-Y	5	K413	1.48E+08	5.23	3.14E+08	6.95	2.12	0.00024
P00738	Haptoglobin	K-LPECEADDGCPK*PPEIAHGYVEHSVR-Y	5	K94	2.10E+06	7.91	1.35E+07	9.53	6.43	0.00011
		K-LPECEADDGCPK*PPEIAHGYVEHSVR-Y	4	K94	2.42E+07	8.41	6.53E+07	5.69	2.70	< 0.0001

K*, carbamylated lysine residue, Q*, deamidated glutamine residue, N*, deamidated asparagine residue; -, trypsin cleavage site; MI, Myocardial infarction; Ctrl, Control; %CV, percentage coefficient variation.

Table 4.2. List of significantly regulated carbamylated peptides and proteins in human plasma. (Continued)

Protein Accession	Protein description	Peptide sequence	Charge state	K-Carmylation Residue Position	Mean Peak Area (Ctrl)	%CV (Ctrl)	Mean Peak Area (MI)	%CV (MI)	MI:Ctrl	P value
P02647	Apolipoprotein A-I	K-ETEGLRQEMSK*DLEEVK-A	3	K112	2.28E+08	9.11	6.09E+08	8.76	2.67	0.00032
		K-VQPYLDDFQKK*WQ*EEM*ELYR-Q	4	K131	6.31E+07	2.68	1.53E+08	2.16	2.43	< 0.0001
P04114	Apolipoprotein B-100	R-MN*FK*QELN*GNTK-S	2	K3441	3.60E+07	6.80	5.65E+07	9.47	1.57	0.00376
P02763	Alpha-1-acid glycoprotein 1	K-TYM*LAFDVNDEK*N*WGLSVYADKPETTK-E	4	K137	4.30E+06	6.83	1.92E+07	5.13	4.46	< 0.0001
P01024	Complement C3	R-NN*NEK*DMALTAFVLISLQEAK-D	3	K1139	3.56E+08	0.51	7.84E+08	9.29	2.20	0.00052
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	R-GMADQ*DGLK*PTIDKPSEDSPPLEM*LGPR-R	4	K621	2.49E+08	11.75	4.03E+08	2.06	1.62	0.00093
B4E1F0	Plasma protease C1 inhibitor	K-K*VETN*MAFSPFSIASLLTQVLLGAGENTK-T	3	K167	2.20E+09	9.59	4.38E+09	5.84	1.99	0.00034
P01876	Ig alpha-1 chain C region	K-K*GDTFSCMVGHEALPLAFTQK-T	4	K307	6.60E+06	6.32	2.15E+07	4.83	3.25	< 0.0001
P01859	Ig gamma-2 chain C region	R-K*CCVECPPCPAPPVAGPSVFLFPPKPK-D	4	K101	4.67E+06	11.00	1.49E+07	10.07	3.20	0.00036
P01834	Ig kappa chain C region	K-ADYEKHK*VYACEVTHQGLSSPVTK-S	4	K82	1.23E+09	2.93	1.87E+09	1.43	1.52	< 0.0001
P01597	Ig kappa chain V-I region DEE	R-AGQSVN*K*YLNWYQQKPGK-A	3	K31	4.40E+08	1.74	1.18E+09	1.41	2.69	< 0.0001
P0CG06	Ig lambda-3 chain C regions	K-ADSSPAK*AGVETTTPSK-Q	3	K50	1.65E+07	5.55	2.51E+07	8.68	1.52	0.00321

K*, carbamylated lysine residue; Q*, deamidated glutamine residue, N*, deamidated asparagine residue; -, trypsin cleavage site; MI, Myocardial infarction; Ctrl, Control; %CV, percentage coefficient variation.

Chapter 4

Advances in proteomic techniques have afforded the assessment and quantitative evaluation of site-specific carbamylation on specific protein targets for peptide biomarker studies. The determination of specific carbamylation sites on peptide provides practicability and sensitivity for the development of standardized assays using for clinical usage. Elevation of non-enzymatic protein carbamylation in plasma has been tightly linked to chronic renal failure (CRF) accelerated by raised plasma urea, the increased levels of c-LDL and c-HDL in uremic patients have been shown and thought to promote the development of uremic atherosclerosis (187-191). Several investigations have showed that c-LDL apoliprotein B (APOB), confers multiple pro-atherogenic effects , such as induction of endothelial cell death (187), impaired clearance from plasma (366) and induction of smooth muscle cells expression of adhesion molecules and proliferation (367). Whereas, c-HDL apolipoprotein A1 (APOA1) have been reported to possess compromised functionality in reverse cholesterol transport from macrophages and loss of anti-inflammatory and anti-oxidant activities, thus promoting atherogenesis (353, 368). In this study, under normal renal functions, carbamylated peptides of APOB (R-MN*FK*QELN*GNTK-S, ~1.57 fold-change, p = 0.0038, Figure 4.3A) and APOA1 (K-ETEGLRQEMSK*DLEEVK-A, ~ 2.67-fold change, p = 0.00032, 4.3B) were elevated in patients who have had MI, presumably driven by inflammation. These observations are in good agreement with the study by Wang et al (182), the only report to date that demonstrates the direct association of MPO-mediated protein carbamylation to inflammation and atherogenesis independent of urea load.

Carbamylated ALB levels are strongly predictive of mortality risk in patients with end-stage renal disease (369, 370). The susceptibility of a protein to irreversible carbamylation depends on various factors, such as the availability of K residue to isocyanic acid, the concentration of isocyanic acid and the biological life-span of the protein (371). The progressive stable accumulation of homocitruline residues are more likely to occur on high abundance and long-lived protein (348). ALB is considered a long-lived circulatory protein with half-life of approximately 2 weeks (372). The concentration of carbamylated ALB is significantly correlated with concentrations of urea in the circulation, therefore quantified carbamylated ALB

provides an index of overall carbamylation burden (186). In inflammatory CKD, the adverse effects on various cellular response induced by the interaction of carbamylated ALB have been evidently presented, such as the inhibition of polymorphonuclear neutrophil respiratory burst (350), the stimulation of glomerular mesangial cells proliferation and collagen deposition (373) and the induction of nephrotoxicity (181). Here, all carbamylated peptides of ALB were significantly raised in MI group, with peptide K-VFDEFKPLVEEPQNLIK*QNCEL FEQLGEYKFQNALLVR-Y exhibiting the most regulated fold-change of ~3.34 ($p < 0.0001$, Figure 4.3C). To date, the direct relationship between non-uremic systemic measures of carbamylated ALB and incidents of CAD burden has not yet been shown.

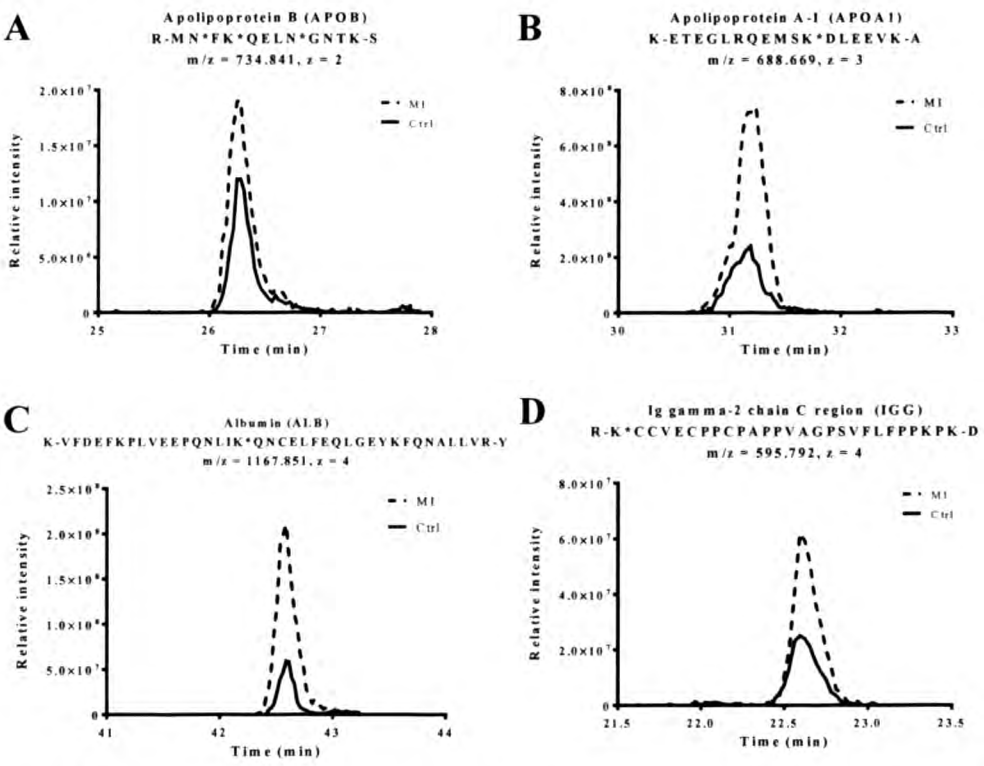

Figure 4.3. Extracted ion chromatograph (XIC) of differentially regulated carbamylated peptides. XIC of significantly ($p < 0.05$) differentially regulated carbamylated peptides identified in patients who have had myocardial infarction (MI) and control (Ctrl) subjects. (A) XIC of carbamylated peptide of APOB (R-MN*FK*QELN*GNTK-S) (B) XIC of carbamylated peptide of APOA1 (K-ETEGLRQEMSK*DLEEVK-A). (C) XIC of carbamylated peptide of ALB (K-VFDEFKPLVEEPQNLIK*QNCELF EQLGEYKFQNALLVR-Y). (D) XIC of carbamylated peptide of IGG (K-K*GDTFSCMVGH EALPLAFTQK-T). K*, carbamylated lysine residue; Q*, deamidated glutamine residue, N*, deamidated asparagine residue.

Currently, available knowledge on circulatory immunoglobulins (IGGs) is considerably lesser and generally in context of inflammatory RA. Carbamylated IGGs have been found in the inflamed synovial fluid of RA patients and sera of dialyzed patients (374). It has been demonstrated that carbamylated antibodies exert an inhibitory effect on classical complement activation, due to impaired binding to complement C1q, resulting in decreased clearance of immune complexes (374). Increased levels of carbamylated peptides of IGGs were notice in the MI group, with the most regulated fold-change of 3.25 ($p < 0.0001$), displayed by K-K*GDTFSCMVGH EALPLAFTQK-T (Figure 4.3D). Carbamylated peptides and proteins are sources of antigens and should elicit the antibody formation against specific carbamylated epitopes. In the context of atherosclerosis, it is tempting to speculate that the loss of IGG effector function due to carbamylation could be an evasion mechanism that delay clearance of immunogenic carbamylated antigens, such as cLDL, thus promoting atherogenesis.

Of substantial interest, seven novel carbamylated proteins showed significant ($p < 0.05$) up-regulation in MI group were identified. Among which, haptoglobin (c-HP), serotransferrin (c-TF) and alpha-2-macroglobulin (c-A2M) were identified with at least two high confident carbamylated peptide entry. While alpha-1-acid glycoprotein 1 (c-ORM1), complement C3 (c-C3), inter-alpha-trypsin inhibitor heavy chain H1 (c-ITIH1) and plasma protease C1 inhibitor (c-SERPING1) were identified with one high confident carbamylated peptide. Taking into consideration that the degree of carbamylation on proteins is likely to reflect the inflammatory burden post-MI. Taking reference from most prominent regulated fold change between MI and Ctrl, and the pertinent differential changes evidently shown in Figure 4.4, c-HP (K-LPECEADDGCPK*PPEIAHGYVEH SVR-Y, ~ 6.43-fold change, $p = 0.00011$, Figure 4.4A), c-TF (R-SMGGK*EDLIWELLNQ*AQEHFG KDK-S, ~ 7.38-fold change, $p = 0.00033$, Figure 4.4B) and c-ORM1 (K-TYM*LAFDVND EK*N*WGLSVYAD KPETTK-E, ~ 4.46-fold change, $p < 0.0001$, Figure 4.4C) displayed obvious potential as novel inflammatory peptide markers of MI.

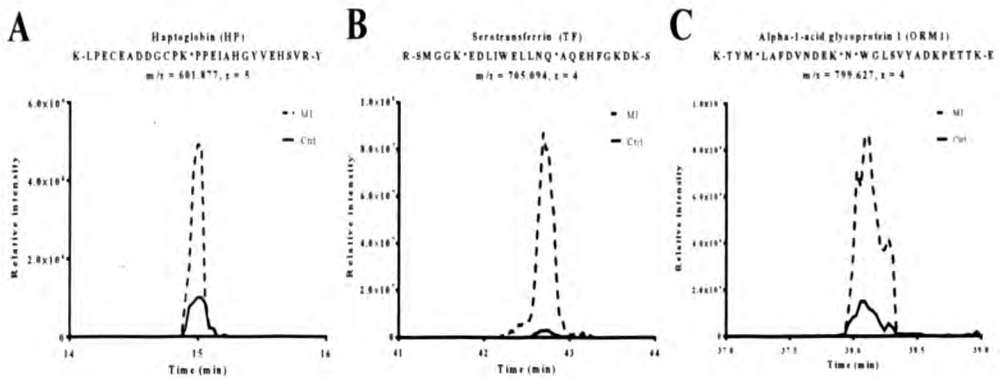

Figure 4.4. Extracted ion chromatograph (XIC) of proposed carbamylated peptide markers. XIC of potential novel inflammatory peptide markers of myocardial injury, identified in patients who have had myocardial infarction (MI) and control (Ctrl) subjects. (A) XIC of carbamylated peptide of HP (K-LPECEADDGCPK*PPEIAHGYVEH SVR-Y, ~ 6.43-fold change), (B) XIC of carbamylated peptide of TF (R-SMGGK*EDLIWELLNQ*AQEHFGKDK-S, ~ 7.38-fold change). (C) XIC of carbamylated peptide of ORM1 (K-TYM*LAFDVNDEK*N*WGLSVYAD KPETTK-E, ~ 4.46-fold change). K*, carbamylated lysine residue; Q*, deamidated glutamine residue, N*, deamidated asparagine residue.

Beyond the effects of declined kidney function, we have showed the significant elevation of 30 signature carbamylated peptides on 15 specific protein in MI compared to Ctrl, which is likely a consequence of inflammatory driven carbamylation, resulting from the enhanced inflammatory state triggered after myocardial injury. This is the first differential quantitative assessment on characterized carbamylated plasma peptides/proteins that likely reflects in vivo carbamylation driven by inflammatory-derived cyanate in human plasma.

4.4.5 Novel inflammatory carbamylated markers predictive of MI

The involvement of carbamylation in the pathogenesis of several human pathologies, particularly in CKD and atherosclerosis, have led to its rise of prominence. As measured values of carbamylated proteins are reflective of chronic uremia and or chronic inflammation burden, there is considerable interest for their usage as diagnostics biomarkers. To date, two carbamylated proteins, carbamylated hemoglobin (c-Hb) (375) and c-LDL (376) are currently being evaluated for their predictive values in renal failure and atherosclerosis respectively. Based on the findings from differential quantitative assessment of carbamylated peptides, c-HP

(K-LPECEADDGCPK*PPEIAHGYVEH SVR-Y, ~ 6.43-fold change, p = 0.00011, Figure 4.4A), c-TF (R-SMGGK*EDLIWELLNQ*AQEHFG KDK-S, ~ 7.38-fold change, p = 0.00033, Figure 4.4B) and c-ORM1 (K-TYM*LAFDVND EK*N*WGLSVYAD KPETTK-E, ~ 4.46-fold change, p < 0.0001, Figure 4.4C) were proposed as potential inflammatory peptides markers of MI.

Protein carbamylation alters the structural and biological properties of proteins, resulting in the partial or complete loss of function and mechanical properties (189). As these findings are novel, the speculated pathological association of c-HP, c-TF and c-ORM1 in atherosclerotic vascular disease discussed here were directed at the loss of protein function. The raised plasma levels of positive acute phase proteins HP and ORM, are indicative of intense inflammatory response following myocardial injury. It has been reported that plasma HP levels were raised in the initial phase of acute MI before hemolysis kicks in (377). In plasma, HP binds to free haemoglobin (Hb) to inhibit its oxidative activity and to remove it from the circulation, preventing iron loss through the kidneys and renal damage (378). In addition to being an antioxidant, HP modulate many aspects of inflammatory response such as immunosuppression of lymphocytes function (379) , anti-bacterial activity (380), inhibition of nitric oxide produced by activated macrophages (381), inhibition of pro-inflammatory prostaglandins (382) and aids in wound repair by promoting of angiogenesis (383). We speculate that c-HP, presumably with a loss of function, may result in the deleterious pro-inflammatory and pro-oxidation effect that favors the pathogenesis of atherosclerotic conditions.

Despite the first description of ORM1 in 1950s, the exact biological function remains ambiguous (384). ORM1 have been described as an immunomodulatory agent, and at high concentration mediates several anti-inflammatory processes such as the inhibition the activation of polymorphonuclear neutrophils (384, 385), inhibition of platelet aggregation (386), inhibition of proinflammatory interleukin 1 activity (387) and acceleration of wound healing (388). In addition, ORM1 have been shown to reduce ischemia-reperfusion (I/R) -induced apoptosis and inflammation (248). Based on these experimental findings, ORM1 exert protective effects under periods

of intense inflammation and tissue injury, similar to that of MI. As protein carbamylation often always result in the loss of protein function, one can raise the hypothesis that c-ORM1, may possess altered, suppressed or abrogated anti-inflammatory effects, hence aggravating inflammation associated atherosclerotic conditions or tissue injury.

Negative phase acute protein TF, are iron-binding proteins that transports free iron between sites of adsorption, sites of degradation, sites of utilization and sites of storage (389). Elevated levels of c-TF in MI group were observed in this study. Increased levels of plasma TF have been associated with iron deficiency (390). Iron-bound TF is recognized generally by proliferating cells expressing TF receptor (TFR) and internalized via endocytosis for utilization (391). The decreased levels of TF during inflammation is in part due to the increased expression of surface TFR on maturing macrophages (392). Reports have shown that TFR expression on lymphocytes were essential for cell division and proliferation (393, 394). Apart from its iron transport capacity, it has been suggested that TF confers immunologic roles, specifically in early T cell differentiation in the periphery and thymus (390). Keeping with this, it is tempting to speculate that TRF is unable to recognize c-TF during inflammation, hence the increased in circulatory levels were detected. As a result, elevated levels of c-TF may cause the decreased delivery of iron to macrophages containing LDL at inflammation sites, causes macrophage defense dysfunction, thus promoting the formation of foam cells, the key initial steps in promoting atherogenesis.

Despite considerable carbamylation studies performed to date, the potential implication of carbamylated proteins in the progression of age-dependent atherosclerosis, in the absence of uremia, unfortunately has been elusive. In addition, systemic protein damages as a consequence of carbamylation under inflammatory conditions have never been assessed. These novel findings raised the possibility that the abovementioned carbamylated peptides/proteins may have causal roles in favouring atherogenesis and thromboembolic events, their potential values as novel predictive indicators of MI risk, necessitates further examinations in a valid outcome

study. Therapeutic modulation of these proteins may influence the rate of atherosclerosis progression. The new discovery in this study will serve as a platform for future molecular investigations to unravel the roles of disease relevant carbamylated peptides of c-HP, c-ORM1 and c-TF.

4.5 Conclusion

In this study, label-free differential carbamylated peptide analyses were performed using plasma obtained from MI and Ctrl groups. In order to minimize experimental-induced carbamylation, the use of urea was avoided during sample preparation in this study. The occurrence of total plasma protein carbamylation in the absence of uremia was estimated to be approximately 4% each in both MI and Ctrl groups. A total of 30 high confident (FDR $\leq$ 1%, $p < 0.05$) carbamylated peptides assigned to 15 proteins were found to be significantly elevated ($p < 0.05$) in the MI group, with at least 1.5-ratio fold change. Among which, seven novel carbamylated proteins (c-HP, c-TF, c-ORM1, c-SERPING1, c-C3, c-A2M and c-ITIH4) were identified. Of substantial interest, the prominent up-regulation of c-HP (K-LPECEADDGCPK*PPEIAHGYVEH SVR-Y, ~ 6.43-fold change, $p = 0.00011$), c-TF (R-SMGGK*EDLIWELLNQ*AQEHFGKDK-S, ~ 7.38-fold change, $p = 0.00033$) and c-ORM1 (K-TYM*LAFDVNDEK*N*WGLSVYAD KPETTK-E, ~ 4.46-fold change, $p < 0.0001$) have made them the best candidate inflammatory peptide markers of MI. We speculated that the loss of HP, TF and ORM1 function due to carbamylation would may drive the development of inflammation associated atherosclerotic conditions and impair tissue healing. The diagnostic values and possible utility of these novel carbamylated peptides in the treatment and management of atherosclerotic CAD, warrants further examinations in a valid outcome study.

Chapter 5

Novel Prolong Ultracentrifugation with Electrostatic Repulsion-Hydrophilic Interaction Chromatography [PUC-ERLIC] Method Facilitates the Simultaneous Recovery of Secretory and Extracellular vesicular Glycoproteins from Plasma

Chapter 5

5.1 Abstract

Pathological secretions of glycoproteins and extracellular vesicles (EVs) present in plasma are excellent sources for disease biomarkers. Unfortunately, detection of these circulatory proteins remains challenging as they are present in relatively low abundance in highly complex plasma. Intensive research have laid the groundwork for the effective enrichment of either glycoproteins or EVs from biofluids, for biomarker development. A single method that affords parallel enrichment of both structures from biofluids would be highly desirable for biomarker discovery and useful for clinical application. In order to harvest a broader spectrum of candidate biomarkers for use in mass spectrometry-based proteomic profiling, we have developed an enrichment method that couples prolong ultracentrifugation (PUC) with electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) [PUC-ERLIC] to facilitate the simultaneous recovery of both glycoproteins and EV from crude plasma. Following PUC, the recovery of a pigmentation suspension found to be highly concentrated in soluble and EV-derived glycoproteins was described for the first time. Through this novel systematic PUC-ERLIC approach, we reported the confident (FDR $\leq$ 1%) identification of 127 secretory and EV-derived glycoproteins, 599 unique glycopeptides and 361 unique glycosylation sites in a single LC-MS/MS run. Without the use of immunodepeletion strategy, 48 low-abundant glycoproteins with concentration ranging from pg/mL – ng/mL were detected in crude plasma. The novel method described here mediates the paucity in our knowledge on EV-derived glycoproteomics, and will benefit similar future plasma-based EV glycoproteomics studies. This method can be readily applied to any type of biofluids and assist in the detection of novel souble/EV-derived biomarkers in a wide variety of human diseases.

5.2 Introduction

Human plasma is one of the most comprehensive and ideal source for biomarker discovery. In addition to resident proteins, plasma contains repertoire of other sub-proteomes including glycoproteins and extracellular vesicles (EVs) that are secreted or released from organs and compartments throughout the body, mirroring the

pathophysiological status of the body (395). Protein glycosylation, particularly asparagine-linked (N-linked), is the most prevalent and a major post-translational modification (PTM) in plasma (169, 170). Circulatory N-linked glycosylated proteins are typically secreted, enclosed in EVs or shed from plasma membrane (168). Although plasma glycoproteins are present in low abundance, they play crucial roles in both biological and pathological processes, including cell-cell interactions, signal transductions, protein stability, developmental process, immune defenses and aberrant protein glycosylation has been correlated to diseases such as cancer and neurodegenerative disorders (160-165). Hence, strategies that afford the enrichment of these low abundant entities for characterization of aberrant protein glycosylation, would generate clinically meaningful information on human pathologies, facilitating discovery of novel PTM biomarkers.

Although the study of N-glycoproteome have been made possible by mass spectrometry (MS)-based techniques, detection of these low-abundant glycopeptides are analytically challenging due to the structural heterogeneity of glycosylation and the vast dynamic range of protein abundance in plasma (396). In order achieve in-depth glycoprotein detection, numerous enrichment approaches that selectively enrich for glycoproteome from complex biofluids prior to MS analyses have been developed. Selective enrichment of glycoproteins and glycopeptides in biofluids can be attain using lectin affinity based method (171, 172), hydrazide chemistry based-solid phase extraction method (173), boronic acid method (397), size-exclusion chromatography (398), hydrophilic interaction chromatography (399, 400) and electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) (401, 402), with the first two more widely applied.

Membrane-bound vesicles termed as EVs are released by almost all cell types into the circulation under both normal and pathological conditions (192). Functional studies have evidently demonstrated that EVs play key roles in numerous physiological processes including regulation of cell signaling (193, 194), cell-cell communication (195, 196), blood coagulation(403), immune surveillance (197, 198), antigen presentation (404), stem cell maintenance (199) and tissue repair(200).

Additionally, EVs have been tightly implicated in inflammation (405), neurodegenerative diseases (199, 201), vascular diseases (202, 203), tumorigenesis (204, 205), diabetes (406) and infectious diseases (206, 207). The involvement of EVs in a myriad of biological and pathological process have presented them as highly attractive bioactive targets for disease biomarkers. Similar to glycoproteins, plasma EVs are present in low abundance, as such they have to be selectively isolated for their exploitation. Besides differential centrifugation followed by ultracentrifugation (UC) being gold standard and widely applied method for EVs isolation (208), density gradient centrifugation (209), filtration techniques (210), microfluidic device (407), System Biosciences ExoQuick™ precipitation (208) and immune-affinity capture beads (211) are alternative methods utilized for isolation of EVs.

In view of the persistent interest for plasma glycoproteins and EVs as attractive sources of potential clinical biomarkers, a single method that affords parallel enrichment of both structures from bodily fluids would be highly desirable for biomarker discovery. In the present study, we introduced a single enrichment strategy that couples prolonged ultracentrifugation (PUC) with ERLIC [PUC-ERLIC] for extraction of both secretory and EV in crude plasma. Without the use of any immunodepletion strategies, we showed that simultaneous recovery of high- and low-abundant secretory and EV-enriched glycoproteins are achievable with PUC-ERLIC method. Through this novel approach, a total of 599 unique N-glycopeptides and 361 unique N-glycosylation sites assigned to 127 glycoproteins were confidently (FDR $\leq$1%) identified in crude plasma, including 48 glycoproteins with concentration ranging from pg/mL – ng/mL. The method described here can be readily applied to any type of biological fluid and contribute to the identification of novel biomarkers of human diseases.

5.3 Materials and Methods

5.3.1 Chemicals

All water and acetonitrile (ACN) used in this experiment were of high performance liquid chromatography (HPLC) grade (Thermo Scientific, Waltham, MA). All chemicals were purchased from Sigma-Aldrich (St Louis, MO) unless stated otherwise. All buffers used in proteomic sample preparation were supplemented with protease inhibitors (1:50, v/v) and phosphatase inhibitors (1:10, v/v) (Roche Diagnostics, Mannheim, DE).

5.3.2 Human plasma

The study was approved by the National Healthcare Group Domain Specific Review Board (NHG DSRB). Blood was collected into lithium-heparin vacutainers from patients diagnosed with stable angina pectoris (NMI, n=20) for period more than one month prior to coronary artery bypass graft surgery (CABG) and processed immediately. Plasma aliquots were cryopreserved at -80°C until subsequent proteomic processing. Written informed consent of participation was obtained from all recruited patients.

5.3.3 Simultaneous enrichment of secretory glycoprotein and extracellular vesicles

In order to normalize biological variations, for each group, plasma from individual patients were equivalently combined to a final volume of 5 mL. 5mL of pooled plasma was diluted with 25mL of 1 x phosphate buffer saline (PBS) and then differentially centrifuged at 200 x g (30min), 2,000 x g (30min) and 12,000 x g (60min) at 4°C to exclude intact cells and cellular debris. The resultant supernatant was aliquoted into a 25 x 89mm polycarbonate tube (Type 50.2 Ti rotor, Beckman Coulter, CA) and enriched plasma glycoproteins and extracellular vesicles were simultaneously pelleted at 200,000 x g (18h, 4°C) using a Beckman L100-XP Ultracentrifuge (Beckman Coulter, CA). The enriched glycoprotein and extracellular

vesicular fraction was then re-suspended in 1 x PBS and pelleted again at 200,000 x g (18h, 4°C) to remove any residual contaminants.

5.3.4 Cryo-electron microscopy (Cryo-EM)

Electron microscope grids coated with holey carbon film (R2/2 Quantifoil) were glow discharged, and a 4μL droplet of extracellular vesicle-enriched suspension was deposited onto the grid at 99% humidity. Excess liquid was blotted with filter paper and plunged into liquid ethane (Vitrobot, FEI Company). Cryo grids were imaged using a field emission gun transmission electron microscope operated at 80kV (Arctica, FEI Company), and equipped with a direct electron detector (Falcon II, FEI Company). Images were recorded at a nominal magnification of 23,500 X.

5.3.5 Western blot analyses

A total of 50 μg secretory and extracellular vesicle proteins were lysed in 4 x Laemmli protein sample buffer (BioRad, Hercules, CA), resolved on 8-12% polyacrylamide gels, and transferred onto 0.45 μM nitrocellulose membranes (BioRad, Hercules, CA). Membranes were blocked and incubated overnight at 4°C with anti-Alix antibody (1:1000 Ab88743; Abcam, Cambridge, *MA*), anti-CD9 antibody (1:1000 sc-13118; Santa Cruz Biotechnology, CA) and anti-CD81 antibody (1:1000 sc-9158; Santa Cruz Biotechnology, CA), followed by incubation with horseradish-peroxidase-conjugated secondary antibody (Santa Cruz Biotechnology, CA; 1:4000, sc-2371, sc-2370), and then immunodetection using an enhanced chemiluminescence assay (Millipore Coporation, Billerica, MA).

5.3.6 In-solution tryptic digestion

Proteomic sample preparation was performed according to previously described methods designed to minimize experimentally-induced deamidation (306, 307), except for minor modifications. Briefly, the secretory and extracellular vesicle-enriched glycoproteins were re-suspended in 2 mL lysis buffer containing 8 M urea and 50 mM ammonium acetate (pH 6.0). Plasma proteins were quantified in a 96-

well plate using a bicinchoninic acid (BCA) assay according to the manufacturer's protocol. Disulfide bonds were reduced by incubating 200 µg protein in 20 mM dithiothreitol (DTT) for 3 h at 30°C, and were then alkylated in the dark using 55 mM iodoacetamide (IAA) for 1h at room temperature. Prior to tryptic digestion, the urea concentration was diluted to less than 1 M using 50 mM ammonium acetate buffer (pH 6.0) to ensure optimal trypsin activity. Proteins were enzymatically digested overnight at 37°C using sequencing-grade trypsin (Promega, Madison, *WI*) at a 1:100 ratio (w/w, trypsin:protein). The tryptic peptides were desalted using a Sep-Pak C18 cartridge (Waters, Milford, *MA)* and the eluted peptides were dried in a vacuum concentrator.

5.3.7 Selective enrichment of glycosylated peptides

Selective enrichment of glycosylated peptides was performed accordingly to a previously described method (408), with minor modifications. Briefly, vacuum-dried peptides were reconstituted in 200 µL mobile phase A (80% acetonenitrile [ACN] containing 0.1% formic acid [FA]) and fractionated using a PolyWAX LP weak anion-exchange column (4.6 × 200 mm, 5µm, 300Å; PolyLC, Columbia, MD) on a Shimadzu Prominence UFLC system (Kyoto, JP). The UV spectra of the peptides were collected at 280 nm. Mobile phase A and mobile phase B (10% ACN, 2% FA) were used with a 90min gradient of 0 - 5% B over 8min and 5−28% B over 37min followed by 45min at 100% B (constant flow rate of 1mL/min). Forty-five separate fractions were collected and vacuum dried. Dried peptides in each fraction were then reconstituted in 3% ACN, 0.1% FA for subsequent LC-MS/MS analysis.

5.3.8 Deglycosylation

Glycopeptides were deglycosylated using PNGase F (New England BioLabs, Beverly, *MA*) in 50 mM AA as previously described (306, 307) to minimize experimentally-induced deamidation. Peptides were deglycosylated overnight at 37°C and then dried using a vacuum concentrator prior to reconstitution in LC-MS/MS compatible buffer (3% ACN, 0.1% FA).

5.3.9 LC-MS/MS

The dried peptides were injected and analyzed using a LC-MS/MS system that comprised a Dionex Ultimate 3000 RSLC nano-HPLC system coupled to an online LTQ-FT Ultra™ linear ion trap mass spectrometer. Approximately 3 µg of peptides from each fraction were injected into a Zorbax peptide trap column (Agilent Technologies, Santa Clara, CA) via the auto-sampler of the Dionex system, and were subsequently resolved in a capillary column (75µm × 10cm) which was packed with C18 AQ (5µm, 300Å; Bruker-Michrom, Auburn, CA), and run at a flow rate of 300nL/min. Buffer A (0.1% FA in HPLC water) and buffer B (0.1% FA in ACN) were used to establish the 60min gradient; starting with 1 min of 5–8% B, 44min of 8–32% B, 7min of 32–55% B, 1min of 55–90% B and 2min of 90% B, followed by re-equilibration in 5% B for 5 min. The samples were ionized in an ADVANCE™ CaptiveSpray™ Source (Bruker-Michrom) with an electrospray potential of 1.5kV. The LTQ-FT Ultra was set to perform data acquisition in the positive ion mode. A full MS scan (350–1600 m/z range) was acquired in the FT-ICR cell at a resolution of 100,000 and a maximum ion accumulation time of 1000ms. The automatic gain control target for FT was set at 1×10^6, and precursor ion charge state screening was activated. The linear ion trap was used to collect peptides and measure the peptide fragments generated by CID. The default automatic gain control setting was used in the linear ion trap (full MS target 3.0×10^4, MSn 1×10^4). The 10 most intense ions above a 500 count threshold were selected for fragmentation in CID (MS2), which was performed concurrently with a maximum ion accumulation time of 200 ms. Dynamic exclusion was activated for the process, with a repeat count of 1 and exclusion duration of 60s. For CID, the activation Q was set at 0.25, isolation width (m/z) was 2.0, activation time was 30ms, and normalized collision energy was 35%. Duplicate LC-MS/MS runs per fractions were performed.

5.3.10 Data analyses

The extract_msn program (version 4.0) as found in Bioworks Browser 3.3 (Thermo Electron, Bremen, Germany) was used to extract tandem MS spectra in dta format

from the raw data of LTQ-FT ultra. Protein identification was performed by querying against the extracted Uniprot Human database (Released on 20131129; 176946 sequences, 70141034 residues), by means of an in-house Mascot server (version 2.4.1, Matrix Science, Boston, MA). Target-decoy search strategy was employed for the estimation of false positive identification. The search was limited to a maximum of 2 missed trypsin cleavages; # 13 C of 2; peptide precursor mass tolerances of 5.1 ppm; and 0.8 Da mass tolerance for fragment ions. Carbamidomethylation of cysteine residues was set as a fixed modification, while oxidation of methionine residues and deamidation of asparagine residues were set as variable modifications. Mascot results were exported to csv file format and then further processed in Microsoft Excel. Identified peptides were sorted from smallest to largest Mascot peptide expect value (a measure of random match probability), and largest to smallest ion score. The resultant peptide list was used for calculation of false discovery rate [FDR = 2*(decoy hits/total hits)*100%], using an in-house script. The FDR associated with the searched dataset was 1% at the peptide level. In order to reduce the presence of outliers in our dataset, only peptides with ion scores greater than homology or identity scores were selected for further analysis.

5.3.11 Data Annotation

Several open source online bioinformatics software tools including STRAP V1.5 (409), AmiGO V2 (340, 410) and DAVID V6.7 (411) were used to determine plasma glycoproteins function. Peptide sequences were submitted to NCBI BLASTP (222, 223) search against a non-redundant protein sequences (nr) using NCBI server default settings.

5.4 Results and Discussion

5.4.1 Development of a novel PUC-ERLIC strategy for glycoproteome enrichment

Circulatory glycoproteins and EVs are highly sought after as potential clinical biomarkers. Unfortunately, the parallel isolation of both structures from plasma have

not been imaginable due to technical and complexity challenges. In attempt to achieve the simultaneous recovery of both entities from crude plasma, modification to the convention UC method were made to simplify the complexity of plasma. As displayed in Figure 5.1, crude plasma was diluted five-folds prior to UC to reduce the viscosity and centrifugation speed and time were increased from 100, 000 x g, 2 h to 200, 000 x g, 18 h to improve the efficiency of sedimentation. With this approach, an obvious yellow suspension was recovered for the first time and subsequent proteomics processing revealed the suspension to be highly enriched in soluble glycoproteins and EVs. Selective enrichment of N-soluble and EV glycopeptides were achieved using ERLIC. Finally, PNGasse F was used to release the N-linked glycan the glycopeptides prior analyses in LC-MS/MS.

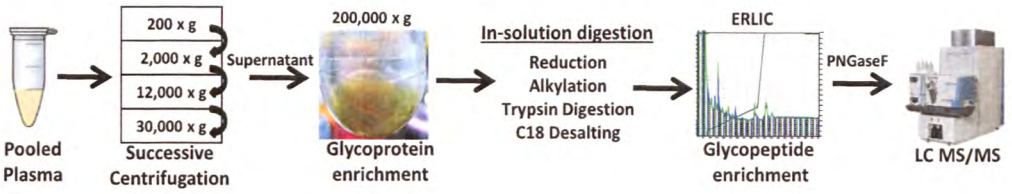

Figure 5.1. Work flow of novel enrichment strategy. PUC-ERLIC proteomics workflow showing plasma sample preparation for the simultaneous isolation of secretory and extracellular vesicle-enriched glycoproteins prior to mass spectrometry. Yellow suspension collected after PUC was found to be enriched in glycoproteins.

5.4.2 Plasma EVs enrichment by PUC

Differential centrifugation followed by UC has been considered the gold standard for EVs precipitation, the limitation of this method is the co-precipitation of protein aggregates, apoptotic bodies and other membrane fragments (412). We acknowledge this presence of impurities in our EVs preparation and hence termed the isolated vesicular fractions in this study as EV-enriched. Successful enrichment of EVs was confirmed by cryo-EM, as depicted in Figures 5.2A and 5.2B, harvested EVs were circular, membrane-bound and approximately 50 – 100 nm in diameter, displaying size range similar to that of exosomes (194). Additionally, the isolation of EVs by PUC were reaffirmed by the detection of broadly used exosome markers Alix, CD9 and CD81 (413) in western blot analyses (Figure 5.2C). Proteomics analyses revealed the detection of several exosome-specific proteins including FLOT1,

TSG101, HSP70, ALIX/PDCD6IP, CD9, CD81 and CD63. Collectively, we have firmly demonstrated the enrichment of EVs from crude plasma using PUC.

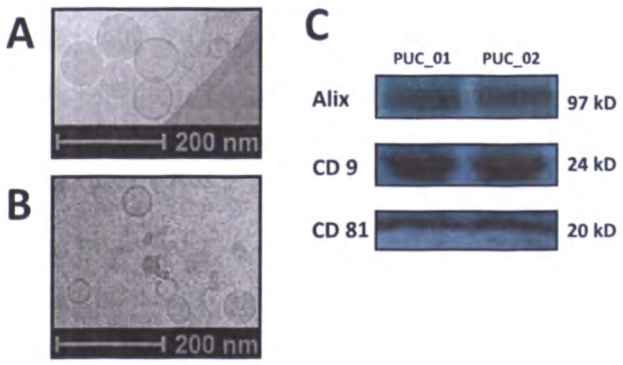

Figure 5.2. Verification of plasma extracellular vesicles (EVs) isolation. Plasma EVs isolated using prolonged ultracentrifugation (PUC). (A) Electron micrographs of harvested extracellular vesicles (sizes 50–100nm, recorded at 23,500 X magnification. (B) Electron micrographs of extracellular vesicles recorded on carbon (23,500 X magnification with a defocus of -6μM). (C) Western blot analyses of harvested proteins using extracellular vesicle markers Alix, CD9 and CD81. PUC-01 and PUC-02 are biological replicates.

5.4.3 Simultaneous recovery of soluble glycoproteins and EVs facilitated by PUC

While standard UC typically uses 2 h for the isolation of EVs, in the present study extended UC duration was implemented to improve the sedimentation efficiency of EVs from viscous plasma. In order to assess the advantages of PUC overuc, data obtained from both approaches were compared. Visually (Figure 5.3A), using the same volume of starting material, pellet recovered from standard UC was translucent with a tint of yellow, whereas pellet obtained from PUC was notably larger with darker shade of yellow, and encircled by a pigmented dense suspension. PUC demonstrated better pellet recovery as compared to UC. Despite the extensive application of UC in EV isolation, this is the first report that describe the sedimentation of pigmented plasma EV-suspension that is also highly enriched in secretory glycoproteins, to our knowledge.

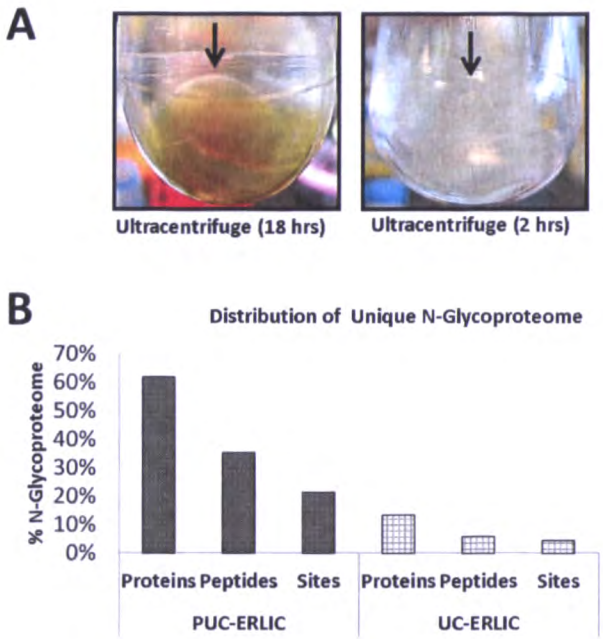

Figure 5.3. Enrichment of plasma secretory and extracellular vesicle glycoproteins. Prolonged ultracentrifugation (PUC) enables simultaneous recovery of secretory and EV-enriched glycoproteins from human plasma. (A) Comparison of fractions recovered from 5mL total plasma after 18h or 2h of UC. Yellow pellet recovered after PUC (18h) was significantly larger and more visible than the pellet recovered after UC (2h). (B) Proportion of unique N-glycoproteins, N-glycopeptides and N-glycosylated sites obtained when using PUC-ERLIC or standard UC-ERLIC approaches (both with PNGase F treatment). Significant glycoprotein enrichment was observed after PUC-ERLIC. The glycosylation percentage was calculated using the number of unique glycosylated products divided by the sum of all unique glycosylated and unmodified peptides.

Next, the glycoproteome of the EV-enriched suspension obtained by standard UC-ERLIC and PUC-ERLIC were examined. As shown in Figure 5.3B, PUC-ERLIC approach exhibited improvement in the identification of N-glycosylated proteins by approximately four-fold when compared to UC-ERLIC approach (detailed list of glycoproteome can be found in supplemental data S2, worksheet PNGaseAll01 in Ref. (168)), indicating that enrichment of soluble N-glycoproteins was achieved with PUC-ERLIC approach. Furthermore, of the 127 glycoproteins identified in PUC-ERLIC approach, 99 proteins were annotated by GO as secreted proteins. Our above-mentioned results showed that plasma secretory glycoproteins and EVs were simultaneously sedimented in a form of yellow suspension, facilitated by PUC. We speculate that highly-branched glycoproteins exist as larger and denser constituents in plasma than unmodified proteins, hence are pelleted alongside with EVs with

extended UC. We have demonstrated the feasibility of PUC-ERLIC approach in the parallel isolation of soluble and EVs from crude plasma.

5.4.4 Assessment of PUC-ERLIC approach in glycoproteome enrichment

In attempt to prove that the successful enrichment of glycoproteins is attributed by the novel PUC-ERLIC strategy, and not by PUC approach alone or ERLIC approach alone, data obtained from these three approaches were assessed. Depicted in Figure 5.4A, the percentage of unique plasma N-glycoproteins identified was highest in PUC-ERLIC approach (~62%), followed by ERLIC alone approach (~34%) and PUC alone approach (~24%). The lower N-glycoproteome detected in ERLIC alone and PUC alone approach demonstrates the classic analytical challenge in characterizing the low abundant glycoproteins in crude plasma. In PUC alone approach, the total unique glycosylated proteins identified was about 24%, affirming the recovery of EV-enriched glycoproteins with PUC. The Venn diagram (Figure 5.4B) of overlapping glycoproteins identified by PUC-ERLIC approach, PUC alone approach and ERLIC alone approach, showed an overlap of approximately 43% (59 proteins) among the three approaches, demonstrating fairly good complementation. Additionally, PUC-ERLIC approach contributed to the highest number of unique glycosylated proteins exclusively detected in crude human plasma. We have clearly showed that the enrichment of glycoprotein is indeed credited to the synergistic performance of PUC-ERLIC strategy when used in combination.

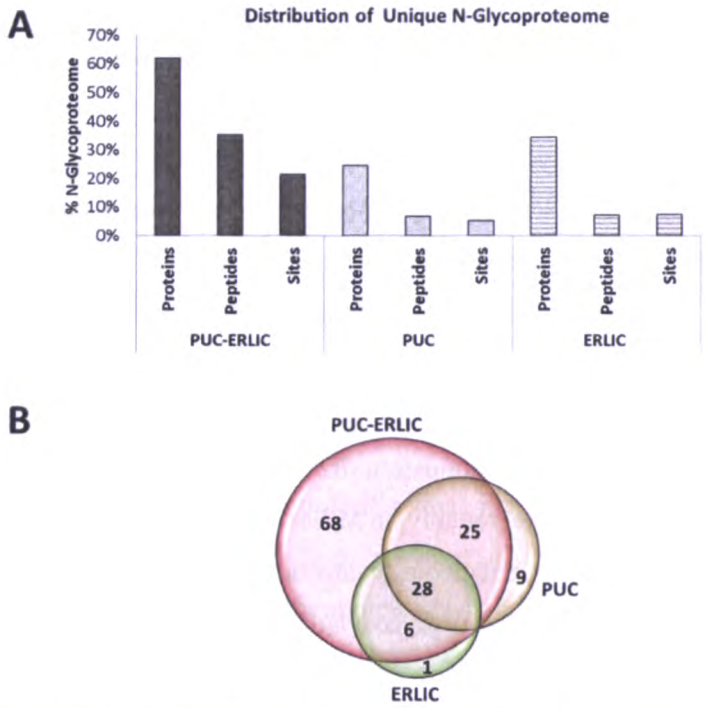

Figure 5.4. Glycoprotein enrichment by PUC-ERLIC method. PUC-ERLIC facilitates enrichment of the N-glycoproteome of human plasma. (A) Proportion of unique N-glycoproteins, N-glycopeptides and N-glycosylated sites obtained using PUC-ERLIC, PUC alone, or ERLIC alone (all with prior PNGase F treatment). The glycosylation percentage was calculated using the number of unique glycosylated products divided by the sum of all unique glycosylated and unmodified peptides. (B) Venn diagram of overlapping plasma N-glycoproteins obtained using PUC-ERLIC, PUC alone, or ERLIC alone.

5.4.5 Evaluation of false positive N- glycosylation sites

Using an overall confidence level of $\geq$ 99%, FDR $\leq$ 1%, , queries against Mascot search engine returned a total of 599 unique N-glycopeptides and 361 unique N-glycosylation sites assigned to 127 distinct glycoproteins in crude human plasma using PUC -ERLIC enrichment strategy. All pertinent MS/MS spectrum of all glycosylated peptides were manually checked and can be viewed in supplemental information 1 from Ref.(168). We found 94 - 98% of the total detected N-glycosylated sites PUC-ERLIC matching to the conserved consensus motif N-X-S/T, with N-X-T motif being more common than the N-X-S motif. Minor proportion (2 - 6%) of N-glycosites matched the rare N-X-C motif.

Chapter 5

N-linked glycosylation is characterized by the addition carbohydrate moiety to the consensus N-X-Serine(S)/Threonine(T)/Cysteine(C) tri-peptide sequence motif, where X can be any non-proline amino acid. This conserved consensus N-X-S/T/C motif representing N-glycosylation is used as a criterion for the assignment of N-glycopeptides. PNGase F treatment is most commonly used for the deglycosylation of N-glycopeptides via deamidation of N to D (+0.984 Da) at the site of glycan attachment. The mass increment of the deglycosylated peptides can be monitored in MS/MS. In N-linked glycoproteomic experiments, the accurate assignment of N-glycosylation sites may be hampered by the use of alkaline reaction buffers that encourages non-enzymatic deamidation of N to aspartic acid (D) (306), mimicking PNGase F deamidation. Regrettably, in MS/MS PNGase F-mediated deamidation is not distinguishable from spontaneous or chemically-induced Asn deamidation (414). The use of acidic digestion and deglycosylation treatment at corresponding pH 6 and pH 5, have been shown to minimize chemically-induced deamidation (306, 307).

The false-positive assignment of N-glycosylation sites was assessed in this study by the independent PUC-ERLIC reference experiment performed without PNGase F treatment. We identified four falsely assigned N-glycosylated peptides (Table 5.1) with consensus N-X-S/T/C motif, amounting to 0.68% FDR of total N-glycopeptides in the present study, these peptides were disregarded and not used for downstream analyses. In addition, 252 (12%) unique deamidated peptides out of 2171 total unique peptides that were not within the N-linked sequon were also detected in plasma without PNGase F treatment, a majority of which contains the common deamidation N-G or N-S motif (415, 416) at the +1 position, presumably these peptides were deamidated in vivo. Therefore, given the measures taken to minimize chemically-induced deamidation and the low observed FDR in false assignment of N-glycopeptides, we consider that set of N-glycosylation site assignments obtained from PUC-ERLIC were reliably based on PNGase F-mediated deamidation.

Table 5.1. Proteins and peptides identified in human plasma with false glycosylation assignment.

Accession	Protein Description	Peptide Sequence	Peptide Score	N-X-S/T/C Position
P01616	Ig kappa chain V-II region MIL	RFSGSGSGT**NFT**LKI	95.68	N74
A6NNI4	CD9 antigen	KAIHYAL**NCC**GLAGGVEQFISDICPKK	41.73	N82
U3KQ34	Interleukin-1 receptor-associated kinase 1-binding protein 1	KMQ**NIC**NFLVEKL	50.55	N53
P08603	Complement factor H	KSPDVI**NGS**PISQKI	57.14	N217

Deamidated sites that match the consensus N-X-S/T/C glycosylation motifs are underlined and bold.

5.4.6 Enrichment of low-abundant plasma glycoproteins and EVs by PUC-ERLIC

Plasma N-glycosylated proteins identified using PUC-ERLIC were compared against published datasets obtained from previous glycoproteomics studies that employed alternative selective enrichment methods on similar LC-MS/MS platform. With the use of solid phase extraction of N-linked glycoproteins (SPEG), Yang et al. reported the identification of 462 glycopeptides assigned to 185 N-glycoproteins in crude human plasma. (417). Drake et al reported the identification of 227 N-glycosylation sites covering 119 N-glycoproteins from MARS-depleted plasma using Sambucus nigra agglutinin (SNA) and Aleuria aurantia lectin (AAL) affinity capture workflows (418). Displayed in a venn diagram in Figure 5.5, of 229 the N-glycosylated proteins identified 130 (~57%) glycosylated proteins were found to be in common among the three studies (the full comparison analyses can be found in supplemental data S4, worksheet CompareStudy in Ref. (168) in). PUC-ERLIC accounted for at least 48% overlapped of glycosylated proteins, indicating that PUC-ERLIC glycoprotein enrichment method is at least as efficient as established methods such as SPEG and lectin affinity workflows. Of importance, PUC-ERLIC is the only method that enables parallel enrichment of soluble glycoproteins and EVs within a single experiment.

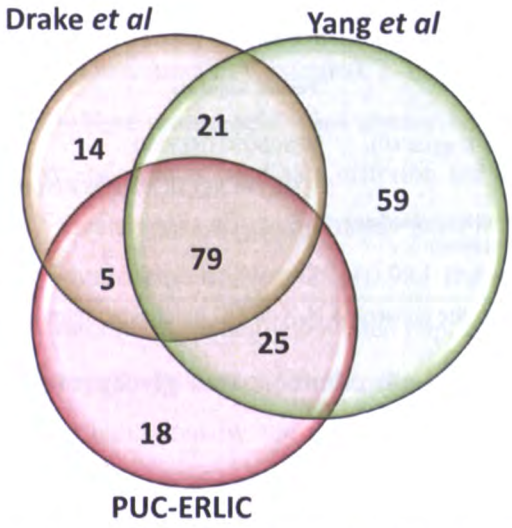

Figure 5.5. Venn diagram of three glycoproteomics studies. Venn diagram showing the N-glycosylated plasma proteins identified using PUC-ERLIC (current study), SPEG approach (Yang et al (417)) or affinity capture method (Drake et al (418)).

We surveyed the list of plasma EVs studies consolidated in Exocarta (http://www.exocarta.org/), a public web-based repository for EVs (419), and determine the percentage of known EVs proteins identified within this study. As listed in Table 5.2, 58 out of 127 glycosylated protein have been cataloged in the human plasma reference studies in Exocarta (152, 420, 421), providing further confidence in PUC-ERLIC capability in the isolation of EV-derived glycoproteins, in addition to soluble glycoproteins. To our understanding, we are the first to report on vesicular glycoproteins isolation from human plasma, hence the undocumentation of 69 glycosylated proteins obtained from PUC-ERLIC in Exocarta was expected.

Next, we addressed the dynamic range of detection within this study by conducting literature surveys to gauge the approximate plasma concentration of each glycoproteins identified. In plasma, albumin alone constitutes more than 50% of plasma and the 12 most abundant components accounts to about 95% of the total protein content (422, 423). Therefore, it is no surprise that 62% of glycosylated proteins identified in this study were among the top 150 high abundant plasma components, having concentrations ranging from µg/mL - mg/mL in plasma (Table 5.3). Of interest, the detection of 48 low abundant proteins present in the pg/mL –

ng/mL concentration range were not hampered by the co-existence of high abundant proteins. Among the low-abundant proteins identified, vascular endothelial growth factor receptor 3 (FLT4) have concentration reported to at 4.1 ng/mL (424), multiple epidermal growth factor-like domains protein 8 (MEGF8) was estimated to have 4.3 pg/mL concentration in plasma (424) and cysteine-rich secretory protein 3 (CRISP3) have plasma concentration of 6.3 pg/mL (425).

Table 5.2. Extracellular vesicle-associated glycoproteins as cataloged in Exocarta.

No.	UniProt ID	Gene Symbol	Protein Score	Protein Description	Extracellular vesicle proteins	Ref
1	P04217	A1BG	5598	Alpha-1B-glycoprotein	Y	(420)
2	P01023	A2M	8107	Alpha-2-macroglobulin	Y	(420)
3	P43652	AFM	3787	Afamin	Y	(420)
4	C9JV77	AHSG	81282	Alpha-2-HS-glycoprotein	Y	(420)
5	P02768	ALB	285547	Serum albumin	Y	(420)
6	P02760	AMBP	11966	Protein AMBP	Y	(420)
7	P04114	APOB	1640	Apolipoprotein B-100	Y	(421)
8	C9JF17	APOD	5406	Apolipoprotein D	Y	(421)
9	F8WF14	BCHE	901	Cholinesterase	Y	(420)
10	H0YFH3	C1R	1036	Complement C1r subcomponent	Y	(420)
11	F8WCZ6	C1S	381	Complement C1s subcomponent	Y	(420)
12	P01024	C3	6503	Complement C3	Y	(420)
13	P0C0L5	C4B	52520	Complement C4-B	Y	(420)
14	P04003	C4BPA	2570	C4b-binding protein alpha chain	Y	(420)
15	Q5VVQ7	C4BPB	5187	C4b-binding protein beta chain	Y	(421)
16	P07357	C8A	1189	Complement component C8 alpha chain	Y	(420)
17	P02748	C9	7324	Complement component C9	Y	(420)
18	B4E1Z4	CFB	6621	Complement factor B	Y	(420)
19	P08603	CFH	35443	Complement factor H	Y	(420)
20	Q03591	CFHR1	2208	Complement factor H-related protein 1	Y	(420)
21	Q02985	CFHR3	791	Complement factor H-related protein 3	Y	(420)
22	P10909	CLU	8364	Clusterin	Y	(421)
23	P00450	CP	50977	Ceruloplasmin	Y	(420)
24	P22792	CPN2	417	Carboxypeptidase N subunit 2	Y	(420)
25	E9PIT3	F2	2800	Thrombin light chain	Y	(420)
26	Q9Y6R7	FCGBP	980	IgGFc-binding protein	Y	(421)
27	P02671	FGA	1132	Fibrinogen alpha chain	Y	(420)
28	P02675	FGB	14861	Fibrinogen beta chain	Y	(420)
29	C9JC84	FGG	10296	Fibrinogen gamma chain	Y	(420)
30	F8W7G7	FN1	259	Fibronectin	Y	(420)
31	P00738	HP	166762	Haptoglobin	Y	(420)
32	P00739	HPR	136434	Haptoglobin-related protein	Y	(420)
33	P02790	HPX	32600	Hemopexin	Y	(420)
34	P04196	HRG	17014	Histidine-rich glycoprotein	Y	(420)
35	P01591	IGJ	1829	Immunoglobulin J chain	Y	(421)
36	H7C4N8	ITGB1	105	Integrin beta-1	Y	(420)
37	F5H7E1	ITIH1	1655	Inter-alpha-trypsin inhibitor heavy chain H1	Y	(420)
38	Q5T985	ITIH2	15178	Inter-alpha-trypsin inhibitor heavy chain H2	Y	(420)
39	Q14624-2	ITIH4	5879	Isoform 2 of Inter-alpha-trypsin inhibitor heavy chain H4	Y	(420)
40	P01042-2	KNG1	47217	Isoform LMW of Kininogen-1	Y	(420)
41	H0YCG2	LAMP2	242	Lysosome-associated membrane glycoprotein 2	Y	(152)
42	Q08380	LGALS3BP	3599	Galectin-3-binding protein	Y	(420)
43	Q07954	LRP1	562	Prolow-density lipoprotein receptor-related protein 1	Y	(421)
44	F8W876	MASP1	78	Mannan-binding lectin serine protease 1 heavy chain	Y	(421)
45	O00187	MASP2	85.18	Mannan-binding lectin serine protease 2	Y	(421)
46	Q13201	MMRN1	272	Multimerin-1	Y	(421)
47	P02763	ORM1	3105	Alpha-1-acid glycoprotein 1	Y	(420)
48	P19652	ORM2	968	Alpha-1-acid glycoprotein 2	Y	(420)
49	P01833	PIGR	81	Polymeric immunoglobulin receptor	Y	(421)
50	P27169	PON1	15059	Serum paraoxonase/arylesterase 1	Y	(420)
51	P01009	SERPINA1	28362	Alpha-1-antitrypsin	Y	(421)
52	G3V5I3	SERPINA3	16582	Alpha-1-antichymotrypsin	Y	(420)
53	P01008	SERPINC1	16622	Antithrombin-III	Y	(420)
54	B4E1F0	SERPING1	20410	Plasma protease C1 inhibitor	Y	(420)
55	P02787	TF	27433	Serotransferrin	Y	(420)
56	P07996	THBS1	87	Thrombospondin-1	Y	(421)
57	P04004	VTN	18138	Vitronectin	Y	(421)
58	P04275	VWF	281	Von Willebrand factor	Y	(421)

Evidence of cataloged extracellular vesicle proteins obtained from listed references.

Table 5.3. Estimated concentrations of the identified plasma glycoproteins.

No	UniProt Accession No.	No. Unique Peptide(s)	Protein Description	Top 150 plasma proteins [a]	Conc. in plasma (mg/mL)[b]	Ref
1	P04217	3	Alpha-1B-glycoprotein	Y	0.133	(424)
2	P01023	6	Alpha-2-macroglobulin	Y	1.609	(424)
3	P43652	4	Afamin	Y	0.035	(424)
4	C9JV77	10	Alpha-2-HS-glycoprotein	Y	0.082	(426)
5	P02768	2	Serum albumin	Y	33.3	(424)
6	P02760	19	Protein AMBP	Y	0.048	(426)
7	K7EKF6	2	Angiopoietin-related protein 6	N	*NA	*NA
8	H0Y512	2	Adipocyte plasma membrane-associated protein	N	0.000029	(426)
9	P04114	5	Apolipoprotein B-100	Y	0.2989	(424)
10	K7EMC3	1	Apolipoprotein C-IV	Y	4.6	(426)
11	C9JF17	9	Apolipoprotein D	Y	0.082	(426)
12	P02749	12	Beta-2-glycoprotein 1	Y	0.0725	(424)
13	Q5SRP5	3	Apolipoprotein M	Y	0.0015	(426)
14	O75882	9	Attractin	Y	0.00066	(426)
15	F8WF14	10	Cholinesterase	Y	0.00017	(426)
16	B4DLJ9	7	Biotinidase	Y	0.00049	(426)
17	P02745	3	Complement C1q subcomponent subunit A	Y	0.0012	(426)
18	H0YFH3	3	Complement C1r subcomponent	Y	0.0043	(426)
19	Q9NZP8	2	Complement C1r subcomponent-like protein	N	0.00026	(426)
20	F8WCZ6	2	Complement C1s subcomponent	Y	0.0052	(426)
21	P01024	1	Complement C3	Y	1.73	(425)
22	P0C0L5	3	Complement C4-B	Y	0.104	(424)
23	P04003	12	C4b-binding protein alpha chain	Y	0.011	(426)
24	Q5VVQ7	8	C4b-binding protein beta chain	Y	0.00053	(426)
25	P13671	1	Complement component C6	Y	0.0037	(426)
26	P07357	4	Complement component C8 alpha chain	Y	0.0024	(426)
27	F5GY80	1	Complement component C8 beta chain	Y	0.0028	(426)
28	P02748	3	Complement component C9	Y	0.01393	(424)
29	H0Y2P0	1	CD44 antigen	N	0.000035	(426)
30	P33151	1	Cadherin-5	N	0.00011	(426)
31	E9PPW2	1	Cadherin-related family member 5	N	0.0000044	(426)
32	B4E1Z4	16	Complement factor B	Y	0.00126	(424)
33	P08603	25	Complement factor H	Y	0.0287	(424)
34	Q03591	2	Complement factor H-related protein 1	Y	0.057	(426)
35	Q02985	1	Complement factor H-related protein 3	Y	0.00014	(426)
36	B1ALQ8	1	Complement factor H-related protein 4	Y	0.000059	(426)
37	Q5VYL6	1	Complement factor H-related protein 5	Y	0.000011	(426)
38	G3XAM2	7	Complement factor I light chain	Y	0.0000057	(426)
39	P10909	9	Clusterin	Y	0.0914	(424)
40	J3KRP0	2	Beta-Ala-His dipeptidase	Y	0.00023	(426)
41	P00450	12	Ceruloplasmin	Y	0.0649	(424)
42	Q96IY4	4	Carboxypeptidase B2	Y	0.0007	(426)
43	P22792	3	Carboxypeptidase N subunit 2	Y	0.002	(426)

*NA, data not available; a, Top 150 plasma protein annotation obtained from Hortin et al(422); b, Approximate plasma concentration in mg/mL (obtained from listed references).

Table 5.3. Estimated concentrations of the identified plasma glycoproteins. (Continued)

No	UniProt Accession No.	No. Unique Peptide(s)	Protein Description	Top 150 plasma proteins [a]	Conc. in plasma (mg/mL)[b]	Ref
44	J3KPA1	1	Cysteine-rich secretory protein 3	N	6.30E-09	(427)
45	Q5T382	2	Dopamine beta-hydroxylase	N	0.00011	(426)
46	P03951	5	Coagulation factor XI	N	0.000063	(426)
47	P00748	2	Coagulation factor XII	N	0.0143	(425)
48	P05160	1	Coagulation factor XIII B chain	Y	0.00096	(426)
49	E9PIT3	6	Thrombin light chain	N	0.00941	(424)
50	P12259	5	Coagulation factor V	Y	0.006417	(425)
51	B1AHL2	3	Fibulin-1	Y	0.00062	(426)
52	Q9Y6R7	3	IgGFc-binding protein	N	0.0003	(426)
53	P02671	1	Fibrinogen alpha chain	Y	1.5	(424)
54	P02675	7	Fibrinogen beta chain	Y	0.706	(424)
55	C9JC84	8	Fibrinogen gamma chain	Y	0.0233	(424)
56	E9PD35	1	Vascular endothelial growth factor receptor 3	N	0.0000041	(426)
57	F8W7G7	2	Fibronectin	Y	0.02	(426)
58	Q8NBJ4	1	Golgi membrane protein 1	N	0.0000092	(426)
59	P80108	8	Phosphatidylinositol-glycan-specific phospholipase D	Y	0.00046	(426)
60	Q8IZF2	1	Probable G-protein coupled receptor 116	N	0.000012	(426)
61	Q9ULI3	1	Protein HEG homolog 1	N	0.0000066	(426)
62	D6RAR4	1	Hepatocyte growth factor activator	N	0.00024	(426)
63	P00738	69	Haptoglobin	Y	1.1	(424)
64	P00739	1	Haptoglobin-related protein	Y	0.0407	(428)
65	P02790	8	Hemopexin	Y	0.4762	(424)
66	P04196	2	Histidine-rich glycoprotein	Y	0.035	(426)
67	P05362	2	Intercellular adhesion molecule 1	N	0.000071	(426)
68	J3QQX6	1	Intercellular adhesion molecule 2	N	0.000037	(426)
69	P35858	2	Insulin-like growth factor-binding protein complex acid labile subunit	N	0.0015	(426)
70	B3KWK7	3	Insulin-like growth factor binding protein 3, isoform CRA	Y	0.0057	(426)
71	P01876	6	Ig alpha-1 chain C region	Y	*NA	*NA
72	P01877	1	Ig alpha-2 chain C region	Y	*NA	*NA
73	P01857	1	Ig gamma-1 chain C region	Y	*NA	*NA
74	P01859	1	Ig gamma-2 chain C region	Y	*NA	*NA
75	P01860	1	Ig gamma-3 chain C region	Y	*NA	*NA
76	P01861	1	Ig gamma-4 chain C region	Y	0.32	(426)
77	P01871	9	Ig mu chain C region	Y	0.32	(426)
78	P01591	3	Immunoglobulin J chain	Y	0.0056	(426)
79	H7C4N8	1	Integrin beta-1	N	0.000011	(426)
80	F5H7E1	6	Inter-alpha-trypsin inhibitor heavy chain H1	Y	0.024	(426)
81	Q5T985	3	Inter-alpha-trypsin inhibitor heavy chain H2	Y	0.021	(426)
82	Q06033	2	Inter-alpha-trypsin inhibitor heavy chain H3	Y	0.002	(426)
83	Q14624-2	6	Isoform 2 of Inter-alpha-trypsin inhibitor heavy chain H4	Y	0.042	(426)
84	H0YAC1	8	Plasma kallikrein heavy chain	Y	0.0026	(426)

*NA, data not available; a, Top 150 plasma protein annotation obtained from Hortin et al(422); b, Approximate plasma concentration in mg/mL (obtained from listed references).

Table 5.3. Estimated concentrations of the identified plasma glycoproteins. (Continued)

No	UniProt Accession No.	No. Unique Peptide(s)	Protein Description	Top 150 plasma proteins [a]	Conc. in plasma (mg/mL)[b]	Ref
85	P01042-2	15	Isoform LMW of Kininogen-1	Y	1.085	(424)
86	H0YCG2	1	Lysosome-associated membrane glycoprotein 2	N	0.0000073	(426)
87	P04180	1	Phosphatidylcholine-sterol acyltransferase	N	0.00022	(426)
88	Q08380	4	Galectin-3-binding protein	N	0.00044	(426)
89	Q16609	1	Putative apolipoprotein(a)-like protein 2	Y	0.0886	(424)
90	P02750	4	Leucine-rich alpha-2-glycoprotein	Y	0.0027	(426)
91	Q07954	1	Prolow-density lipoprotein receptor-related protein 1	N	0.00002	(426)
92	P51884	5	Lumican	Y	0.004	(426)
93	Q9Y5Y7	1	Lymphatic vessel endothelial hyaluronic acid receptor 1	N	0.0000062	(426)
94	F8W876	2	Mannan-binding lectin serine protease 1 heavy chain	N	0.00024	(426)
95	O00187	1	Mannan-binding lectin serine protease 2	N	0.0028	(426)
96	Q7Z7M0	1	Multiple epidermal growth factor-like domains protein 8	N	4.30E-09	(426)
97	Q13201	3	Multimerin-1	N	0.0000072	(426)
98	G3XAK1	3	Hepatocyte growth factor-like protein alpha chain	N	0.00025	(426)
99	F8W6L6	1	Myosin-10	N	*NA	*NA
100	P02763	8	Alpha-1-acid glycoprotein 1	Y	0.000383	(424)
101	P19652	3	Alpha-1-acid glycoprotein 2	Y	0.22	(426)
102	Q96PD5	11	N-acetylmuramoyl-L-alanine amidase	N	0.014	(426)
103	E7ER40	1	PHD finger protein 3	N	*NA	*NA
104	Q6UXB8	1	Peptidase inhibitor 16	N	0.000041	(426)
105	P01833	2	Polymeric immunoglobulin receptor	N	0.000025	(426)
106	B3KUE5	3	Phospholipid transfer protein	N	0.000064	(426)
107	P27169	3	Serum paraoxonase/arylesterase 1	Y	0.0077	(426)
108	Q15166	1	Serum paraoxonase/lactonase 3	Y	0.000011	(426)
109	H0Y5A1	2	Prostaglandin-H2 D-isomerase	N	0.012	(426)
110	P35542	1	Serum amyloid A-4 protein	Y	0.03	(426)
111	P14151	3	L-selectin	N	0.000936	(424)
112	D6REX5	9	Selenoprotein P	Y	0.00051	(426)
113	P01009	6	Alpha-1-antitrypsin	Y	0.1374	(424)
114	G3V5I3	12	Alpha-1-antichymotrypsin	Y	0.11	(426)
115	P29622	3	Kallistatin	Y	0.0011	(426)
116	G3V350	2	Corticosteroid-binding globulin	Y	0.0012	(426)
117	P05543	2	Thyroxine-binding globulin	Y	0.0013	(426)
118	P01008	9	Antithrombin-III	Y	0.058	(425)
119	P05546	8	Heparin cofactor 2	Y	0.044	(425)
120	I3L4N7	1	Pigment epithelium-derived factor	Y	0.00736	(424)
121	B4E1F0	8	Plasma protease C1 inhibitor	Y	0.012	(426)
122	B0FWH6	2	Sex hormone-binding globulin	Y	0.00026	(426)
123	B5MCY1	1	Tudor domain-containing protein 15	N	*NA	*NA
124	P02787	13	Serotransferrin	Y	1.5	(424)
125	P07996	2	Thrombospondin-1	N	0.00051	(426)
126	P04004	6	Vitronectin	Y	0.0197	(424)
127	P04275	3	Von Willebrand factor	Y	0.007698	(424)

*NA, data not available; a, Top 150 plasma protein annotation obtained from Hortin et al(422); b, Approximate plasma concentration in mg/mL (obtained from listed references).

In order to increase the reliability of our dataset, strict criteria of peptide FDR ≤ 1% and narrow precursor ion mass tolerance of 5 ppm was applied to eliminates ambiguous identification and miss-assignment of C13 peak of the native peptide as deamidated peptide (363, 364). In total, 137 unique glycosylated peptides were found assigned to the 48 low abundant glycoproteins. The MS/MS spectra and fragment ion assignment of 137 statistically matched ($p < 0.05$) low abundant peptides were manually curated and checked. In addition, most PNGase F-mediated deamidated peptides including peptide ALLTNVSSVALGSR peptide assigned to MEGF8, peptide LVIQNANVSAMYK assigned to FLT4 and peptide DSCKASCNCSNSIY assigned to CRISP3 were confidently identified with high peptide mascot score and low expect values, indicating the unambiguous identification of homologous protein. Furthermore, all low abundant glycosylated peptides (D residue) and their native non-modified counterparts (N residue) were submitted to NCBI BLASTP (222, 223) for complementary sequence hormonology search. Consistency between BLASTP protein homologues hits and Mascot protein hits were observed, > 90% sequence identity among species with no peptide sequence habouring homology to another protein in NCBI were found.

Our enrichment method has efficiently enabled the detection of numerous low abundant plasma glycoproteins obtained from heart disease patients, which is of great importance in the discovery of candidate biomarkers of human pathologies. With biological relevance to heart disease, we detected lymphatic vessel endothelial hyaluronic acid receptor 1 (LYVE1), a lymphangiogenic glycoprotein factor found on lymphatic endothelial cells have been reported to be significantly overexpressed and involved in myocardium remodeling in the peri-infarction areas following myocardial infarction (429, 430). The identification of cardiac-specific remodeling glycoproteins LYVE1 suggest its potential as a biomarkers indicative for myocardium healing. Taken together, PUC-ERLIC is an efficient method that simplifies plasma complexity, enabling dynamic range detection of greater than 10^8 in crude plasma, offering potential to uncover novel glycoproteomic biomarkers of

diseases. In addition to having good sample preparation procedures as the first step for overcoming dynamic range issues in plasma, it is also mandatory to have well maintained sensitive detection analytical system in order to increase the depth of glycoproteomics analysis in complex biological samples.

5.4.7 Classification of glycosylated proteins derived from PUC-ERLIC

The secretory and EV-derived glycoproteins identified with PUC-ERLIC approach were annotated by Gene Ontology (GO) for an overview of the preferential molecular functions, biological processes and subcellular localization that glycoproteins are associated with. As shown in Figure 5.6A, GO annotated molecular function indicates that binding activity (50.6%) was particularly overrepresented, followed by catalytic activity (18.3%) and enzyme regulator (15.2%), of note, these annotation molecular functions are known molecular roles of glycosylated proteins. GO biological functions categorization (Figure 5.6B) associates regulation (28.0%), cellular processes (16%) and metabolic processes or responses to stimuli (12%) as the principal biological functions of glycosylated proteins. Immune system processes, localization and interaction with cells and organisms were represent in equal proportion (9%). Other annotated biological functions such as developmental process (6%) and growth (1%) were also observed.

Consistent with the topological locations glycosyltransferase in the endoplasmic reticulum (ER) and golgi apparatus, N-glycosylated proteins are thought to predominate as secreted proteins and extracellular membrane proteins (162). GO subcellular localization annotation predicts that the majority of identified glycoproteins (69.4%) were mainly extracellular proteins and plasma membrane proteins as shown in Figure 5.6C. The extraction of membrane proteins, which constitute up to 20-30% of the total proteome, has been challenging due to its low abundance and inherent hydrophobicity (431). In our results, substantial recovery (20.3%) glycosylated plasma membrane proteins was achieved by PUC-ERLIC approach. In addition, cadherin-related family member 5 (CDHR5) and lysosome-associated membrane glycoprotein 2 (LAMP2) were detected as EV-specific

proteins (GO:0070062). However, 9.9% of glycosylated proteins in our dataset were annotated in atypical locations including cytoplasmic, nuclear or mitochondrial compartments; further investigation revealed these annotations to be nonexclusive.

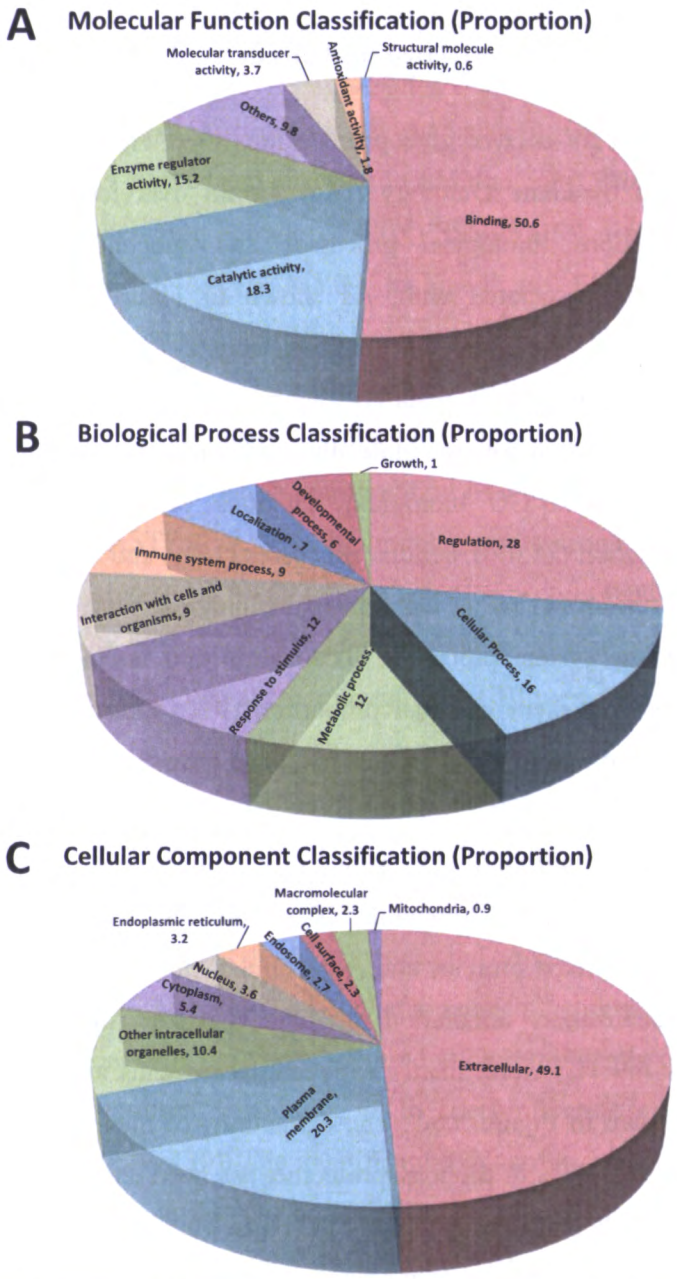

Figure 5.6. Gene Ontology (GO) functional annotation of glycoproteins. GO-based functional annotation showing the proportionate distribution of identified glycoproteins according to their (A) molecular functions, (B) biological processes and (C) cellular localization. Predictions were obtained using STRAP v1.5(409), AmiGO v2(340, 410) and DAVID v6.7(411, 432).

Chapter 5

5.5 Conclusion

In order to harvest a broader spectrum of candidate biomarkers for use in mass spectrometry-based proteomic profiling, we have developed an enrichment method that couples PUC with electrostatic repulsion-hydrophilic interaction chromatography ERLIC [PUC-ERLIC] to facilitate the simultaneous recovery of both glycoproteins and EV from crude plasma. We described the first visual recovery of a yellow EV-containing fraction that was found to be highly concentrated in glycoproteins that could probed as prognostic/diagnostic biomarkers. Comparative analyses with published studies that utilized SPEG and lectin affinity workflows for selective enrichment of N-glycosylated proteins showed that PUC-ERLIC is as efficient as established method for glycoprotein enrichment. Through this novel systematic PUC-ERLIC approach, we reported the confident (FDR $\leq$ 1%) identification of 127 secretory and EV-derived glycoproteins, 599 unique glycopeptides and 361 unique glycosylation sites in a single LC-MS/MS run. Without the use of immunodepeletion strategy, 48 low-abundant glycoproteins with concentration ranging from pg/mL – ng/mL were detected in crude plasma, overcoming the complexity and technical challenges associated with delineating the protein composition of human plasma. This method can be readily applied to any type of biofluids and assist in the detection of novel souble/EV-derived biomarkers in a wide variety of human diseases. This study is thus expected to benefit similar future plasma-based EV glycoproteomics studies.

Chapter 6

Conclusion and Future Work

Chapter 6

6.1 Conclusion

Alongside with global population ageing, the burden of CAD will remain significant. Effective screening and diagnostic methods that stratify high risk patients at risk for CAD, prior to any clinical event are thereby urgently needed. The identification of efficacious plasma biomarkers that are sensitive and specific for CAD, remain as one of the greatest challenges in modern medicine. Gladly, biomarker discovery has taken a leap forward with the advent and development of MS-based proteomics approaches, enabling the interrogation of the proteome and sub-proteomes of plasma. At present, the incomplete understanding in the mechanisms and triggers for plaque destabilization and eventual rupture, creates a significant knowledge gap in the pathophysiological concept underlying atherosclerosis-associated stable angina and the sudden onset of MI. In keeping with this, currently there is no plasma-based screening assays that effectively discriminates underlying atherosclerotic conditions from impending myocardial injury. We recognized that the development of innovative techniques that improves the detectable range of plasma proteome, are essential in advancing the field of plasma proteomics-based biomarker discovery research. In order the address the aforementioned issues, of the four parts (I - IV) (Figure 1.3) in my PhD research work, three studies showcased the clinical relevance and maximal utility of the plasma proteome and sub-proteome for CAD proteomics-based biomarker discovery studies, and the successful development of a novel technique for the simultaneous enrichment of glycoproteins and EV from plasma was demonstrated in the final study.

In chapter 2, we developed a multi-marker panel consisting of eight novel proteins that reliably predicts and discriminates atherosclerosis from myocardial injury. In the discovery phase, we analyzed pooled plasma obtained from MI, NMI and Ctrl samples by iTRAQ-based LC–MS/MS. Of the 371 proteins quantified, 27 proteins that displayed significant differential expression ($\geq$ 2-fold change, $p < 0.05$) in 114_{MI}:113_{Ctrl} and 115_{NMI}:113_{Ctrl} were selected as atherosclerotic-specific markers, and 26 proteins with at least two-ratio fold change in 114_{MI}:115_{NMI} were flagged as myocardial injury-specific markers. GO analysis predicted some subsets of

biological processes and pathways distinctive to either atherosclerotic-specific markers or myocardial injury-specific markers. In the verification phase, label-free LC-MRM-MS was used to screen 53 candidate biomarkers in pooled plasma. Based on stringent selection criteria, 59 peptides representing eight atherosclerotic-specific markers, 15 myocardial injury-specific markers and three house-keeping proteins were assembled and quantified in 49 individual plasma samples. Label-free MRM quantification verification analyses identified 12 out of 23 protein candidates with significant ($p < 0.05$) differential abundance and in agreement with discovery proteomics experiments. Among the 12 candidate markers, two (F10, MST1) atherosclerotic-specific markers and six (ORM2, SERPINA3, CPN2, LUM, ACTG1, NAGLU) myocardial injury-specific markers were found to be novel. We suggest that pathogenesis of atherosclerosis and/or myocardial tissue injury may be accelerated by elevated levels of F10, SERPINA3, CPN2 and decreased levels of LUM. Whilst atheroprotective and/or of tissue repair activities may be induced with the increase abundance in MST1, ORM2, ACTG1 and decreased expression of NAGLU. Functional pathological linkage of these proteins to CAD, have suggested their involvement in endothelial dysfunction (F10, MST1), plaque destabilization (SERPINA3, CPN2, LUM) and protective mechanisms against further damage (ORM2, ACTG1, NAGLU), which present as protein signatures poised to differentiate underlying atherosclerotic conditions from myocardial injury. We have showed that capability of iTRAQ-label-free MRM based approach in the discovery and verification of novel candidate biomarkers of atherosclerosis and myocardial injury in undepleted plasma. The diagnostic values and possible clinical utility of these candidate proteins warrants further assessment in a valid outcome study.

The enrichment of plasma sub-proteomes are commonly pursued as alternative method for the search of biomarkers within the plasma, through the reduction of proteome complexity. EVs are attractive targets with numerous potential applications in clinical utility, at present proteomics studies on circulatory EVs in MI-related research are limited and their involvement in myocardial ischemic injury has yet to be determined. In Chapter 3, we performed a global ion intensity label-

free differential proteome analyses of plasma EVs obtained from MI and NMI patients. The confidence and reliability of our quantitative dataset was demonstrated by the consistency in the number of identification (FDR <1%) on protein, peptide and spectral levels and by good correlation of extracted ion intensities between biological replicates. Functional comparative proteomics profiling of perturbed EV proteome of MI with NMI were represented for the first time, to the best of our knowledge. The categorization of patients presented with CAD into MI and NMI increases the probability of detecting myocardial injury specific markers that defines plaque rupture. We have identified a panel of six potential EV markers of myocardial injury from three pathways, namely complement activation (C1Q1A, ~3.23 fold-change, $p = 0.012$; C5, ~1.27 fold-change, $p = 0.087$), lipid metabolism (APOD, ~1.86-fold change, $p = 0.033$; APOC3, ~2.63-fold change, $p = 0.029$) and platelet activation (GP1BA, ~9.18-fold change, $p < 0.0001$; PPBP, ~4.72-fold change, $p = 0.027$) that showed significant ($p < 0.05$) association with post-infarction response. Altered increased expression of EV-derived C1Q1A, C5, APOC3, GP1BA and PPBP were thought to aggravate post-infarction inflammation and myocardial tissue damage. On contrary, elevated levels of APOD might confer properties that promote plaque stability and atheroprotection. Our differential proteome data have showed EVs mediation in myocardial damage and myocardial healing processes following MI. In addition, we are the first to report the paradoxical decreased levels of EV-derived fibrinogen after MI, we infer that this observation could possibly imply a compensatory attempt to suppress coagulation after MI. We recognized that massive complement activation is the main trigger to exaggerated inflammatory reaction after MI, that could evoke adverse post-infarcted remodeling and pathogenesis of heart failure. Independent validation of differentially expressed complement proteins C1QA ($p = 0.005$) and C5 ($p = 0.0021$) were performed using the Luminex-single plex assay and results were aligned with our quantitative proteomics dataset. Future larger studies will be important for confirmation of our data for enhancement of knowledge on the pathophysiological roles untaken by EVs following MI.

Due to the long-standing belief that significant occurrence of protein carbamylation can only be attributed during hyperuremic states. Majority of the important findings

in protein carbamylation studies to date, were mainly investigated on patients or on atherosclerotic animal models with existing uremic conditions. Current knowledge on the association of protein carbamylation in CAD, under normal renal function is limited, and their exploitation as biomarkers are relatively unexplored. In Chapter 4, label-free differential carbamylated peptide analyses were performed using plasma obtained from MI and Ctrl groups. In attempt to accurately characterize the carbamylated plasma proteome, precautions in sample preparation were taken to minimize experimental-induced carbamylation by urea. The overall analytical variability among the triplicate experiment in terms of CV was < 7%. Under normal renal function, carbamylated plasma proteins have accounted to approximately 4% of total plasma proteome each in both MI and Ctrl group. A total of 30 high confident (FDR ≤ 1%, $p < 0.05$) carbamylated peptides assigned to 15 proteins were found to be significantly elevated ($p < 0.05$) in the MI group, with at least 1.5-ratio fold change. Of the 7 novel carbamylated proteins identified, the prominent regulation of c-HP (K-LPECEADDGCPK*PPEIAHGYVEH SVR-Y, ~ 6.43-fold change, $p = 0.00011$), c-TF (R-SMGGK*EDLIWELLNQ*AQEHFGKDK-S, ~ 7.38-fold change, $p = 0.00033$) and c-ORM1 (K-TYM*LAFDVNDEK*N*WGLSVYAD KPETTK-E, ~ 4.46-fold change, $p < 0.0001$) have present them as potential inflammatory peptide markers of MI. We speculated that the loss of HP, TF and ORM1 function due to carbamylation, favors the pathogenesis of atherosclerosis, by induction of pro-inflammatory activities and promotion of foam cells formation. Beyond the effects of impaired renal functions, this is the first report that describes the total in vivo plasma protein carbamylation and assessment of differential quantitation on carbamylated peptides. The new findings in this study should direct future molecular investigations aimed at unravelling roles carbamylated peptides of c-HP, c-ORM1 and c-TF in atherogenesis.

The involvement of plasma glycoproteins and EVs in a myriad of human pathologies have presented them as highly sought after clinical biomarkers. In order to harvest a broader spectrum of candidate biomarkers for use in mass spectrometry-based proteomic profiling, in Chapter 5, we developed a novel systematic PUC-ERLIC approach that affords simultaneous enrichment of low abundant secretory and EV-

derived glycoproteins from crude plasma. We described the first visual recovery of a yellow EV-containing fraction that was found to be highly enriched in glycoproteins that could probed as prognostic/diagnostic biomarkers. The presence of EV isolation by PUC have been confirmed by cryo-EM, western blot as well as proteomics analyses. The high enrichment specificity and reliable in the assignment of N-glycosylation sites in this study were supported by the low FDR (0.68%) attributed by chemically-induced deamidation. Comparative analyses with published studies that utilized SPEG and lectin affinity workflows for selective enrichment of N-glycosylated proteins showed that PUC-ERLIC is as efficient as established method for glycoprotein enrichment. In a single LC-MS/MS run, we have confidently (FDR $\leq$ 1%). reported a total of 127 plasma-derived glycoproteins, 599 unique glycopeptides and 361 unique glycosylation sites. Without having to remove the high abundance plasma proteins, our novel enrichment strategy enabled the detection of 48 low-abundant glycoproteins with concentration ranging from pg/mL – ng/mL in crude plasma, overcoming the complexity and technical challenges associated with delineating the protein composition of human plasma. The identification of cardiac-specific remodeling glycoprotein, LYVE1 in plasma of heart disease patients demonstrated the feasibility of PUC-ERLIC in identification of identify potential biomarkers with specific relevance to human pathologies. The presented proteomics dataset not only mediates the paucity in our knowledge on EV-derived glycoproteomics, but it also demonstrates great potential as a method for candidate biomarker discovery.

With the availability of numerous shotgun-based quantitative proteomics approaches to date, the tools employed in the respective CAD biomarker discovery study described here were influenced by the following factors, (i) sample type, (ii) multiplexity of the experiment, (iii) availability of analytical platform and software, and (iv) cost of experiment. In this thesis, iTRAQ-labeled (Chapter 2) and label-free (Chapter 2, 3, 4) shotgun proteomics techniques have been applied successfully on different analytical platforms (In Chapter 2: TripleTOF® 5600 hybrid quadrupole time-of-flight MS and TSQ Vantage triple quadrupole MS; In Chapter 3: LTQ-FT Ultra linear ion trap MS; In Chapter 4: Q-Exactive™ hybrid quadrupole orbitrap MS)

for the unbiased characterization, and comparative expression study of the proteome, and the sub-proteome of human plasma, obtained from healthy and CAD-affected individuals.

In efforts to ensure that findings from each respective study are meaningful, stringent search criteria of a narrow precursor ion mass tolerance window of 5 ppm, and a strict FDR $\leq$ 1% were consistently implemented in all proteomics dataset, regardless of the proteomics platforms that were used, to essentially exclude false positive measured masses, enabling reasonable unambiguous identifications of high fidelity peptides (363, 364). In addition, the employment of optimized plasma processing workflows, and the stability and accuracy of the analytical system have also largely attributed to the presentation of high quality, and reliable MS quantitative information in each respective study. Consequently, the successful detection of aberrantly expressed proteins, that have been previously documented or established as biomarkers of CAD (e.g. Chapter 2: CRP, SAA1; Chapter 3: CRP, APOB, APOA2; Chapter 4: APOB, APOA1, ALB), have provided an additional support for MS data quality assurance. Following these abovementioned criteria, along with critical functional linkage analyses of aberrantly expressed proteins, we showed that all respective employed proteomics tools described in Chapters 2 – 4, possess the robustness and capability in discovering novel disease biomarkers that are predictive of CAD.

As different proteomics experimental methods were utilised for the quantitative assessment of the systemic plasma proteome, the low abundant plasma EVs and the carbamylated proteome respectively. It is therefore inappropriate to conduct a direct comparison based on profile of proteins, to determine which platform served best for the discovery of novel biomarkers. Additionally, we would like to emphasize that at present, there is no single technique that offers complete proteome coverage (433). Hence, comparison will be explored in relation to the weighted merits and limitation of each quantitative approach, in meeting the research objectives for each respective chapter.

Chapter 6

In chapter 2, direct sample preparation on undepleted plasma was performed for three experimental groupings (Ctl, MI and NMI), using iTRAQ-labeled based quantitative approach, and label-free MRM based approach for the discovery, and verification of eight novel plasma candidate biomarkers, that specifically discriminates atherosclerosis from myocardial injury. iTRAQ has been a popular and successful labelling technique used for the comparative analysis of proteins from a variety of clinical specimens, including plasma (86, 110-117). The advantage of iTRAQ-based quantification over label-free quantitation is that since labelling was performed at peptide level, several labeled peptides of the same protein under different conditions can be identified, and quantified multiple times for high confidence protein detection within a single analysis (101). Consequently, once labeled samples are mixed, all subsequent separation and MS/MS conditions are identical, resulting in low experimental variability, as shown in our results of % CV < 9% (Table 2.2). Labelling approaches have also been considered to be more accurate in protein abundance measurement, as it bypasses the problems of ion-suppressive effects arising from co-eluting peptides (94). With the availability of advanced data analysis nodes embedded in ProteinPilot™ software, analysis of iTRAQ quantitation data has been made simple and straightforward. However, the complexity of the labeling process, the variability arising from labeling efficiencies, and the cost of these labeling reagents, are common downsides of iTRAQ in achieving reproducible and reliable results (107, 108). Fortunately, these aforesaid concerns were largely circumvented in this study (Figure 2.3), as processing protocol were firmly established in the laboratory, and efficient trainings procedures were in placed to gain familiarity and competency of this technique.

It is well-documented that the direct analysis of crude plasma by MS poses technical limitation in the detection of low-abundant proteins (87). Regarding the depletion strategies available, there is often the trade-off between achieving higher coverage of plasma proteome, and the non-specific removal of potentially important albumin and/or solid support-bound proteins during the depletion process (434). Hence, to overcome the sensitivity concerns related to plasma, without any inadvertent protein losses, the combination of first-dimension offline ERLIC with second-dimension

online RPLC was implemented to reduce the complexity of plasma, where peptides are separated based on the differences in isoelectric point (pI) and polarity, respectively (402, 435, 436). The approximate plasma concentration of the 23 verified candidate biomarkers were assessed through literature surveys (437-439), and 8 proteins (ACTG, ATRN, F10, FLNA, LPA, MST1, NAGLU and TLN1) were estimated to be present in ng/mL, while the remaining 15 proteins were found to be present in μg/mL. With the application of appropriate tools, we did not suffer largely from limitations in detection of low abundant plasma proteins, as the dynamic range of nearly 9 orders of magnitude was achieved in this study, the concentration where clinically relevant biomarkers are expected (142).

Label-free targeted MRM-LC-MS targeted approach was used in the verification of 27 candidate markers of atherosclerosis, and 26 candidate markers of myocardial injury (Table 2.3). This cost-effective rational strategy possess the capacity and efficiency to screen and prioritize initial long candidate list in a single workflow, before significant resources is invested in the development of quantitative assays. Unlike ELISA, this method permits the analysis of every possible candidate proteins at low cost and high throughput. The re-evaluation on the relative abundance for each candidate protein in individual plasma samples, will determine whether the discovery-phase detected difference is real and disease-specific. However, the development of sensitive MRM assay for each targeted protein requires careful planning and operator's experience, and the manual construction and refinement of best performing peptides and transitions can be laborious and lengthy. In addition, data analysis and data handling on the huge quantitative datasets generated can be challenging and tedious. Nevertheless, with relevant experience, high quality and reliable quantitative information for each targeted protein were achieved with this approach (Table 2.4). Though with their limitations, we showed that the combination of iTRAQ and MRM proteomics are efficient, robust and convenient platforms that fulfils the requirement for the high confident identification (low FDR), and verification of numerous potentially useful plasma biomarkers on CAD cohorts. This systematic and high throughput strategy presented here is well orchestrated for unbiased multiplexed quantitative discovery, the prioritization of candidates with the

used of well- characterized technologies, will facilitate the development of clinically relevant plasma protein biomarker panel for CAD.

In chapter 3 and 4, an inexpensive label-free peak intensity quantitative approach, was employed for the differential comparative analysis of two different sample groups in each chapter. Label-free sample processing procedure is generally straightforward and hassle-free, unlike multistep labelling protocol. Peak intensity measurement has been shown to correlate linearly with a wide range of protein abundance in complex biological samples (133-136). However, in cases where thousands of peptides from complex sample are analysed within a LC-MS analysis, ion-suppression effects from co-eluting peptides could result in non-linear responses (440). Fortunately, suppression effects and the detection of low abundance proteins can be reduced by the reliance of multidimensional LC (101), such as ERLIC-RP-LC. In addition, lower complexity sub-proteomic plasma EV proteins (Chapter 3) and carbamylated modified proteins (Chapter 4), are better suited for analysis using this label-free quantitative approach. As data filtering and normalization processes are essential steps, to correct for bias and systemic variability across multiple samples and multiple datasets, for the extraction of accurate quantitative information, the complexity in data processing for label-free quantitation hence increases with the multiplexity of the experiment. Herein, data processing was manageable as only two samples were compared, and reasonably low inconsistency and good reproducibility in identification and quantitation across multiple LC-MS/MS datasets were attained in both Chapter 3 (Figure 3.3) and 4 (Figure 4.2), and reliable results were attained distinguish significant real differences from experimental variability ($p < 0.05$).

However, the verification and validation of plasma-sub proteome with immunoassay-based approach can be costly, time-consuming and laborious, as selective enrichment of the desired sub-proteome, which are present in low yield, have to be extracted from individual clinical samples. As a result, in view of the ease and practicability for the validation of EV-derived candidate biomarkers in Chapter 3, commercial EV isolation kit was used for the isolation of plasma EVs from 43 individual CAD patients. It is therefore more cost-saving for the initial verification

to be performed using targeted MRM-LC-MS approach, so that maximal number of candidate proteins can be rapidly assessed quantitatively and prioritized for further investigation. Although label-free techniques are generally considered inferior to labeled techniques, in terms of quantitation accuracy and inability for multiplexing, we showed that with careful consideration and techniques, label-free approaches can be reliable and less- rigorous quantitative tools for the discovery of candidate biomarkers, and the comparative analysis of lower abundance plasma sub proteome. It would have been ideal to be able to recommend a single technique that is most suited for the discovery of novel CAD biomarkers. However, based on the above data and discussion, it is clear that there is no perfect single method that is free from analytical problems associated with - global or targeted, absolute and relative protein quantitation. Demonstrated in my thesis, it is possible to obtained useful and reliable quantitation data by the careful matching of the biological questions with the appropriate quantitative proteomics approach.

Overall in my thesis, careful selection and systematic integration of several proteomics strategies have proven useful in the discovery of novel plasma biomarkers of CAD, derived from (i) undepleted plasma proteome, (ii) EV plasma proteome and (iii) carbamylated plasma proteome. These novel prioritized candidates warrant further statistical assessment in large scale verification, and clinical validation studies on larger patient cohort, for their potential diagnostic values. Hopefully, these promising biomarkers will possess high predictive values necessary for accurate patient risk stratification, and in near future, be established as diagnostic biomarkers. Depending on the availability of resources, immunoassay-based or targeted AQUA MRM-MS-based approach can be utitlised in pilot studies ($n > 200$ per group) verification on individual clinical samples, to determine the credentials of the candidate biomarkers for potential clinical utility, and improve the success rate of advancing the most clinically promising subset of candidates into clinical validation studies. The validation ($n > 2000$ per group) of highly credentialed candidates flagged during verification, should be performed using the gold standard immunoassay-based approach (e.g. ELISA), to achieve sensitive and accurate quantitative measurements. Through multivariate bioinformatics and receiver

operating characteristic (ROC) curve analyses of the validated datasets, there are good reasons to believe that the panel of biomarkers derived at the end will deliver a sensitivity, and specificity protein signature having predictive value to complement with existing screening tools, to improve the diagnosis, stratification and long-term prognosis of CAD.

The application of proteomics in biomarker studies is not limited to the identification and evaluation of biomarker candidates, we have generated new hypotheses into the pathophysiological states that specifies for atherosclerosis and myocardial injury. Future functional investigations are essential and important for the confirmation of the hypothesized pathophysiological roles untaken by candidate proteins in the progression of CAD. The development of the novel enrichment technique described in this thesis, can be readily applied to any type of biological fluid and contribute to the identification of novel soluble/EV-derived biomarkers in a wide variety of human diseases in addition to CAD, and is expected to benefit similar future plasma-based EV glycoproteomics studies. These studies will have meaningful impacts on the continued health and well-being of this ageing society. With any luck, the findings in this thesis would permit the development of effective therapeutic approaches aiming at preventing and treating major CV complications, and thereby reducing CV-related mortality, morbidity and associated socioeconomic cost. Regardless of the proteomics strategy that was used, we are optimistic that with rationalized and systematic workflow in which couples biomarker discovery and validation, protein diagnostic will eventually will find clinical utility in near future.

References

References

1. Abegunde, D. O., Mathers, C. D., Adam, T., Ortegon, M., and Strong, K. (2007) The burden and costs of chronic diseases in low-income and middle-income countries. *The Lancet* 370, 1929-1938

2. Mathers, C., Fat, D. M., and Boerma, J. T. (2008) *The global burden of disease: 2004 update*, World Health Organization

3. Vilahur, G., Badimon, J. J., Bugiardini, R., and Badimon, L. (2014) The burden of cardiovascular risk factors and coronary heart disease in Europe and worldwide. *European Heart Journal Supplements* 16, A7-A11

4. Yusuf, S., Reddy, S., Ôunpuu, S., and Anand, S. (2001) Global burden of cardiovascular diseases part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* 104, 2746-2753

5. Critchley, J. A., and Capewell, S. (2003) Mortality risk reduction associated with smoking cessation in patients with coronary heart disease: a systematic review. *Jama* 290, 86-97

6. Teo, K. K., Ounpuu, S., Hawken, S., Pandey, M., Valentin, V., Hunt, D., Diaz, R., Rashed, W., Freeman, R., and Jiang, L. (2006) Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *The Lancet* 368, 647-658

7. Stamler, J., Vaccaro, O., Neaton, J. D., and Wentworth, D. (1993) Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes care* 16, 434-444

8. Hu, F. B., Stampfer, M. J., Haffner, S. M., Solomon, C. G., Willett, W. C., and Manson, J. E. (2002) Elevated risk of cardiovascular disease prior to clinical diagnosis of type 2 diabetes. *Diabetes care* 25, 1129-1134

9. Hubert, H. B., Feinleib, M., McNamara, P. M., and Castelli, W. P. (1983) Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation* 67, 968-977

10. Wilson, P. W., D'Agostino, R. B., Sullivan, L., Parise, H., and Kannel, W. B. (2002) Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. *Archives of internal medicine* 162, 1867-1872

References

11. Castelli, W. P., Garrison, R. J., Wilson, P. W., Abbott, R. D., Kalousdian, S., and Kannel, W. B. (1986) Incidence of coronary heart disease and lipoprotein cholesterol levels: the Framingham Study. *Jama* 256, 2835-2838

12. Kannel, W. B., Castelli, W. P., Gordon, T., and McNamara, P. M. (1971) Serum cholesterol, lipoproteins, and the risk of coronary heart disease: the Framingham Study. *Annals of Internal Medicine* 74, 1-12

13. Dahlöf, B., Devereux, R. B., Kjeldsen, S. E., Julius, S., Beevers, G., de Faire, U., Fyhrquist, F., Ibsen, H., Kristiansson, K., and Lederballe-Pedersen, O. (2002) Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol. *The Lancet* 359, 995-1003

14. MacMahon, S., Peto, R., Collins, R., Godwin, J., Cutler, J., Sorlie, P., Abbott, R., Neaton, J., Dyer, A., and Stamler, J. (1990) Blood pressure, stroke, and coronary heart disease: part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet* 335, 765-774

15. Mackay, J., Mensah, G. A., Mendis, S., and Greenlund, K. (2004) *The atlas of heart disease and stroke*, World Health Organization

16. Vasan, R. S. (2006) Biomarkers of cardiovascular disease molecular basis and practical considerations. *Circulation* 113, 2335-2362

17. Canto, J. G., Kiefe, C. I., Rogers, W. J., Peterson, E. D., Frederick, P. D., French, W. J., Gibson, C. M., Pollack, C. V., Ornato, J. P., and Zalenski, R. J. (2011) Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. *Jama* 306, 2120-2127

18. Moohebati, M., Falsoleiman, H., Dehghani, M., Fazlinezhad, A., Daloee, M. H., Esmaeili, H., Parizadeh, S. M. R., Tavallaie, S., Rahsepar, A. A., and Paydar, R. (2011) Serum inflammatory and immune marker response after bare-metal or drug-eluting stent implantation following percutaneous coronary intervention. *Angiology* 62, 184-190

19. Phinikaridou, A., Qiao, Y., and Hamilton, J. A. (2012) Stable and Vulnerable Atherosclerotic Plaques. *Ultrasound and Carotid Bifurcation Atherosclerosis*, 3-25

References

20. Joshi, F. R., Lindsay, A. C., Obaid, D. R., Falk, E., and Rudd, J. H. (2012) Non-invasive imaging of atherosclerosis. *European Heart Journal-Cardiovascular Imaging* 13, 205-218

21. Klingenberg, R., Hasun, M., Corti, R., and Lüscher, T. F. (2012) Clinical Manifestations of Atherosclerosis. *Inflammation and Atherosclerosis*, 39-58

22. Narula, J., and Strauss, H. (2007) The popcorn plaques. *Nature medicine* 13, 532

23. Shah, P. K. (2003) Mechanisms of plaque vulnerability and rupture. *Journal of the American college of cardiology* 41, S15-S22

24. Blake, G., and Ridker, P. (2002) Inflammatory bio-markers and cardiovascular risk prediction. *Journal of internal medicine* 252, 283-294

25. Blake, G. J., and Ridker, P. M. (2001) Novel clinical markers of vascular wall inflammation. *Circulation research* 89, 763-771

26. Wu, J. C., Bengel, F. M., and Gambhir, S. S. (2007) Cardiovascular Molecular Imaging 1. *Radiology* 244, 337-355

27. Packard, R. R. S., and Libby, P. (2008) Inflammation in atherosclerosis: from vascular biology to biomarker discovery and risk prediction. *Clinical chemistry* 54, 24-38

28. Vengrenyuk, Y., Carlier, S., Xanthos, S., Cardoso, L., Ganatos, P., Virmani, R., Einav, S., Gilchrist, L., and Weinbaum, S. (2006) A hypothesis for vulnerable plaque rupture due to stress-induced debonding around cellular microcalcifications in thin fibrous caps. *Proceedings of the National Academy of Sciences* 103, 14678

29. Shah, P. K. (2007) Molecular mechanisms of plaque instability. *Current opinion in lipidology* 18, 492

30. Wong, N. D. (2014) Epidemiological studies of CHD and the evolution of preventive cardiology. *Nature Reviews Cardiology* 11, 276-289

31. Tobin, K. J. (2010) Stable angina pectoris: what does the current clinical evidence tell us. *J Am Osteopath Assoc* 110, 364-370

32. Yeghiazarians, Y., Braunstein, J. B., Askari, A., and Stone, P. H. (2000) Unstable angina pectoris. *New England Journal of Medicine* 342, 101-114

References

33. Antman, E., Bassand, J.-P., Klein, W., Ohman, M., Sendon, J. L. L., Rydén, L., Simoons, M., and Tendera, M. (2000) Myocardial infarction redefined—a consensus document of the Joint European Society of Cardiology/American College of Cardiology committee for the redefinition of myocardial infarction: the Joint European Society of Cardiology/American College of Cardiology Committee. *Journal of the American College of Cardiology* 36, 959-969

34. Zipes, D. P., and Wellens, H. J. (1998) Sudden cardiac death. *Circulation* 98, 2334-2351

35. Reichlin, T., Hochholzer, W., Bassetti, S., Steuer, S., Stelzig, C., Hartwiger, S., Biedert, S., Schaub, N., Buerge, C., and Potocki, M. (2009) Early diagnosis of myocardial infarction with sensitive cardiac troponin assays. *New England Journal of Medicine* 361, 858-867

36. Hazinski, M. F., and Field, J. M. (2010) 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science. *Circulation* 122, S639-S946

37. Fox, K. A., Mehta, S. R., Peters, R., Zhao, F., Lakkis, N., Gersh, B. J., and Yusuf, S. (2004) Benefits and Risks of the Combination of Clopidogrel and Aspirin in Patients Undergoing Surgical Revascularization for Non–ST-Elevation Acute Coronary Syndrome The Clopidogrel in Unstable angina to prevent Recurrent ischemic Events (CURE) Trial. *Circulation* 110, 1202-1208

38. Boersma, E., Harrington, R. A., Moliterno, D. J., White, H., Théroux, P., Van de Werf, F., de Torbal, A., Armstrong, P. W., Wallentin, L. C., and Wilcox, R. G. (2002) Platelet glycoprotein IIb/IIIa inhibitors in acute coronary syndromes: a meta-analysis of all major randomised clinical trials. *The Lancet* 359, 189-198

39. Basu, S., Senior, R., Raval, U., van der Does, R., Bruckner, T., and Lahiri, A. (1997) Beneficial Effects of Intravenous and Oral Carvedilol Treatment in Acute Myocardial Infarction A Placebo-Controlled, Randomized Trial. *Circulation* 96, 183-191

40. Freemantle, N., Cleland, J., Young, P., Mason, J., and Harrison, J. (1999) β Blockade after myocardial infarction: systematic review and meta regression analysis. *Bmj* 318, 1730-1737

References

41. Armstrong, P. W., Chang, W.-C., Wallentin, L., Goldstein, P., Granger, C. B., Bogaerts, K., Danays, T., and Van de Werf, F. (2006) Efficacy and safety of unfractionated heparin versus enoxaparin: a pooled analysis of ASSENT-3 and-3 PLUS data. *Canadian Medical Association Journal* 174, 1421-1426

42. Tea, K. K., Yusuf, S., Pfeffer, M., Kober, L., Hall, A., Pogue, J., Latini, R., and Collins, R. (2002) Effects of long-term treatment with angiotensin-converting-enzyme inhibitors in the presence or absence of aspirin: a systematic review. *The Lancet* 360, 1037-1043

43. Voors, A. A., de Kam, P. J., van den Berg, M. P., Borghi, C., Hochman, J. S., van Veldhuisen, D. J., and van Gilst, W. H. (2005) Acute administration of angiotensin converting enzyme inhibitors in thrombolysed myocardial infarction patients is associated with a decreased incidence of heart failure, but an increased re-infarction risk. *Cardiovascular drugs and therapy* 19, 119-124

44. Correia, L. C., Spósito, A. C., Lima, J. C., Magalhães, L. P., Passos, L. C., Rocha, M. S., D'Oliveira, A., and Esteves, J. P. (2003) Anti-inflammatory effect of atorvastatin (80 mg) in unstable angina pectoris and non–Q-wave acute myocardial infarction. *The American journal of cardiology* 92, 298-301

45. Lenderink, T., Boersma, E., Gitt, A. K., Zeymer, U., Wallentin, L., Van de Werf, F., Hasdai, D., Behar, S., and Simoons, M. L. (2006) Patients using statin treatment within 24 h after admission for ST-elevation acute coronary syndromes had lower mortality than non-users: a report from the first Euro Heart Survey on acute coronary syndromes. *European heart journal* 27, 1799-1804

46. Hannan, E. L., Wu, C., Walford, G., Culliford, A. T., Gold, J. P., Smith, C. R., Higgins, R. S., Carlson, R. E., and Jones, R. H. (2008) Drug-eluting stents vs. coronary-artery bypass grafting in multivessel coronary disease. *New England Journal of Medicine* 358, 331-341

47. Serruys, P. W., Morice, M.-C., Kappetein, A. P., Colombo, A., Holmes, D. R., Mack, M. J., Ståhle, E., Feldman, T. E., van den Brand, M., and Bass, E. J. (2009) Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. *New England Journal of Medicine* 360, 961-972

References

48. Danesh, J., Whincup, P., Walker, M., Lennon, L., Thomson, A., Appleby, P., Rumley, A., and Lowe, G. D. (2001) Fibrin D-dimer and coronary heart disease prospective study and meta-analysis. *Circulation* 103, 2323-2327

49. Puntmann, V. (2009) How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. *Postgraduate medical journal* 85, 538-545

50. Rifai, N., Gillette, M. A., and Carr, S. A. (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature biotechnology* 24, 971-983

51. Petricoin, E. F., Belluco, C., Araujo, R. P., and Liotta, L. A. (2006) The blood peptidome: a higher dimension of information content for cancer biomarker discovery. *Nature Reviews Cancer* 6, 961-967

52. Bouma, B. N., and Mosnier, L. O. (2003) Thrombin activatable fibrinolysis inhibitor (TAFI) at the interface between coagulation and fibrinolysis. *Pathophysiology of haemostasis and thrombosis* 33, 375-381

53. Ridker, P. M., Rifai, N., Stampfer, M. J., and Hennekens, C. H. (2000) Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation* 101, 1767-1772

54. Blankenberg, S., Luc, G., Ducimetière, P., Arveiler, D., Ferrières, J., Amouyel, P., Evans, A., Cambien, F., and Tiret, L. (2003) Interleukin-18 and the risk of coronary heart disease in European men the prospective epidemiological study of myocardial infarction (PRIME). *Circulation* 108, 2453-2459

55. Packard, C. J., O'Reilly, D. S., Caslake, M. J., McMahon, A. D., Ford, I., Cooney, J., Macphee, C. H., Suckling, K. E., Krishna, M., and Wilkinson, F. E. (2000) Lipoprotein-associated phospholipase A2 as an independent predictor of coronary heart disease. *New England Journal of Medicine* 343, 1148-1155

56. Meuwese, M. C., Stroes, E. S., Hazen, S. L., van Miert, J. N., Kuivenhoven, J. A., Schaub, R. G., Wareham, N. J., Luben, R., Kastelein, J. J., and Khaw, K.-T. (2007) Serum myeloperoxidase levels are associated with the future risk of coronary artery disease in apparently healthy individuals: the EPIC-Norfolk Prospective Population Study. *Journal of the American College of Cardiology* 50, 159-165

References

57. Ridker, P. M., Rifai, N., Pfeffer, M., Sacks, F., Lepage, S., Braunwald, E., Cholesterol, and Investigators, R. E. (2000) Elevation of tumor necrosis factor-α and increased risk of recurrent coronary events after myocardial infarction. *Circulation* 101, 2149-2153

58. Ridker, P. M., Hennekens, C. H., Roitman-Johnson, B., Stampfer, M. J., and Allen, J. (1998) Plasma concentration of soluble intercellular adhesion molecule 1 and risks of future myocardial infarction in apparently healthy men. *The Lancet* 351, 88-92

59. Hwang, S.-J., Ballantyne, C. M., Sharrett, A. R., Smith, L. C., Davis, C. E., Gotto, A. M., and Boerwinkle, E. (1997) Circulating adhesion molecules VCAM-1, ICAM-1, and E-selectin in carotid atherosclerosis and incident coronary heart disease cases the atherosclerosis risk in communities (ARIC) study. *Circulation* 96, 4219-4225

60. Ridker, P. M., Buring, J. E., and Rifai, N. (2001) Soluble P-selectin and the risk of future cardiovascular events. *Circulation* 103, 491-495

61. Varo, N., de Lemos, J. A., Libby, P., Morrow, D. A., Murphy, S. A., Nuzzo, R., Gibson, C. M., Cannon, C. P., Braunwald, E., and Schönbeck, U. (2003) Soluble CD40L risk prediction after acute coronary syndromes. *Circulation* 108, 1049-1052

62. Ballantyne, C. M., Hoogeveen, R. C., Bang, H., Coresh, J., Folsom, A. R., Heiss, G., and Sharrett, A. R. (2004) Lipoprotein-associated phospholipase A2, high-sensitivity C-reactive protein, and risk for incident coronary heart disease in middle-aged men and women in the Atherosclerosis Risk in Communities (ARIC) study. *Circulation* 109, 837-842

63. Luc, G., Bard, J.-M., Juhan-Vague, I., Ferrieres, J., Evans, A., Amouyel, P., Arveiler, D., Fruchart, J.-C., and Ducimetiere, P. (2003) C-reactive protein, interleukin-6, and fibrinogen as predictors of coronary heart disease The PRIME study. *Arteriosclerosis, thrombosis, and vascular biology* 23, 1255-1261

64. Johnson, B. D., Kip, K. E., Marroquin, O. C., Ridker, P. M., Kelsey, S. F., Shaw, L. J., Pepine, C. J., Sharaf, B., Merz, C. N. B., and Sopko, G. (2004) Serum Amyloid A as a Predictor of Coronary Artery Disease and Cardiovascular Outcome

References

in Women The National Heart, Lung, and Blood Institute–Sponsored Women's Ischemia Syndrome Evaluation (WISE). *Circulation* 109, 726-732

65. Vogiatzi, G., Tousoulis, D., and Stefanadis, C. (2009) The role of oxidative stress in atherosclerosis. *Hellenic J Cardiol* 50, 402-409

66. Koba, S., and Hirano, T. (2011) Dyslipidemia and atherosclerosis]. *Nihon rinsho. Japanese journal of clinical medicine* 69, 138

67. Rimm, E. B., Stampfer, M. J., Ascherio, A., Giovannucci, E., Colditz, G. A., and Willett, W. C. (1993) Vitamin E consumption and the risk of coronary heart disease in men. *New England Journal of Medicine* 328, 1450-1456

68. Hertog, M. G., Feskens, E. J., Kromhout, D., Hollman, P., and Katan, M. (1993) Dietary antioxidant flavonoids and risk of coronary heart disease: the Zutphen Elderly Study. *The Lancet* 342, 1007-1011

69. Meisinger, C., Baumert, J., Khuseyinova, N., Loewel, H., and Koenig, W. (2005) Plasma oxidized low-density lipoprotein, a strong predictor for acute coronary heart disease events in apparently healthy, middle-aged men from the general population. *Circulation* 112, 651-657

70. Wallenfeldt, K., Fagerberg, B., Wikstrand, J., and Hulthe, J. (2004) Oxidized low-density lipoprotein in plasma is a prognostic marker of subclinical atherosclerosis development in clinically healthy men. *Journal of internal medicine* 256, 413-420

71. Ahmed, E., Trifunovic, J., Stegmayr, B., Hallmans, G., and Lefvert, A. K. (1999) Autoantibodies against oxidatively modified LDL do not constitute a risk factor for stroke a nested case-control study. *Stroke* 30, 2541-2546

72. Braun, S., Ndrepepa, G., von Beckerath, N., Mehilli, J., Gorchakova, O., Vogt, W., Schömig, A., and Kastrati, A. (2005) Lack of association between circulating levels of plasma oxidized low-density lipoproteins and clinical outcome after coronary stenting. *American heart journal* 150, 550-556

73. Naruko, T., Ueda, M., Ehara, S., Itoh, A., Haze, K., Shirai, N., Ikura, Y., Ohsawa, M., Itabe, H., and Kobayashi, Y. (2006) Persistent high levels of plasma oxidized low-density lipoprotein after acute myocardial infarction predict stent restenosis. *Arteriosclerosis, thrombosis, and vascular biology* 26, 877-883

References

74. Tsimikas, S., Kiechl, S., Willeit, J., Mayr, M., Miller, E. R., Kronenberg, F., Xu, Q., Bergmark, C., Weger, S., and Oberhollenzer, F. (2006) Oxidized phospholipids predict the presence and progression of carotid and femoral atherosclerosis and symptomatic cardiovascular disease: five-year prospective results from the Bruneck study. *Journal of the American College of Cardiology* 47, 2219-2228

75. Vassalle, C., Botto, N., Andreassi, M. G., Berti, S., and Biagini, A. (2003) Evidence for enhanced 8-isoprostane plasma levels, as index of oxidative stressin vivo, in patients with coronary artery disease. *Coronary artery disease* 14, 213-218

76. Blankenberg, S., Rupprecht, H. J., Bickel, C., Torzewski, M., Hafner, G., Tiret, L., Smieja, M., Cambien, F., Meyer, J., and Lackner, K. J. (2003) Glutathione peroxidase 1 activity and cardiovascular events in patients with coronary artery disease. *New England Journal of Medicine* 349, 1605-1613

77. Vaziri, N., Kennedy, S. C., Kennedy, D., and Gonzales, E. (1992) Coagulation, fibrinolytic, and inhibitory proteins in acute myocardial infarction and angina pectoris. *The American journal of medicine* 93, 651-657

78. Whincup, P., Danesh, J., Walker, M., Lennon, L., Thomson, A., Appleby, P., Rumley, A., and Lowe, G. (2002) Von Willebrand factor and coronary heart disease. *Eur Heart J* 23, 1764-1770

79. Rumley, A., Lowe, G., Sweetnam, P., Yarnell, J., and Ford, R. (1999) Factor VIII, von Willebrand factor and the risk of major ischaemic heart disease in the Caerphilly Heart Study. *British journal of haematology* 105, 110-116

80. Lowe, G., Danesh, J., Lewington, S., Walker, M., Lennon, L., Thomson, A., Rumley, A., and Whincup, P. (2004) Tissue plasminogen activator antigen and coronary heart disease. *European heart journal* 25, 252-259

81. Smith, A., Patterson, C., Yarnell, J., Rumley, A., Ben-Shlomo, Y., and Lowe, G. (2005) Which hemostatic markers add to the predictive value of conventional risk factors for coronary heart disease and ischemic stroke? The Caerphilly Study. *Circulation* 112, 3080-3087

82. Arrell, D. K., Neverova, I., and Van Eyk, J. E. (2001) Cardiovascular proteomics evolution and potential. *Circulation research* 88, 763-773

References

83. Arab, S., Gramolini, A. O., Ping, P., Kislinger, T., Stanley, B., van Eyk, J., Ouzounian, M., MacLennan, D. H., Emili, A., and Liu, P. P. (2006) Cardiovascular proteomics: tools to develop novel biomarkers and potential applications. *Journal of the American College of Cardiology* 48, 1733-1741

84. Usui-Aoki, K., Kyo, M., Kawai, M., Murakami, M., Imai, K., Shimada, K., and Koga, H. (2006) Protein and antibody microarrays: Clues towards biomarker discovery. *Frontiers in Drug Design & Discovery* 2, 23-33

85. Meng, Z., and Veenstra, T. D. (2011) Targeted mass spectrometry approaches for protein biomarker verification. *Journal of proteomics* 74, 2650-2659

86. Maurya, P., Meleady, P., Dowling, P., and Clynes, M. (2007) Proteomic approaches for serum biomarker discovery in cancer. *Anticancer research* 27, 1247-1255

87. Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., and Domon, B. (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Molecular & Cellular Proteomics* 6, 1809-1817

88. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *Journal of biological chemistry* 250, 4007-4021

89. Klose, J. (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. *Humangenetik* 26, 231-243

90. Wilkins, M. R., Sanchez, J.-C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., and Williams, K. L. (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology and genetic engineering reviews* 13, 19-50

91. Tyers, M., and Mann, M. (2003) From genomics to proteomics. *Nature* 422, 193-197

92. Cho, W. (2007) Contribution of oncoproteomics to cancer biomarker discovery. *Mol Cancer* 6, 25-27

93. Ullah, M. F., and Aatif, M. (2009) The footprints of cancer development: Cancer biomarkers. *Cancer treatment reviews* 35, 193-200

94. Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., Van Sluyter, S. C., and Haynes, P. A. (2011) Less label,

more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11, 535-553

95. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry* 389, 1017-1031

96. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering* 11, 49-79

97. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics* 1, 376-386

98. Gruhler, A., Schulze, W. X., Matthiesen, R., Mann, M., and Jensen, O. N. (2005) Stable isotope labeling of Arabidopsis thaliana cells and quantitative proteomics by mass spectrometry. *Molecular & Cellular Proteomics* 4, 1697-1709

99. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-648

100. Harsha, H., Molina, H., and Pandey, A. (2008) Quantitative proteomics using stable isotope labeling with amino acids in cell culture. *Nature protocols* 3, 505-516

101. Elliott, M. H., Smith, D. S., Parker, C. E., and Borchers, C. (2009) Current trends in quantitative proteomics. *Journal of Mass Spectrometry* 44, 1637-1660

102. Mann, M. (2006) Functional and quantitative proteomics using SILAC. *Nature reviews Molecular cell biology* 7, 952-958

103. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., and Daniels, S. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics* 3, 1154-1169

104. Ye, H., Sun, L., Huang, X., Zhang, P., and Zhao, X. (2010) A proteomic approach for plasma biomarker discovery with 8-plex iTRAQ labeling and SCX-LC-MS/MS. *Molecular and cellular biochemistry* 343, 91-99

References

105. Zieske, L. R. (2006) A perspective on the use of iTRAQ™ reagent technology for protein complex and profiling studies. *Journal of experimental botany* 57, 1501-1508

106. Pottiez, G., Wiederin, J., Fox, H. S., and Ciborowski, P. (2012) Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *Journal of proteome research* 11, 3774-3781

107. Bantscheff, M., Boesche, M., Eberhard, D., Matthieson, T., Sweetman, G., and Kuster, B. (2008) Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Molecular & Cellular Proteomics* 7, 1702-1713

108. Kuzyk, M. A., Ohlund, L. B., Elliott, M. H., Smith, D., Qian, H., Delaney, A., Hunter, C. L., and Borchers, C. H. (2009) A comparison of MS/MS-based, stable-isotope-labeled, quantitation performance on ESI-quadrupole TOF and MALDI-TOF/TOF mass spectrometers. *Proteomics* 9, 3328-3340

109. DeSouza, L. V., Romaschin, A. D., Colgan, T. J., and Siu, K. M. (2009) Absolute quantification of potential cancer markers in clinical tissue homogenates using multiple reaction monitoring on a hybrid triple quadrupole/linear ion trap tandem mass spectrometer. *Analytical chemistry* 81, 3462-3470

110. Bouchal, P., Roumeliotis, T., Hrstka, R., Nenutil, R., Vojtesek, B., and Garbis, S. D. (2008) Biomarker discovery in low-grade breast cancer using isobaric stable isotope tags and two-dimensional liquid chromatography-tandem mass spectrometry (iTRAQ-2DLC-MS/MS) based quantitative proteomic analysis. *Journal of proteome research* 8, 362-373

111. Ralhan, R., DeSouza, L. V., Matta, A., Tripathi, S. C., Ghanny, S., Gupta, S. D., Bahadur, S., and Siu, K. M. (2008) Discovery and verification of head-and-neck cancer biomarkers by differential protein expression analysis using iTRAQ labeling, multidimensional liquid chromatography, and tandem mass spectrometry. *Molecular & Cellular Proteomics* 7, 1162-1173

112. Chen, Y.-T., Chen, C.-L., Chen, H.-W., Chung, T., Wu, C.-C., Chen, C.-D., Hsu, C.-W., Chen, M.-C., Tsui, K.-H., and Chang, P.-L. (2010) Discovery of novel bladder cancer biomarkers by comparative urine proteomics using iTRAQ technology. *Journal of proteome research* 9, 5803-5815

References

113. Shi, M., Caudle, W. M., and Zhang, J. (2009) Biomarker discovery in neurodegenerative diseases: a proteomic approach. *Neurobiology of disease* 35, 157-164

114. Guerreiro, N., Gomez-Mancilla, B., Williamson, B., Minkoff, M., and Guertin, S. (2009) Proteomic profiling of cerebrospinal fluid by 8-plex iTRAQ reveals potential biomarker candidates of Alzheimer's disease. *Clinical Proteomics* 5, 114-124

115. Zhang, X., Yin, X., Yu, H., Liu, X., Yang, F., Yao, J., Jin, H., and Yang, P. (2012) Quantitative proteomic analysis of serum proteins in patients with Parkinson's disease using an isobaric tag for relative and absolute quantification labeling, two-dimensional liquid chromatography, and tandem mass spectrometry. *Analyst* 137, 490-495

116. Jing, L., Parker, C. E., Seo, D., Hines, M. W., Dicheva, N., Yu, Y., Schwinn, D., Ginsburg, G. S., and Chen, X. (2011) Discovery of biomarker candidates for coronary artery disease from an APOE-knock out mouse model using iTRAQ-based multiplex quantitative proteomics. *Proteomics* 11, 2763-2776

117. Yin, X., Subramanian, S., Hwang, S.-J., O'Donnell, C. J., Fox, C. S., Courchesne, P., Muntendam, P., Gordon, N., Adourian, A., and Juhasz, P. (2014) Protein Biomarkers of New-Onset Cardiovascular Disease Prospective Study From the Systems Approach to Biomarker Research in Cardiovascular Disease Initiative. *Arteriosclerosis, thrombosis, and vascular biology* 34, 939-945

118. Levy, D., Yin, X., Lyas, A., Ho, J., Courchesne, P. L., Subramanian, S., Hwang, S.-J., O'Donnell, C. J., Ramachandran, V. S., and Benjamin, E. J. (2013) Protein Biomarkers of New-Onset Atherosclerotic Cardiovascular Disease. *Circulation* 128, A15535

119. Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Analytical chemistry* 73, 2836-2842

120. Wang, Y. K., Ma, Z., Quinn, D. F., and Fu, E. W. (2001) Inverse 18O labeling mass spectrometry for the rapid identification of marker/target proteins. *Analytical chemistry* 73, 3742-3750

References

121. Schnölzer, M., Jedrzejewski, P., and Lehmann, W. D. (1996) Protease-catalyzed incorporation of 18O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Electrophoresis* 17, 945-953

122. Vaslow, F. (1955) Kinetics of the chymotrypsin catalyzed oxygen exchange of N-acetyl-3: 5-dibromo-L-tyrosine. *Biochimica et biophysica acta* 16, 601-602

123. Silver, M. S., Stoddard, M., and Stein, T. P. (1970) Mechanism of the pepsin-catalyzed exchange of carboxylic acids with water-18O. *Journal of the American Chemical Society* 92, 2883-2890

124. Ong, S.-E., and Mann, M. (2005) Mass spectrometry–based proteomics turns quantitative. *Nature chemical biology* 1, 252-262

125. Pan, S., Zhang, H., Rush, J., Eng, J., Zhang, N., Patterson, D., Comb, M. J., and Aebersold, R. (2005) High throughput proteome screening for biomarker detection. *Molecular & Cellular Proteomics* 4, 182-190

126. Lemoine, J., Fortin, T., Salvador, A., Jaffuel, A., Charrier, J.-P., and Choquet-Kastylevsky, G. (2012) The current status of clinical proteomics and the use of MRM and MRM3 for biomarker validation.

127. Kirkpatrick, D. S., Gerber, S. A., and Gygi, S. P. (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* 35, 265-273

128. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences* 100, 6940-6945

129. Ahn, Y. H., Shin, P. M., Oh, N. R., Park, G. W., Kim, H., and Yoo, J. S. (2012) A lectin-coupled, targeted proteomic mass spectrometry (MRM MS) platform for identification of multiple liver cancer biomarkers in human plasma. *Journal of Proteomics*

130. Wang, M., You, J., Bemis, K. G., Tegeler, T. J., and Brown, D. P. (2008) Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. *Briefings in functional genomics & proteomics* 7, 329-339

References

131. Liu, H., Sadygov, R. G., and Yates, J. R. (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical chemistry* 76, 4193-4201

132. Megger, D. A., Bracht, T., Meyer, H. E., and Sitek, B. (2013) Label-free quantification in clinical proteomics. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1834, 1581-1590

133. Bondarenko, P. V., Chelius, D., and Shaler, T. A. (2002) Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Analytical chemistry* 74, 4741-4749

134. Chelius, D., and Bondarenko, P. V. (2002) Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *Journal of proteome research* 1, 317-323

135. Higgs, R. E., Knierman, M. D., Gelfanova, V., Butler, J. P., and Hale, J. E. (2005) Comprehensive label-free method for the relative quantification of proteins from biological samples. *Journal of proteome research* 4, 1442-1450

136. Zhang, B., VerBerkmoes, N. C., Langston, M. A., Uberbacher, E., Hettich, R. L., and Samatova, N. F. (2006) Detecting differential and correlated protein expression in label-free shotgun proteomics. *Journal of proteome research* 5, 2909-2918

137. Quintana, L. F., Campistol, J. M., Alcolea, M. P., Bañon-Maneus, E., Sol-González, A., and Cutillas, P. R. (2009) Application of label-free quantitative peptidomics for the identification of urinary biomarkers of kidney chronic allograft dysfunction. *Molecular & Cellular Proteomics* 8, 1658-1673

138. Xie, F., Liu, T., Qian, W.-J., Petyuk, V. A., and Smith, R. D. (2011) Liquid chromatography-mass spectrometry-based quantitative proteomics. *Journal of Biological Chemistry* 286, 25443-25449

139. Hanash, S. M., Pitteri, S. J., and Faca, V. M. (2008) Mining the plasma proteome for cancer biomarkers. *Nature* 452, 571-579

140. Devine, D., and Schubert, P. (2011) Proteomic applications in blood transfusion: working the jigsaw puzzle. *Vox Sanguinis* 100, 84-91

References

141. Veenstra, T. D., Conrads, T. P., Hood, B. L., Avellino, A. M., Ellenbogen, R. G., and Morrison, R. S. (2005) Biomarkers: mining the biofluid proteome. *Molecular & Cellular Proteomics* 4, 409-418

142. Polanski, M., and Anderson, N. L. (2006) A list of candidate cancer biomarkers for targeted proteomics. *Biomarker insights* 1, 1

143. Anderson, N. L., and Anderson, N. G. (2002) The human plasma proteome history, character, and diagnostic prospects. *Molecular & Cellular Proteomics* 1, 845-867

144. States, D. J., Omenn, G. S., Blackwell, T. W., Fermin, D., Eng, J., Speicher, D. W., and Hanash, S. M. (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nature biotechnology* 24, 333-338

145. Surinova, S., Schiess, R., Hüttenhain, R., Cerciello, F., Wollscheid, B., and Aebersold, R. (2010) On the development of plasma protein biomarkers. *Journal of proteome research* 10, 5-16

146. Echan, L. A., Tang, H. Y., Ali-Khan, N., Lee, K. B., and Speicher, D. W. (2005) Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 5, 3292-3303

147. Tang, H. Y., Ali-Khan, N., Echan, L. A., Levenkova, N., Rux, J. J., and Speicher, D. W. (2005) A novel four-dimensional strategy combining protein and peptide separation methods enables detection of low-abundance proteins in human plasma and serum proteomes. *Proteomics* 5, 3329-3342

148. Lei, T., He, Q. Y., Wang, Y. L., Si, L. S., and Chiu, J. F. (2008) Heparin chromatography to deplete high-abundance proteins for serum proteomics. *Clinica Chimica Acta* 388, 173-178

149. Wolters, D. A., Washburn, M. P., and Yates, J. R. (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Analytical chemistry* 73, 5683-5690

150. Zhang, H., and Aebersold, R. (2006) Isolation of glycoproteins and identification of their N-linked glycosylation sites. *METHODS IN MOLECULAR BIOLOGY-CLIFTON THEN TOTOWA-* 328, 177

References

151. Oda, Y., Nagasu, T., and Chait, B. T. (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nature biotechnology* 19, 379-382

152. Caby, M.-P., Lankar, D., Vincendeau-Scherrer, C., Raposo, G., and Bonnerot, C. (2005) Exosomal-like vesicles are present in human blood plasma. *International immunology* 17, 879-887

153. Hanash, S., and Celis, J. E. (2002) The Human Proteome Organization. *Molecular & Cellular Proteomics* 1, 413-414

154. Omenn, G. S., Blackwell, T. W., Fermin, D., Eng, J., Speicher, D. W., and Hanash, S. M. (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nature biotechnology* 24, 333-338

155. Greening, D. W., and Simpson, R. J. (2010) A centrifugal ultrafiltration strategy for isolating the low-molecular weight ($\leq$25K) component of human plasma proteome. *Journal of proteomics* 73, 637-648

156. Omenn, G. S., Adamski, M., Blackwell, T. W., Menon, R., Hermjakob, H., Apweiler, R., Haab, B. B., Simpson, R. J., Eddes, J. S., and Kapp, E. A. (2005) Overview of the HUPO Plasma Proteome Project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 5, 3226-3245

157. Griffin, N. M., Yu, J., Long, F., Oh, P., Shore, S., Li, Y., Koziol, J. A., and Schnitzer, J. E. (2010) Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nature biotechnology* 28, 83-89

158. Jensen, O. N. (2004) Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Current opinion in chemical biology* 8, 33-41

159. Aebersold, R., Rist, B., and Gygi, S. P. (2000) Quantitative proteome analysis: methods and applications. *Annals of the New York Academy of Sciences* 919, 33-47

160. Zielinska, D. F., Gnad, F., Schropp, K., Wisniewski, J. R., and Mann, M. (2012) Mapping N-glycosylation sites across seven evolutionarily distant species

reveals a divergent substrate proteome despite a common core machinery. *Mol Cell* 46, 542-548

161. Ruhaak, L. R., Koeleman, C. A., Uh, H. W., Stam, J. C., van Heemst, D., Maier, A. B., Houwing-Duistermaat, J. J., Hensbergen, P. J., Slagboom, P. E., Deelder, A. M., and Wuhrer, M. (2013) Targeted biomarker discovery by high throughput glycosylation profiling of human plasma alpha1-antitrypsin and immunoglobulin A. *PLoS One* 8, e73082

162. Zielinska, D. F., Gnad, F., Wisniewski, J. R., and Mann, M. (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141, 897-907

163. Kaji, H., Shikanai, T., Sasaki-Sawa, A., Wen, H., Fujita, M., Suzuki, Y., Sugahara, D., Sawaki, H., Yamauchi, Y., Shinkawa, T., Taoka, M., Takahashi, N., Isobe, T., and Narimatsu, H. (2012) Large-scale identification of N-glycosylated proteins of mouse tissues and construction of a glycoprotein database, GlycoProtDB. *J Proteome Res* 11, 4553-4566

164. Berven, F. S., Ahmad, R., Clauser, K. R., and Carr, S. A. (2010) Optimizing performance of glycopeptide capture for plasma proteomics. *J Proteome Res* 9, 1706-1715

165. Durand, G., and Seta, N. (2000) Protein glycosylation and diseases: blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring. *Clinical chemistry* 46, 795-805

166. Lai, Z. W., Nice, E. C., and Schilling, O. (2013) Glycocapture-based proteomics for secretome analysis. *Proteomics* 13, 512-525

167. Ahn, Y. H., Kim, K. H., Shin, P. M., Ji, E. S., Kim, H., and Yoo, J. S. (2012) Identification of low-abundance cancer biomarker candidate TIMP1 from serum with lectin fractionation and peptide affinity enrichment by ultrahigh-resolution mass spectrometry. *Anal Chem* 84, 1425-1431

168. Cheow, E. S. H., Sim, K. H., de Kleijn, D., Lee, C. N., Sorokin, V., and Sze, S. K. (2015) Simultaneous enrichment of plasma soluble and extracellular vesicular glycoproteins using prolonged ultracentrifugation-ERLIC approach. *Molecular & Cellular Proteomics*, mcp. O114. 046391

References

169. Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nature Reviews Molecular Cell Biology* 13, 448-462

170. Zhang, H., Li, X.-j., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nature biotechnology* 21, 660-666

171. Gonzalez-Begne, M., Lu, B., Liao, L., Xu, T., Bedi, G., Melvin, J. E., and Yates, J. R., 3rd (2011) Characterization of the human submandibular/sublingual saliva glycoproteome using lectin affinity chromatography coupled to multidimensional protein identification technology. *J Proteome Res* 10, 5031-5046

172. Jung, K., Cho, W., and Regnier, F. E. (2009) Glycoproteomics of plasma based on narrow selectivity lectin affinity chromatography. *J Proteome Res* 8, 643-650

173. Tian, Y., Zhou, Y., Elliott, S., Aebersold, R., and Zhang, H. (2007) Solid-phase extraction of N-linked glycopeptides. *Nature Protocols* 2, 334-339

174. Xia, Y., Dobaczewski, M., Gonzalez-Quesada, C., Chen, W., Biernacka, A., Li, N., Lee, D.-W., and Frangogiannis, N. G. (2011) Endogenous thrombospondin 1 protects the pressure-overloaded myocardium by modulating fibroblast phenotype and matrix metabolism. *Hypertension* 58, 902-911

175. Nishioka, T., Onishi, K., Shimojo, N., Nagano, Y., Matsusaka, H., Ikeuchi, M., Ide, T., Tsutsui, H., Hiroe, M., and Yoshida, T. (2010) Tenascin-C may aggravate left ventricular remodeling and function after myocardial infarction in mice. *American Journal of Physiology-Heart and Circulatory Physiology* 298, H1072-H1078

176. Oka, T., Xu, J., Kaiser, R. A., Melendez, J., Hambleton, M., Sargent, M. A., Lorts, A., Brunskill, E. W., Dorn, G. W., and Conway, S. J. (2007) Genetic manipulation of periostin expression reveals a role in cardiac hypertrophy and ventricular remodeling. *Circulation research* 101, 313-321

177. Parker, B. L., Palmisano, G., Edwards, A. V., White, M. Y., Engholm-Keller, K., Lee, A., Scott, N. E., Kolarich, D., Hambly, B. D., and Packer, N. H. (2011) Quantitative N-linked glycoproteomics of myocardial ischemia and reperfusion

injury reveals early remodeling in the extracellular environment. *Molecular & Cellular Proteomics* 10, M110. 006833

178. Cubedo, J., Padró, T., and Badimon, L. (2014) Glycoproteome of human apolipoprotein AI: N-and O-glycosylated forms are increased in patients with acute myocardial infarction. *Translational Research* 164, 209-222

179. Stark, G. R., Stein, W. H., and Moore, S. (1960) Reactions of the cyanate present in aqueous urea with amino acids and proteins. *Journal of Biological Chemistry* 235, 3177-3181

180. Kraus, L. M., Gaber, L., Handorf, C. R., Marti, H.-P., and Kraus Jr, A. (2001) Carbamoylation of glomerular and tubular proteins in patients with kidney failure: a potential mechanism of ongoing renal damage. *Swiss medical weekly* 131, 139-134

181. Gross, M.-L., Piecha, G., Bierhaus, A., Hanke, W., Henle, T., Schirmacher, P., and Ritz, E. (2011) Glycated and carbamylated albumin are more "nephrotoxic" than unmodified albumin in the amphibian kidney. *American Journal of Physiology-Renal Physiology* 301, F476-F485

182. Wang, Z., Nicholls, S. J., Rodriguez, E. R., Kummu, O., Hörkkö, S., Barnard, J., Reynolds, W. F., Topol, E. J., DiDonato, J. A., and Hazen, S. L. (2007) Protein carbamylation links inflammation, smoking, uremia and atherogenesis. *Nature medicine* 13, 1176-1184

183. Pietrement, C., Gorisse, L., Jaisson, S., and Gillery, P. (2013) Chronic increase of urea leads to carbamylated proteins accumulation in tissues in a mouse model of CKD.

184. Jaisson, S., Gorisse, L., Pietrement, C., and Gillery, P. (2012) Quantification of plasma homocitrulline using hydrophilic interaction liquid chromatography (HILIC) coupled to tandem mass spectrometry. *Analytical and bioanalytical chemistry* 402, 1635-1641

185. J'Neka, S. C., Sandoval, P. C., Liu, G., Chou, C.-L., Hoffert, J. D., and Knepper, M. A. (2013) Endogenous carbamylation of renal medullary proteins.

186. Berg, A. H., Drechsler, C., Wenger, J., Buccafusca, R., Hod, T., Kalim, S., Ramma, W., Parikh, S. M., Steen, H., and Friedman, D. J. (2013) Carbamylation of

serum albumin as a risk factor for mortality in patients with kidney failure. *Science translational medicine* 5, 175ra129-175ra129

187. Ok, E., Basnakian, A. G., Apostolov, E. O., Barri, Y. M., and Shah, S. V. (2005) Carbamylated low-density lipoprotein induces death of endothelial cells: a link to atherosclerosis in patients with kidney disease. *Kidney international* 68, 173-178

188. Apostolov, E. O., Ray, D., Savenka, A. V., Shah, S. V., and Basnakian, A. G. (2010) Chronic uremia stimulates LDL carbamylation and atherosclerosis. *Journal of the American Society of Nephrology* 21, 1852-1857

189. Jaisson, S., Pietrement, C., and Gillery, P. (2011) Carbamylation-derived products: bioactive compounds and potential biomarkers in chronic renal failure and atherosclerosis. *Clinical chemistry* 57, 1499-1505

190. Horkko, S., Huttunen, K., and Kesaniemi, Y. A. (1995) Decreased clearance of low-density lipoprotein in uremic patients under dialysis treatment. *Kidney international* 47, 1732-1740

191. Apostolov, E. O., Basnakian, A. G., Ok, E., and Shah, S. V. (2012) Carbamylated low-density lipoprotein: nontraditional risk factor for cardiovascular events in patients with chronic kidney disease. *Journal of Renal Nutrition* 22, 134-138

192. Cocucci, E., Racchetti, G., and Meldolesi, J. (2009) Shedding microvesicles: artefacts no more. *Trends in cell biology* 19, 43-51

193. Verweij, F. J., van Eijndhoven, M. A., Hopmans, E. S., Vendrig, T., Wurdinger, T., Cahir-McFarland, E., Kieff, E., Geerts, D., van der Kant, R., and Neefjes, J. (2011) LMP1 association with CD63 in endosomes and secretion via exosomes limits constitutive NF-κB activation. *The EMBO journal* 30, 2115-2129

194. Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J. J., and Lötvall, J. O. (2007) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nature Cell Biology* 9, 654-659

195. Mathivanan, S., Ji, H., and Simpson, R. J. (2010) Exosomes: extracellular organelles important in intercellular communication. *Journal of proteomics* 73, 1907-1920

References

196. Scanu, A., Molnarfi, N., Brandt, K. J., Gruaz, L., Dayer, J. M., and Burger, D. (2008) Stimulated T cells generate microparticles, which mimic cellular contact activation of human monocytes: differential regulation of pro-and anti-inflammatory cytokine production by high-density lipoproteins. *Journal of leukocyte biology* 83, 921-927

197. Chaput, N., and Théry, C. (2011) Exosomes: immune properties and potential clinical implementations. pp. 419-440, Springer

198. Nolan, S., Dixon, R., Norman, K., Hellewell, P., and Ridger, V. (2008) Nitric oxide regulates neutrophil migration through microparticle formation. *The American journal of pathology* 172, 265-273

199. Ratajczak, J., Miekus, K., Kucia, M., Zhang, J., Reca, R., Dvorak, P., and Ratajczak, M. (2006) Embryonic stem cell-derived microvesicles reprogram hematopoietic progenitors: evidence for horizontal transfer of mRNA and protein delivery. *Leukemia* 20, 847-856

200. Gatti, S., Bruno, S., Deregibus, M. C., Sordi, A., Cantaluppi, V., Tetta, C., and Camussi, G. (2011) Microvesicles derived from human adult mesenchymal stem cells protect against ischaemia–reperfusion-induced acute and chronic kidney injury. *Nephrology Dialysis Transplantation*, gfr015

201. Bellingham, S. A., Guo, B., Coleman, B., and Hill, A. F. (2012) Exosomes: vehicles for the transfer of toxic proteins associated with neurodegenerative diseases? *Frontiers in physiology* 3, 124

202. Martínez, M. C., Tesse, A., Zobairi, F., and Andriantsitohaina, R. (2005) Shed membrane microparticles from circulating and vascular cells in regulating vascular function. *American Journal of Physiology-Heart and Circulatory Physiology* 288, H1004-H1009

203. Bernal-Mizrachi, L., Jy, W., Jimenez, J. J., Pastor, J., Mauro, L. M., Horstman, L. L., de Marchena, E., and Ahn, Y. S. (2003) High levels of circulating endothelial microparticles in patients with acute coronary syndromes. *American heart journal* 145, 962-970

References

204. Tilley, R. E., Holscher, T., Belani, R., Nieva, J., and Mackman, N. (2008) Tissue factor activity is increased in a combined platelet and microparticle sample from cancer patients. *Thrombosis research* 122, 604-609

205. Davila, M., Amirkhosravi, A., Coll, E., Desai, H., Robles, L., Colon, J., Baker, C., and Francis, J. (2008) Tissue factor-bearing microparticles derived from tumor cells: impact on coagulation activation. *Journal of Thrombosis and Haemostasis* 6, 1517-1524

206. Rozmyslowicz, T., Majka, M., Kijowski, J., Murphy, S. L., Conover, D. O., Poncz, M., Ratajczak, J., Gaulton, G. N., and Ratajczak, M. Z. (2003) Platelet-and megakaryocyte-derived microparticles transfer CXCR4 receptor to CXCR4-null cells and make them susceptible to infection by X4-HIV. *Aids* 17, 33

207. Mack, M., Kleinschmidt, A., Bruhl, H., Klier, C., Nelson, P. J., Cihak, J., Plachy, J., Stangassinger, M., Erfle, V., and Schlondorff, D. (2000) Transfer of the chemokine receptor CCR5 between cells by membrane-derived microparticles: a mechanism for cellular human immunodeficiency virus 1 infection. *Nature Medicine* 6, 769-775

208. Taylor, D. D., Zacharias, W., and Gercel-Taylor, C. (2011) Exosome isolation for proteomic analyses and RNA profiling. *Serum/plasma proteomics*, pp. 235-246, Springer

209. Tauro, B. J., Greening, D. W., Mathias, R. A., Ji, H., Mathivanan, S., Scott, A. M., and Simpson, R. J. (2012) Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. *Methods* 56, 293-304

210. Cheruvanky, A., Zhou, H., Pisitkun, T., Kopp, J. B., Knepper, M. A., Yuen, P. S., and Star, R. A. (2007) Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator. *American Journal of Physiology-Renal Physiology* 292, F1657-F1661

211. Clayton, A., Court, J., Navabi, H., Adams, M., Mason, M. D., Hobot, J. A., Newman, G. R., and Jasani, B. (2001) Analysis of antigen presenting cell derived exosomes, based on immuno-magnetic isolation and flow cytometry. *Journal of immunological methods* 247, 163-174

References

212. Timmers, L., Lim, S. K., Hoefer, I. E., Arslan, F., Lai, R. C., van Oorschot, A. A., Goumans, M. J., Strijder, C., Sze, S. K., and Choo, A. (2011) Human mesenchymal stem cell-conditioned medium improves cardiac function following myocardial infarction. *Stem cell research* 6, 206-214

213. Lai, R. C., Arslan, F., Lee, M. M., Sze, N. S. K., Choo, A., Chen, T. S., Salto-Tellez, M., Timmers, L., Lee, C. N., and El Oakley, R. M. (2010) Exosome secreted by MSC reduces myocardial ischemia/reperfusion injury. *Stem cell research* 4, 214-222

214. Zhang, Z., and Chan, D. W. (2010) The road from discovery to clinical diagnostics: lessons learned from the first FDA-cleared in vitro diagnostic multivariate index assay of proteomic biomarkers. *Cancer Epidemiology Biomarkers & Prevention* 19, 2995-2999

215. Abraham, J. (2010) OVA1 test for preoperative assessment of ovarian cancer. *Community Oncology* 7, 249-250

216. Fuzery, A., Levin, J., Chan, M. M., and Chan, D. W. (2013) Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin Proteomics* 10, 13

217. Bassand, J. P., Hamm, C. W., Ardissino, D., Boersma, E., Budaj, A., Fernández-Avilés, F., Fox, K. A. A., Hasdai, D., Ohman, E. M., and Wallentin, L. (2007) Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes The Task Force for the Diagnosis and Treatment of Non-ST-Segment Elevation Acute Coronary Syndromes of the European Society of Cardiology. *European heart journal* 28, 1598-1660

218. Chan, K. H., and Ng, M. K. C. (2012) Is there a role for coronary angiography in the early detection of the vulnerable plaque? *International Journal of Cardiology*

219. Cullen, P., Baetta, R., Bellosta, S., Bernini, F., Chinetti, G., Cignarella, A., von Eckardstein, A., Exley, A., Goddard, M., and Hofker, M. (2003) Rupture of the atherosclerotic plaque. *Arteriosclerosis, thrombosis, and vascular biology* 23, 535-542

References

220. Deutsch, E. W., Lam, H., and Aebersold, R. (2008) PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO reports* 9, 429-434

221. Pathan, M., Keerthikumar, S., Ang, C. S., Gangoda, L., Quek, C. Y., Williamson, N. A., Mouradov, D., Sieber, O. M., Simpson, R. J., and Salim, A. (2015) Technical brief funrich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics*

222. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *Journal of molecular biology* 215, 403-410

223. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* 25, 3389-3402

224. Song, X., Bandow, J., Sherman, J., Baker, J. D., Brown, P. W., McDowell, M. T., and Molloy, M. P. (2008) iTRAQ experimental design for plasma biomarker discovery. *Journal of proteome research* 7, 2952-2958

225. Tang, H.-Y., Beer, L. A., Barnhart, K. T., and Speicher, D. W. (2011) Rapid verification of candidate serological biomarkers using gel-based, label-free multiple reaction monitoring. *Journal of proteome research* 10, 4005-4017

226. Blankley, R. T., Fisher, C., Westwood, M., North, R., Baker, P. N., Walker, M. J., Williamson, A., Whetton, A. D., Lin, W., and McCowan, L. (2013) A label-free selected reaction monitoring workflow identifies a subset of pregnancy specific glycoproteins as potential predictive markers of early-onset pre-eclampsia. *Molecular & Cellular Proteomics* 12, 3148-3159

227. Choi, S., Kim, J., Yea, K., Suh, P.-G., Kim, J., and Ryu, S. H. (2010) Targeted label-free quantitative analysis of secretory proteins from adipocytes in response to oxidative stress. *Analytical biochemistry* 401, 196-202

228. Ekmekci, H., Sonmez, H., Ekmekci, O. B., Ozturk, Z., Domanic, N., and Kokoglu, E. (2002) Plasma vitronectin levels in patients with coronary atherosclerosis are increased and correlate with extent of disease. *Journal of thrombosis and thrombolysis* 14, 221-225

References

229. Kleinegris, M. C., ten Cate-Hoek, A. J., and ten Cate, H. (2012) Coagulation and the vessel wall in thrombosis and atherosclerosis. *POLSKIE ARCHIWUM MEDYCYNY WEWNĘTRZNEJ* 122, 557-566

230. Camerer, E., Huang, W., and Coughlin, S. R. (2000) Tissue factor-and factor X-dependent activation of protease-activated receptor 2 by factor VIIa. *Proceedings of the National Academy of Sciences* 97, 5255-5260

231. Hamilton, J. R., Frauman, A. G., and Cocks, T. M. (2001) Increased expression of protease-activated receptor-2 (PAR2) and PAR4 in human coronary artery by inflammatory stimuli unveils endothelium-dependent relaxations to PAR2 and PAR4 agonists. *Circulation research* 89, 92-98

232. Damiano, B. P., D'Andrea, M. R., de Garavilla, L., Cheung, W.-m., and Andrade-Gordon, P. (1999) Increased expression of protease activated receptor-2 (PAR-2) in balloon-injured rat carotid artery. *Thromb Haemost* 81, 808-814

233. Li, J., Chanda, D., Shiri-Sverdlov, R., and Neumann, D. (2015) MSP: An emerging player in metabolic syndrome. *Cytokine & growth factor reviews* 26, 75-82

234. Cantaluppi, V., Biancone, L., Romanazzi, G. M., Figliolini, F., Beltramo, S., Galimi, F., Camboni, M. G., Deriu, E., Conaldi, P., and Bottelli, A. (2008) Macrophage stimulating protein may promote tubular regeneration after acute injury. *Journal of the American Society of Nephrology* 19, 1904-1918

235. Lee, K. E., Kim, E. Y., Kim, C. S., Choi, J. S., Bae, E. H., Ma, S. K., Kim, K. K., Lee, J. U., and Kim, S. W. (2013) Macrophage-stimulating protein attenuates gentamicin-induced inflammation and apoptosis in human renal proximal tubular epithelial cells. *Biochemical and biophysical research communications* 434, 527-533

236. Brunelleschi, S., Penengo, L., Lavagno, L., Santoro, C., Colangelo, D., Viano, I., and Gaudino, G. (2001) Macrophage stimulating protein (MSP) evokes superoxide anion production by human macrophages of different origin. *British journal of pharmacology* 134, 1285-1295

237. Shah, A. M. (2000) Inducible nitric oxide synthase and cardiovascular disease. *Cardiovascular research* 45, 148-155

References

238. Dayal, S., Blokhin, I. O., Erger, R. A., Jensen, M., Arning, E., Stevens, J. W., Bottiglieri, T., Faraci, F. M., and Lentz, S. R. (2014) Protective vascular and cardiac effects of inducible nitric oxide synthase in mice with hyperhomocysteinemia.

239. Lindahl, B., Toss, H., Siegbahn, A., Venge, P., and Wallentin, L. (2000) Markers of myocardial damage and inflammation in relation to long-term mortality in unstable coronary artery disease. *New England Journal of Medicine* 343, 1139-1147

240. Heeschen, C., Hamm, C. W., Bruemmer, J., and Simoons, M. L. (2000) Predictive value of C-reactive protein and troponin T in patients with unstable angina: a comparative analysis. *Journal of the American College of Cardiology* 35, 1535-1542

241. Verma, S., Szmitko, P. E., and Ridker, P. M. (2005) C-reactive protein comes of age. *Nature Clinical Practice Cardiovascular Medicine* 2, 29-36

242. Kiernan, U. A., Nedelkov, D., and Nelson, R. W. (2006) Multiplexed mass spectrometric immunoassay in biomarker research: a novel approach to the determination of a myocardial infarct. *Journal of proteome research* 5, 2928-2934

243. Liuzzo, G., Biasucci, L. M., Gallimore, J. R., Grillo, R. L., Rebuzzi, A. G., Pepys, M. B., and Maseri, A. (1994) The prognostic value of C-reactive protein and serum amyloid a protein in severe unstable angina. *New England journal of medicine* 331, 417-424

244. Song, C., Shen, Y., Yamen, E., Hsu, K., Yan, W., Witting, P. K., Geczy, C. L., and Freedman, S. B. (2009) Serum amyloid A may potentiate prothrombotic and proinflammatory events in acute coronary syndromes. *Atherosclerosis* 202, 596-604

245. Frangogiannis, N. G. (2014) The inflammatory response in myocardial injury, repair, and remodelling. *Nature Reviews Cardiology* 11, 255-265

246. Frangogiannis, N. G. (2012) Regulation of the inflammatory response in cardiac repair. *Circulation research* 110, 159-173

247. Schmid, K. (1953) Preparation and Properties of Serum and Plasma Proteins. XXIX. Separation from Human Plasma of Polysaccharides, Peptides and Proteins of Low Molecular Weight. Crystallization of an Acid Glycoprotein1a, b, c. *Journal of the American Chemical Society* 75, 60-68

References

248. Daemen, M. A., Heemskerk, V. H., van't Veer, C., Denecker, G., Wolfs, T. G., Vandenabeele, P., and Buurman, W. A. (2000) Functional protection by acute phase proteins α1-acid glycoprotein and α1-antitrypsin against ischemia/reperfusion injury by preventing apoptosis and inflammation. *Circulation* 102, 1420-1426

249. Williams, J. P., Weiser, M. R., Pechet, T. T., Kobzik, L., Moore, F. D., and Hechtman, H. B. (1997) α1-Acid glycoprotein reduces local and remote injuries after intestinal ischemia in the rat. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 273, G1031-G1035

250. Tilg, H., Vannier, E., Vachino, G., Dinarello, C., and Mier, J. (1993) Antiinflammatory properties of hepatic acute phase proteins: preferential induction of interleukin 1 (IL-1) receptor antagonist over IL-1 beta synthesis by human peripheral blood mononuclear cells. *The journal of experimental medicine* 178, 1629-1636

251. Laine, E., Couderc, R., Roch-Arveiller, M., Vasson, M., Giroud, J., and Raichvarg, D. (1990) Modulation of human polymorphonuclear neutrophil functions by α1acid glycoprotein. *Inflammation* 14, 1-9

252. Farshchian, M., Kivisaari, A., Ala-aho, R., Riihilä, P., Kallajoki, M., Grénman, R., Peltonen, J., Pihlajaniemi, T., Heljasvaara, R., and Kähäri, V.-M. (2011) Serpin peptidase inhibitor clade A member 1 (SerpinA1) is a novel biomarker for progression of cutaneous squamous cell carcinoma. *The American journal of pathology* 179, 1110-1119

253. Bot, I., Bot, M., van Heiningen, S. H., van Santbrink, P. J., Lankhuizen, I. M., Hartman, P., Gruener, S., Hilpert, H., van Berkel, T. J., and Fingerle, J. (2011) Mast cell chymase inhibition reduces atherosclerotic plaque progression and improves plaque stability in ApoE−/− mice. *Cardiovascular research* 89, 244-252

254. Lieberman, J., Schleissner, L., Tachiki, K. H., and Kling, A. S. (1995) Serum α1-antichymotrypsin level as a marker for Alzheimer-type dementia. *Neurobiology of aging* 16, 747-753

255. DeKosky, S. T., Ikonomovic, M. D., Wang, X., Farlow, M., Wisniewski, S., Lopez, O. L., Becker, J. T., Saxton, J., Klunk, W. E., and Sweet, R. (2003) Plasma

and cerebrospinal fluid α1-antichymotrypsin levels in Alzheimer's disease: Correlation with cognitive impairment. *Annals of neurology* 53, 81-90

256. Ma, J., Yee, A., Brewer, H. B., Das, S., and Potter, H. (1994) Amyloid-associated proteins α1-antichymotrypsin and apolipoprotein E promote assembly of Alzheimer β-protein into filaments.

257. Nilsson, L. N., Arendash, G. W., Leighty, R. E., Costa, D. A., Low, M. A., Garcia, M. F., Cracciolo, J. R., Rojiani, A., Wu, X., and Bales, K. R. (2004) Cognitive impairment in PDAPP mice depends on ApoE and ACT-catalyzed amyloid formation. *Neurobiology of aging* 25, 1153-1167

258. Matthews, K. W., Mueller-Ortiz, S. L., and Wetsel, R. A. (2004) Carboxypeptidase N: a pleiotropic regulator of inflammation. *Molecular immunology* 40, 785-793

259. Skidgel, R. A., and Erdös, E. G. (2007) Structure and function of human plasma carboxypeptidase N, the anaphylatoxin inactivator. *International immunopharmacology* 7, 1888-1899

260. Bokisch, V. A., and Müller-Eberhard, H. J. (1970) Anaphylatoxin inactivator of human plasma: its isolation and characterization as a carboxypeptidase. *Journal of Clinical Investigation* 49, 2427

261. Walker, J., Binette, T., Mackova, M., Lambkin, G., Mitchell, L., and Bajzar, L. (2008) Proteolytic cleavage of carboxypeptidase N markedly increases its antifibrinolytic activity. *Journal of Thrombosis and Haemostasis* 6, 848-855

262. Zaman, A., Helft, G., Worthley, S., and Badimon, J. (2000) The role of plaque rupture and thrombosis in coronary artery disease. *Atherosclerosis* 149, 251-266

263. Belyantseva, I. A., Perrin, B. J., Sonnemann, K. J., Zhu, M., Stepanyan, R., McGee, J., Frolenkov, G. I., Walsh, E. J., Friderici, K. H., and Friedman, T. B. (2009) γ-Actin is required for cytoskeletal maintenance but not development. *Proceedings of the National Academy of Sciences* 106, 9703-9708

264. Pasquier, E., Tuset, M.-P., Sinnappan, S., Carnell, M., Macmillan, A., and Kavallaris, M. (2015) γ-Actin plays a key role in endothelial cell motility and neovessel maintenance. *Vascular cell* 7, 2

References

265. Baranwal, S., Naydenov, N. G., Harris, G., Dugina, V., Morgan, K. G., Chaponnier, C., and Ivanov, A. I. (2012) Nonredundant roles of cytoplasmic β-and γ-actin isoforms in regulation of epithelial apical junctions. *Molecular biology of the cell* 23, 3542-3553

266. Wai-man, R. W., Hawe, E., Li, L. K., Miller, G. J., Nicaud, V., Pennacchio, L. A., Humphries, S. E., and Talmud, P. J. (2003) Apolipoprotein AIV gene variant S347 is associated with increased risk of coronary heart disease and lower plasma apolipoprotein AIV levels. *Circulation research* 92, 969-975

267. Kronenberg, F., Stühlinger, M., Trenkwalder, E., Geethanjali, F., Pachinger, O., von Eckardstein, A., and Dieplinger, H. (2000) Low apolipoprotein A-IV plasma concentrations in men with coronary artery disease. *Journal of the American College of Cardiology* 36, 751-757

268. Manpuya, M., Guo, J., and Zhao, Y. (2001) The relationship between plasma apolipoprotein A-IV levels and coronary heart disease. *Chinese medical journal* 114, 275-279

269. Chakravarti, S., Petroll, W. M., Hassell, J. R., Jester, J. V., Lass, J. H., Paul, J., and Birk, D. E. (2000) Corneal opacity in lumican-null mice: defects in collagen fibril structure and packing in the posterior stroma. *Investigative ophthalmology & visual science* 41, 3365

270. Iozzo, R. V. (1997) The family of the small leucine-rich proteoglycans: key regulators of matrix assembly and cellular growth. *Critical Reviews in Biochemistry and Molecular Biology* 32, 141-174

271. Baba, H., Ishiwata, T., Takashi, E., Xu, G., and Asano, G. (2001) Expression and localization of lumican in the ischemic and reperfused rat heart. *Japanese circulation journal* 65, 445-450

272. Onda, M., Ishiwata, T., Kawahara, K., Wang, R., Naito, Z., and Sugisaki, Y. (2002) Expression of lumican in thickened intima and smooth muscle cells in human coronary atherosclerosis. *Experimental and molecular pathology* 72, 142-149

273. Dupuis, L. E., Berger, M. G., Feldman, S., Doucette, L., Fowlkes, V., Chakravarti, S., Thibaudeau, S., Alcala, N. E., Bradshaw, A. D., and Kern, C. B.

References

(2015) Lumican deficiency results in cardiomyocyte hypertrophy with altered collagen assembly. *Journal of molecular and cellular cardiology* 84, 70-80

274. Naughton, B. J., Duncan, F. J., Murrey, D., Ware, T., Meadows, A., McCarty, D. M., and Fu, H. (2013) Amyloidosis, synucleinopathy, and prion encephalopathy in a neuropathic lysosomal storage disease: the CNS-biomarker potential of peripheral blood.

275. Van Hove, J. L., Wevers, R. A., Van Cleemput, J., Moerman, P., Sciot, R., Matthijs, G., Schollen, E., de Jong, J. G., Carey, W. F., and Muller, V. (2003) Late-Onset visceral presentation with cardiomyopathy and without neurological symptoms of adult Sanfilippo A syndrome. *American Journal of Medical Genetics Part A* 118, 382-387

276. Pillarisetti, S., Paka, L., Obunike, J. C., Berglund, L., and Goldberg, I. J. (1997) Subendothelial retention of lipoprotein (a). Evidence that reduced heparan sulfate promotes lipoprotein binding to subendothelial matrix. *Journal of Clinical Investigation* 100, 867

277. Hollmann, J., Schmidt, A., von Bassewitz, D.-B., and Buddecke, E. (1989) Relationship of sulfated glycosaminoglycans and cholesterol content in normal and arteriosclerotic human aorta. *Arteriosclerosis, Thrombosis, and Vascular Biology* 9, 154-158

278. Völker, W., Schmidt, A., Oortmann, W., Broszey, T., Faber, V., and Buddecke, E. (1990) Mapping of proteoglycans in atherosclerotic lesions. *European heart journal* 11, 29-40

279. Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Blaha, M. J., Dai, S., Ford, E. S., Fox, C. S., and Franco, S. (2014) Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation* 129, e28

280. Liu, H.-H., and Li, J.-J. (2015) Aging and dyslipidemia: A review of potential mechanisms. *Ageing research reviews* 19, 43-52

281. Gaziano, T., Reddy, K. S., Paccaud, F., Horton, S., and Chaturvedi, V. (2006) Cardiovascular disease.

References

282. Frangogiannis, N. G., Smith, C. W., and Entman, M. L. (2002) The inflammatory response in myocardial infarction. *Cardiovascular research* 53, 31-47

283. Birdsall, H. H., Green, D. M., Trial, J., Youker, K. A., Burns, A. R., MacKay, C. R., LaRosa, G. J., Hawkins, H. K., Smith, C. W., and Michael, L. H. (1997) Complement C5a, TGF-β1, and MCP-1, in sequence, induce migration of monocytes into ischemic canine myocardium within the first one to five hours after reperfusion. *Circulation* 95, 684-692

284. Patzelt, J., Mueller, K., Breuning, S., Karathanos, A., Schleicher, R., Seizer, P., Gawaz, M., Langer, H., and Geisler, T. (2015) Expression of anaphylatoxin receptors on platelets in patients with coronary heart disease. *Atherosclerosis* 238, 289-295

285. Granger, C. B., Mahaffey, K. W., Weaver, W. D., Theroux, P., Hochman, J. S., Filloon, T. G., Rollins, S., Todaro, T. G., Nicolau, J. C., and Ruzyllo, W. (2003) Pexelizumab, an anti-C5 complement antibody, as adjunctive therapy to primary percutaneous coronary intervention in acute myocardial infarction the complement inhibition in myocardial infarction treated with angioplasty (COMMA) trial. *Circulation* 108, 1184-1190

286. Testa, L., Meco, M., Cirri, S., and Bedogni, F. (2011) pexelizumab and survival in cardiac surgery. *HSR proceedings in intensive care & cardiovascular anesthesia* 3, 23

287. Ekeløf, S., Jensen, S. E., Rosenberg, J., and Gögenur, I. (2014) Reduced Oxidative Stress in STEMI Patients Treated by Primary Percutaneous Coronary Intervention and with Antioxidant Therapy: A Systematic Review. *Cardiovascular drugs and therapy* 28, 173-181

288. Jolly, S., Kane, W., Bailie, M., Abrams, G., and Lucchesi, B. (1984) Canine myocardial reperfusion injury. Its reduction by the combined administration of superoxide dismutase and catalase. *Circulation Research* 54, 277-285

289. Bolli, R., Jeroudi, M. O., Patel, B. S., Aruoma, O. I., Halliwell, B., Lai, E. K., and McCay, P. B. (1989) Marked reduction of free radical generation and contractile dysfunction by antioxidant therapy begun at the time of reperfusion. Evidence that

myocardial" stunning" is a manifestation of reperfusion injury. *Circulation Research* 65, 607-622

290. Hemalatha, K., and Prince, P. (2014) Preventive Effects of Zingerone on Altered Lipid Peroxides and Nonenzymatic Antioxidants in the Circulation of Isoproterenol-Induced Myocardial Infarcted Rats. *Journal of biochemical and molecular toxicology*

291. Ma, Y., Yabluchanskiy, A., and Lindsey, M. L. (2013) Neutrophil roles in left ventricular remodeling following myocardial infarction. *Fibrogenesis Tissue Repair* 6

292. Horckmans, M., Drechsler, M., Weber, C., Sohnlein, O., and Steffens, S. (2014) P160Neutrophils modulate healing after myocardial infarction. *Cardiovascular research* 103, S28-S28

293. Liu, Y., Kalogeris, T., Wang, M., Zuidema, M. Y., Wang, Q., Dai, H., Davis, M. J., Hill, M. A., and Korthuis, R. J. (2012) Hydrogen sulfide preconditioning or neutrophil depletion attenuates ischemia-reperfusion-induced mitochondrial dysfunction in rat small intestine. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 302, G44-G54

294. Jolly, S. R., Kane, W. J., Hook, B. G., Abrams, G. D., Kunkel, S. L., and Lucchesi, B. R. (1986) Reduction of myocardial infarct size by neutrophil depletion: effect of duration of occlusion. *American heart journal* 112, 682-690

295. Buzas, E. I., György, B., Nagy, G., Falus, A., and Gay, S. (2014) Emerging role of extracellular vesicles in inflammatory diseases. *Nature Reviews Rheumatology* 10, 356-364

296. Chaput, N., and Théry, C. (2011) Exosomes: immune properties and potential clinical implementations. *Seminars in immunopathology*, pp. 419-440, Springer

297. Théry, C., Ostrowski, M., and Segura, E. (2009) Membrane vesicles as conveyors of immune responses. *Nature Reviews Immunology* 9, 581-593

298. Ratajczak, M., Kucia, M., Jadczyk, T., Greco, N., Wojakowski, W., Tendera, M., and Ratajczak, J. (2012) Pivotal role of paracrine effects in stem cell therapies in regenerative medicine: can we translate stem cell-secreted paracrine factors and microvesicles into better therapeutic strategies&quest. *Leukemia* 26, 1166-1173

References

299. del Conde, I., Shrimpton, C. N., Thiagarajan, P., and López, J. A. (2005) Tissue-factor–bearing microvesicles arise from lipid rafts and fuse with activated platelets to initiate coagulation. *Blood* 106, 1604-1611

300. Ellis, T. N., and Kuehn, M. J. (2010) Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiology and Molecular Biology Reviews* 74, 81-94

301. Lee, Y., Andaloussi, S. E., and Wood, M. J. (2012) Exosomes and microvesicles: extracellular vesicles for genetic information transfer and gene therapy. *Human molecular genetics* 21, R125-R134

302. Owens, A. P., and Mackman, N. (2011) Microparticles in hemostasis and thrombosis. *Circulation research* 108, 1284-1297

303. Tetta, C., Bruno, S., Fonsato, V., Deregibus, M. C., and Camussi, G. (2011) The role of microvesicles in tissue repair. *Organogenesis* 7, 105-115

304. Choi, D.-S., Lee, J., Go, G., Kim, Y.-K., and Gho, Y. S. (2013) Circulating Extracellular Vesicles in Cancer Diagnosis and Monitoring. *Molecular diagnosis & therapy* 17, 265-271

305. Rautou, P.-E., Vion, A.-C., Amabile, N., Chironi, G., Simon, A., Tedgui, A., and Boulanger, C. M. (2011) Microparticles, vascular function, and atherothrombosis. *Circulation research* 109, 593-606

306. Hao, P., Ren, Y., Alpert, A. J., and Sze, S. K. (2011) Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Mol Cell Proteomics* 10, O111 009381

307. Hao, P., Ren, Y., Datta, A., Tam, J. P., and Sze, S. K. (2014) Evaluation of the Effect of Trypsin Digestion Buffers on Artificial Deamidation. *Journal of proteome research*

308. Witwer, K. W., Buzas, E. I., Bemis, L. T., Bora, A., Lässer, C., Lötvall, J., and Nolte, E. N. (2013) Standardization of sample collection, isolation and analysis methods in extracellular vesicle research. *Journal of extracellular vesicles* 2

309. Ohno, S.-i., Ishikawa, A., and Kuroda, M. (2013) Roles of exosomes and microvesicles in disease pathogenesis. *Advanced drug delivery reviews* 65, 398-401

References

310. Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., Resing, K. A., and Ahn, N. G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular & cellular proteomics* 4, 1487-1502

311. Robbins, P. D., and Morelli, A. E. (2014) Regulation of immune responses by extracellular vesicles. *Nature Reviews Immunology* 14, 195-208

312. Niculescu, F., and Rus, H. (2004) The role of complement activation in atherosclerosis. *Immunologic research* 30, 73-80

313. Wang, J., and Frangogiannis, N. G. (2015) Repair of the Infarcted Myocardium. *Introduction to Translational Cardiovascular Research*, pp. 279-297, Springer

314. Ricklin, D., Hajishengallis, G., Yang, K., and Lambris, J. D. (2010) Complement: a key system for immune surveillance and homeostasis. *Nature immunology* 11, 785-797

315. Keizer, M. P., Pouw, R. B., Kamp, A. M., Patiwael, S., Marsman, G., Hart, M. H., Zeerleder, S., Kuijpers, T. W., and Wouters, D. (2015) TFPI inhibits lectin pathway of complement activation by direct interaction with MASP-2. *European journal of immunology* 45, 544-550

316. Jordan, J. E., Montalto, M. C., and Stahl, G. L. (2001) Inhibition of mannose-binding lectin reduces postischemic myocardial reperfusion injury. *Circulation* 104, 1413-1418

317. Schwaeble, W. J., Lynch, N. J., Clark, J. E., Marber, M., Samani, N. J., Ali, Y. M., Dudler, T., Parent, B., Lhotta, K., and Wallis, R. (2011) Targeting of mannan-binding lectin-associated serine protease-2 confers protection from myocardial and gastrointestinal ischemia/reperfusion injury. *Proceedings of the National Academy of Sciences* 108, 7523-7528

318. Meuwissen, M., Van Der Wal, A., Niessen, H., Koch, K., De Winter, R., Van der Loos, C., Rittersma, S., Chamuleau, S., Tijssen, J., and Becker, A. (2006) Colocalisation of intraplaque C reactive protein, complement, oxidised low density lipoprotein, and macrophages in stable and unstable angina and acute myocardial infarction. *Journal of clinical pathology* 59, 196-201

References

319. Berk, B. C., Weintraub, W. S., and Alexander, R. W. (1990) Elevation of C-reactive protein in "active" coronary artery disease. *The American journal of cardiology* 65, 168-172

320. Ridker, P. M., Glynn, R. J., and Hennekens, C. H. (1998) C-reactive protein adds to the predictive value of total and HDL cholesterol in determining risk of first myocardial infarction. *Circulation* 97, 2007-2011

321. Phillips, N., Waters, D., and Havel, R. (1993) Plasma lipoproteins and progression of coronary artery disease evaluated by angiography and clinical events. *Circulation* 88, 2762-2770

322. Rached, F., Lhomme, M., Camont, L., Gomes, F., Dauteille, C., Robillard, P., Santos, R. D., Lesnik, P., Serrano, C. V., and Chapman, M. J. (2015) Defective Functionality of Small, Dense HDL3 Subpopulations in ST Segment Elevation Myocardial Infarction: Relevance of Enrichment in Lysophosphatidylcholine, Phosphatidic Acid and Serum Amyloid A. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*

323. Hoogeveen, R. C., Gaubatz, J. W., Sun, W., Dodge, R. C., Crosby, J. R., Jiang, J., Couper, D., Virani, S. S., Kathiresan, S., and Boerwinkle, E. (2014) Small Dense Low-Density Lipoprotein-Cholesterol Concentrations Predict Risk for Coronary Heart Disease The Atherosclerosis Risk in Communities (ARIC) Study. *Arteriosclerosis, thrombosis, and vascular biology* 34, 1069-1077

324. Jørgensen, A. B., Frikke-Schmidt, R., Nordestgaard, B. G., and Tybjærg-Hansen, A. (2014) Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *New England Journal of Medicine* 371, 32-41

325. Cohorts, D. (2014) Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *New Engl J Medicine* 371, 22-31

326. Rassart, E., Bedirian, A., Do Carmo, S., Guinard, O., Sirois, J., Terrisse, L., and Milne, R. (2000) Apolipoprotein d. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 1482, 185-198

327. Perdomo, G., and Dong, H. H. (2009) Apolipoprotein D in lipid metabolism and its functional implication in atherosclerosis and aging. *Aging (Albany NY)* 1, 17

References

328. Muffat, J., and Walker, D. W. (2010) Apolipoprotein D: an overview of its role in aging and age-related diseases. *Cell Cycle* 9, 269-273

329. Ganfornina, M. D., Do Carmo, S., Martínez, E., Tolivia, J., Navarro, A., Rassart, E., and Sanchez, D. (2010) ApoD, a glia-derived apolipoprotein, is required for peripheral nerve functional integrity and a timely response to injury. *Glia* 58, 1320-1334

330. Muffat, J., Walker, D. W., and Benzer, S. (2008) Human ApoD, an apolipoprotein up-regulated in neurodegenerative diseases, extends lifespan and increases stress resistance in Drosophila. *Proceedings of the National Academy of Sciences* 105, 7088-7093

331. Vaisar, T., Pennathur, S., Green, P. S., Gharib, S. A., Hoofnagle, A. N., Cheung, M. C., Byun, J., Vuletic, S., Kassim, S., and Singh, P. (2007) Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL. *Journal of Clinical Investigation* 117, 746

332. Kontush, A., and Chapman, M. J. (2006) Functionally defective high-density lipoprotein: a new therapeutic target at the crossroads of dyslipidemia, inflammation, and atherosclerosis. *Pharmacological reviews* 58, 342-374

333. von Hundelshausen, P., and Weber, C. (2007) Platelets as immune cells bridging inflammation and cardiovascular disease. *Circulation research* 100, 27-40

334. Liu, Y., Gao, X.-M., Fang, L., Jennings, N. L., Su, Y., Xu, Q., Samson, A. L., Kiriazis, H., Wang, X.-F., and Shan, L. (2011) Novel role of platelets in mediating inflammatory responses and ventricular rupture or remodeling following myocardial infarction. *Arteriosclerosis, thrombosis, and vascular biology* 31, 834-841

335. Bombeli, T., Schwartz, B. R., and Harlan, J. M. (1998) Adhesion of activated platelets to endothelial cells: evidence for a GPIIbIIIa-dependent bridging mechanism and novel roles for endothelial intercellular adhesion molecule 1 (ICAM-1), αvβ3 integrin, and GPIbα. *The Journal of experimental medicine* 187, 329-339

336. Bray, P. F., Howard, T. D., Vittinghoff, E., Sane, D. C., and Herrington, D. M. (2007) Effect of genetic variations in platelet glycoproteins Ibα and VI on the

risk for coronary heart disease events in postmenopausal women taking hormone therapy. *Blood* 109, 1862-1869

337. Coppinger, J., and Maguire, P. (2007) Insights into the platelet releasate. *Current pharmaceutical design* 13, 2640-2646

338. Wilhelmsen, L., Svärdsudd, K., Korsan-Bengtsen, K., Larsson, B., Welin, L., and Tibblin, G. (1984) Fibrinogen as a risk factor for stroke and myocardial infarction. *New England Journal of Medicine* 311, 501-505

339. Kranendonk, M., de Kleijn, D., Kalkhoven, E., Kanhai, D. A., Uiterwaal, C., Van der Graaf, Y., Pasterkamp, G., and Visseren, F. (2014) Extracellular vesicle markers in relation to obesity and metabolic complications in patients with manifest cardiovascular disease. *Cardiovasc Diabetol* 13, 37

340. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics* 25, 25-29

341. Hughes, B. (2012) Studies of aging, metabolism and the age-dependent disease atherosclerosis in mouse mutants with altered mitochondrial function. MCGILL UNIVERSITY

342. Hashimoto, J., and Ito, S. (2009) Some mechanical aspects of arterial aging: physiological overview based on pulse wave analysis. *Therapeutic advances in cardiovascular disease* 3, 367-378

343. Truscott, R. J. W. (2010) Are ancient proteins responsible for the age-related decline in health and fitness? *Rejuvenation research* 13, 83-89

344. Nicholson, D. H., Harkness, D. R., Benson, W. E., and Peterson, C. M. (1976) Cyanate-induced cataracts in patients with sickle-cell hemoglobinopathies. *Archives of ophthalmology* 94, 927-930

345. Gillery, P., and Jaisson, S. (2013) Usefulness of non-enzymatic post-translational modification derived products (PTMDPs) as biomarkers of chronic diseases. *Journal of proteomics* 92, 228-238

References

346. Kalim, S., Karumanchi, S. A., Thadhani, R. I., and Berg, A. H. (2014) Protein carbamylation in kidney disease: pathogenesis and clinical implications. *American Journal of Kidney Diseases* 64, 793-803

347. Hawkins, C. L. (2014) Role of cyanate in the induction of vascular dysfunction during uremia: more than protein carbamylation&quest. *Kidney international* 86, 875-877

348. Shi, J., Knevel, R., Suwannalai, P., van der Linden, M. P., Janssen, G. M., van Veelen, P. A., Levarht, N. E., van der Helm-van, A. H., Cerami, A., and Huizinga, T. W. (2011) Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proceedings of the National Academy of Sciences* 108, 17372-17377

349. Mydel, P., Wang, Z., Brisslert, M., Hellvard, A., Dahlberg, L. E., Hazen, S. L., and Bokarewa, M. (2010) Carbamylation-dependent activation of T cells: a novel mechanism in the pathogenesis of autoimmune arthritis. *The Journal of Immunology* 184, 6882-6890

350. Jaisson, S., Delevallée-Forte, C., Touré, F., Rieu, P., Garnotel, R., and Gillery, P. (2007) Carbamylated albumin is a potent inhibitor of polymorphonuclear neutrophil respiratory burst. *FEBS letters* 581, 1509-1513

351. Balion, C. M., Draisey, T. F., and Thibert, R. J. (1998) Carbamylated hemoglobin and carbamylated plasma protein in hemodialyzed patients. *Kidney international* 53, 488-495

352. Nicholls, S. J., and Hazen, S. L. (2009) Myeloperoxidase, modified lipoproteins, and atherogenesis. *Journal of lipid research* 50, S346-S351

353. Holzer, M., Gauster, M., Pfeifer, T., Wadsack, C., Fauler, G., Stiegler, P., Koefeler, H., Beubler, E., Schuligoi, R., and Heinemann, A. (2011) Protein carbamylation renders high-density lipoprotein dysfunctional. *Antioxidants & redox signaling* 14, 2337-2346

354. Lowenthal, M. S., Mehta, A. I., Frogale, K., Bandle, R. W., Araujo, R. P., Hood, B. L., Veenstra, T. D., Conrads, T. P., Goldsmith, P., and Fishman, D. (2005) Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clinical chemistry* 51, 1933-1945

References

355. Diamandis, E. P. (2006) Peptidomics for cancer diagnosis: present and future. *Journal of proteome research* 5, 2079-2082

356. Lopez, M. F., Mikulskis, A., Kuzdzal, S., Golenko, E., Petricoin, E. F., Liotta, L. A., Patton, W. F., Whiteley, G. R., Rosenblatt, K., and Gurnani, P. (2007) A novel, high-throughput workflow for discovery and identification of serum carrier protein-bound peptide biomarker candidates in ovarian cancer samples. *Clinical chemistry* 53, 1067-1074

357. Marshall, J., Kupchak, P., Zhu, W., Yantha, J., Vrees, T., Furesz, S., Jacks, K., Smith, C., Kireeva, I., and Zhang, R. (2003) Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *Journal of proteome research* 2, 361-372

358. Kollipara, L., and Zahedi, R. P. (2013) Protein carbamylation: in vivo modification or in vitro artefact? *Proteomics* 13, 941-944

359. McCarthy, J., Hopwood, F., Oxley, D., Laver, M., Castagna, A., Righetti, P. G., Williams, K., and Herbert, B. (2003) Carbamylation of Proteins in 2-D Electrophoresis Myth or Reality? *Journal of proteome research* 2, 239-242

360. Wu, C. C., MacCoss, M. J., Howell, K. E., and Yates, J. R. (2003) A method for the comprehensive proteomic analysis of membrane proteins. *Nature biotechnology* 21, 532-538

361. Mayampurath, A. M., Jaitly, N., Purvine, S. O., Monroe, M. E., Auberry, K. J., Adkins, J. N., and Smith, R. D. (2008) DeconMSn: a software tool for accurate parent ion monoisotopic mass determination for tandem mass spectra. *Bioinformatics* 24, 1021-1023

362. Scherl, A., Tsai, Y. S., Shaffer, S. A., and Goodlett, D. R. (2008) Increasing information from shotgun proteomic data by accounting for misassigned precursor ion masses. *Proteomics* 8, 2791-2797

363. Hao, P., Ren, Y., Alpert, A. J., and Sze, S. K. (2011) Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Molecular & Cellular Proteomics* 10, O111. 009381

364. Nepomuceno, A. I., Gibson, R. J., Randall, S. M., and Muddiman, D. C. (2013) Accurate Identification of Deamidated Peptides in Global Proteomics Using

References

a Quadrupole Orbitrap Mass Spectrometer. *Journal of proteome research* 13, 777-785

365. Lin, M.-F., Williams, C., Murray, M. V., Conn, G., and Ropp, P. A. (2004) Ion chromatographic quantification of cyanate in urea solutions: estimation of the efficiency of cyanate scavengers for use in recombinant protein manufacturing. *Journal of Chromatography B* 803, 353-362

366. Hörkkö, S., Huttunen, K., Kervinen, K., and KESÄNIEMI, Y. A. (1994) Decreased clearance of uraemic and mildly carbamylated low-density lipoprotein. *European journal of clinical investigation* 24, 105-113

367. Asci, G., Basci, A., Shah, S. V., Basnakian, A., Toz, H., Ozkahya, M., Duman, S., and Ok, E. (2008) Carbamylated low-density lipoprotein induces proliferation and increases adhesion molecule expression of human coronary artery smooth muscle cells. *Nephrology* 13, 480-486

368. Holzer, M., Zangger, K., El-Gamal, D., Binder, V., Curcic, S., Konya, V., Schuligoi, R., Heinemann, A., and Marsche, G. (2012) Myeloperoxidase-derived chlorinating species induce protein carbamylation through decomposition of thiocyanate and urea: novel pathways generating dysfunctional high-density lipoprotein. *Antioxidants & redox signaling* 17, 1043-1052

369. Drechsler, C., Kalim, S., Wenger, J. B., Suntharalingam, P., Hod, T., Thadhani, R. I., Karumanchi, S. A., Wanner, C., and Berg, A. H. (2015) Protein carbamylation is associated with heart failure and mortality in diabetic patients with end-stage renal disease. *Kidney international*

370. Koeth, R. A., Kalantar-Zadeh, K., Wang, Z., Fu, X., Tang, W. W., and Hazen, S. L. (2013) Protein carbamylation predicts mortality in ESRD. *Journal of the American Society of Nephrology* 24, 853-861

371. Burska, A. N., Hunt, L., Boissinot, M., Strollo, R., Ryan, B. J., Vital, E., Nissim, A., Winyard, P. G., Emery, P., and Ponchel, F. (2014) Autoantibodies to posttranslational modifications in rheumatoid arthritis. *Mediators of inflammation* 2014

References

372. Fish, J. C., Remmers Jr, A. R., Lindley, J. D., and Sarles, H. E. (1972) Albumin kinetics and nutritional rehabilitation in the unattended home-dialysis patient. *New England Journal of Medicine* 287, 478-481

373. Shaykh, M., Pegoraro, A., Mo, W., Arruda, J., Dunea, G., and Singh, A. K. (1999) Carbamylated proteins activate glomerular mesangial cells and stimulate collagen deposition. *Journal of Laboratory and Clinical Medicine* 133, 302-308

374. Koro, C., Bielecka, E., Dahl-Knudsen, A., Enghild, J. J., Scavenius, C., Brun, J. G., Binder, V., Hellvard, A., Bergum, B., and Jonsson, R. (2014) Carbamylation of immunoglobulin abrogates activation of the classical complement pathway. *European journal of immunology* 44, 3403-3412

375. Wynckel, A., Randoux, C., Millart, H., Desroches, C., Gillery, P., Canivet, E., and Chanard, J. (2000) Kinetics of carbamylated haemoglobin in acute renal failure. *Nephrology Dialysis Transplantation* 15, 1183-1188

376. Apostolov, E. O., Shah, S. V., Ok, E., and Basnakian, A. G. (2005) Quantification of carbamylated LDL in human sera by a new sandwich ELISA. *Clinical chemistry* 51, 719-728

377. Bernard, D. R., Langlois, M. R., Delanghe, J. R., and De Buyzere, M. L. (1997) Evolution of haptoglobin concentration in serum during the early phase of acute myocardial infarction. *Clinical Chemistry and Laboratory Medicine* 35, 85-88

378. Wassell, J. (1999) Haptoglobin: function and polymorphism. *Clinical laboratory* 46, 547-552

379. Sadrzadeh, S. H., and Bozorgmehr, J. (2004) Haptoglobin phenotypes in health and disorders. *American Journal of Clinical Pathology. Pathology Patterns Reviews.* 121, S97-S104

380. Eaton, J. W., Brandt, P., and Lee, J. (1982) Haptoglobin: a natural bacteriostat. *Science* 215, 691-693

381. Maeda, S., Miyauchi, T., Iemitsu, M., Tanabe, T., Goto, K., Yamaguchi, I., and Matsuda, M. (2004) Endothelin receptor antagonist reverses decreased NO system in the kidney in vivo during exercise. *American Journal of Physiology-Endocrinology and Metabolism* 286, E609-E614

References

382. Jue, D.-M., Shim, B.-S., and Kang, Y.-S. (1983) Inhibition of prostaglandin synthase activity of sheep seminal vesicular gland by human serum haptoglobin. *Molecular and cellular biochemistry* 51, 141-147

383. Cid, M. C., Grant, D. S., Hoffman, G. S., Auerbach, R., Fauci, A. S., and Kleinman, H. K. (1993) Identification of haptoglobin as an angiogenic factor in sera from patients with systemic vasculitis. *Journal of Clinical Investigation* 91, 977

384. Fournier, T., Medjoubi-N, N., and Porquet, D. (2000) Alpha-1-acid glycoprotein. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 1482, 157-171

385. Vasson, M. P., Roch-Arveiller, M., Couderc, R., Baguet, J. C., and Raichvarg, D. (1994) Effects of alpha-1 acid glycoprotein on human polymorphonuclear neutrophils: influence of glycan microheterogeneity. *Clinica chimica acta* 224, 65-71

386. Snyder, S., and Coodley, E. L. (1976) Inhibition of Platelet Aggregation by a1-Acid Glycoprotein. *Archives of internal medicine* 136, 778-781

387. Bories, P., Guenounou, M., Feger, J., Kodari, E., Agneray, J., and Durand, G. (1987) Human α 1-acid glycoprotein-exposed macrophages release interleukin 1 inhibitory activity. *Biochemical and biophysical research communications* 147, 710-715

388. Maeda, H., Murakami, O., Kann, M., and Yamane, I. (1980) The growth-stimulating effect of α1-acid glycoprotein in cells in culture. *Experimental Biology and Medicine* 163, 223-227

389. Chasteen, N. D. (1977) Human serotransferrin: structure and function. *Coordination Chemistry Reviews* 22, 1-36

390. Macedo, M. F., and Sousa, M. d. (2008) Transferrin and the transferrin receptor: of magic bullets and other concerns. *Inflammation & Allergy-Drug Targets (Formerly Current Drug Targets-Inflammation & Allergy)* 7, 41-52

391. Aisen, P. (2004) Transferrin receptor 1. *The international journal of biochemistry & cell biology* 36, 2137-2143

References

392. Hirata, T., Bitterman, P. B., Mornex, J.-F., and Crystal, R. (1986) Expression of the transferrin receptor gene during the process of mononuclear phagocyte maturation. *The Journal of Immunology* 136, 1339-1345

393. Besancon, F., Bourgeade, M., and Testa, U. (1985) Inhibition of transferrin receptor expression by interferon-alpha in human lymphoblastoid cells and mitogen-induced lymphocytes. *Journal of Biological Chemistry* 260, 13074-13080

394. Neckers, L. M., and Cossman, J. (1984) Transferrin receptor induction in mitogen-stimulated human T lymphocytes is required for DNA synthesis and cell division and is regulated by interleukin-2 (TCGF). *Thymic Hormones and Lymphokines*, pp. 383-394, Springer

395. Devine, D. V., and Schubert, P. (2011) Proteomic applications in blood transfusion: working the jigsaw puzzle. *Vox Sang* 100, 84-91

396. Pan, S., Chen, R., Aebersold, R., and Brentnall, T. A. (2011) Mass spectrometry based glycoproteomics—from a proteomics perspective. *Molecular & Cellular Proteomics* 10, R110. 003251

397. Li, H., and Liu, Z. (2012) Recent advances in monolithic column-based boronate-affinity chromatography. *TrAC Trends in Analytical Chemistry* 37, 148-161

398. Alvarez-Manilla, G., Atwood, J., 3rd, Guo, Y., Warren, N. L., Orlando, R., and Pierce, M. (2006) Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *J Proteome Res* 5, 701-708

399. Alpert, A. J., Shukla, M., Shukla, A. K., Zieske, L. R., Yuen, S. W., Ferguson, M. A., Mehlert, A., Pauly, M., and Orlando, R. (1994) Hydrophilic-interaction chromatography of complex carbohydrates. *Journal of chromatography. A* 676, 191-122

400. Calvano, C. D., Zambonin, C. G., and Jensen, O. N. (2008) Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry. *J Proteomics* 71, 304-317

References

401. Lewandrowski, U., Lohrig, K., Zahedi, R., Walter, D., and Sickmann, A. (2008) Glycosylation site analysis of human platelets by electrostatic repulsion hydrophilic interaction chromatography. *Clin. Proteomics* 4, 25-36

402. Zhang, H., Guo, T., Li, X., Datta, A., Park, J. E., Yang, J., Lim, S. K., Tam, J. P., and Sze, S. K. (2010) Simultaneous characterization of glyco- and phosphoproteomes of mouse brain membrane proteome with electrostatic repulsion hydrophilic interaction chromatography. *Mol Cell Proteomics* 9, 635-647

403. Leroyer, A., Tedgui, A., and Boulanger, C. (2008) Role of microparticles in atherothrombosis. *Journal of Internal Medicine* 263, 528-537

404. Viaud, S., Terme, M., Flament, C., Taieb, J., André, F., Novault, S., Escudier, B., Robert, C., Caillat-Zucman, S., and Tursz, T. (2009) Dendritic cell-derived exosomes promote natural killer cell activation and proliferation: a role for NKG2D ligands and IL-15Rα. *PLoS ONE* 4, e4942

405. Ardoin, S., Shanahan, J., and Pisetsky, D. (2007) The role of microparticles in inflammation and thrombosis. *Scandinavian journal of immunology* 66, 159-165

406. Sabatier, F., Darmon, P., Hugel, B., Combes, V., Sanmarco, M., Velut, J. G., Arnoux, D., Charpiot, P., Freyssinet, J. M., and Oliver, C. (2002) Type 1 and type 2 diabetic patients display different patterns of cellular microparticles. *Diabetes* 51, 2840-2845

407. Chen, C., Skog, J., Hsu, C.-H., Lessard, R. T., Balaj, L., Wurdinger, T., Carter, B. S., Breakefield, X. O., Toner, M., and Irimia, D. (2010) Microfluidic isolation and transcriptome analysis of serum microvesicles. *Lab on a chip* 10, 505-511

408. Zhang, H., Guo, T., Li, X., Datta, A., Park, J. E., Yang, J., Lim, S. K., Tam, J. P., and Sze, S. K. (2010) Simultaneous characterization of glyco-and phosphoproteomes of mouse brain membrane proteome with electrostatic repulsion hydrophilic interaction chromatography. *Molecular & Cellular Proteomics* 9, 635-647

409. Bhatia, V. N., Perlman, D. H., Costello, C. E., and McComb, M. E. (2009) Software tool for researching annotations of proteins: open-source protein annotation software with data visualization. *Anal Chem* 81, 9819-9823

References

410. Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., and Lewis, S. (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics* 25, 288-289

411. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* 37, 1-13

412. van der Pol, E., Boing, A. N., Harrison, P., Sturk, A., and Nieuwland, R. (2012) Classification, functions, and clinical relevance of extracellular vesicles. *Pharmacological Reviews* 64, 676-705

413. Ohno, S., Ishikawa, A., and Kuroda, M. (2013) Roles of exosomes and microvesicles in disease pathogenesis. *Adv Drug Deliv Rev* 65, 398-401

414. Beck, F., Lewandrowski, U., Wiltfang, M., Feldmann, I., Geiger, J., Sickmann, A., and Zahedi, R. P. (2011) The good, the bad, the ugly: validating the mass spectrometric analysis of modified peptides. *Proteomics* 11, 1099-1109

415. Wright, H. T. (1991) Sequence and structure determinants of the nonenzymatic deamidation of asparagine and glutamine residues in proteins. *Protein Eng* 4, 283-294

416. Robinson, N. E. (2002) Protein deamidation. *Proc Natl Acad Sci U S A* 99, 5283-5288

417. Yang, W., Laeyendecker, O., Wendel, S. K., Zhang, B., Sun, S., Zhou, J.-Y., Ao, M., Moore, R. D., Jackson, J. B., and Zhang, H. (2014) Glycoproteomic Study Reveals Altered Plasma Proteins Associated with HIV Elite Suppressors. *Theranostics* 4, 1153

418. Drake, P. M., Schilling, B., Niles, R. K., Braten, M., Johansen, E., Liu, H., Lerch, M., Sorensen, D. J., Li, B., Allen, S., Hall, S. C., Witkowska, H. E., Regnier, F. E., Gibson, B. W., and Fisher, S. J. (2011) A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. *Anal Biochem* 408, 71-85

419. Mathivanan, S., Fahner, C. J., Reid, G. E., and Simpson, R. J. (2012) ExoCarta 2012: database of exosomal proteins, RNA and lipids. *Nucleic Acids Res* 40, D1241-1244

References

420. Bastos-Amador, P., Royo, F., Gonzalez, E., Conde-Vancells, J., Palomo-Diez, L., Borras, F. E., and Falcon-Perez, J. M. (2012) Proteomic analysis of microvesicles from plasma of healthy donors reveals high individual variability. *J Proteomics* 75, 3574-3584

421. Looze, C., Yui, D., Leung, L., Ingham, M., Kaler, M., Yao, X., Wu, W. W., Shen, R.-F., Daniels, M. P., and Levine, S. J. (2009) Proteomic profiling of human plasma exosomes identifies PPARγ as an exosome-associated protein. *Biochemical and biophysical research communications* 378, 433-438

422. Hortin, G. L., Sviridov, D., and Anderson, N. L. (2008) High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clin Chem* 54, 1608-1616

423. Hortin, G. L., and Sviridov, D. (2010) The dynamic range problem in the analysis of the plasma proteome. *J Proteomics* 73, 629-636

424. Percy, A. J., Chambers, A. G., Yang, J., and Borchers, C. H. (2013) Multiplexed MRM-based quantitation of candidate cancer biomarker proteins in undepleted and non-enriched human plasma. *Proteomics* 13, 2202-2215

425. Percy, A. J., Chambers, A. G., Yang, J., Domanski, D., and Borchers, C. H. (2012) Comparison of standard- and nano-flow liquid chromatography platforms for MRM-based quantitation of putative plasma biomarker proteins. *Anal Bioanal Chem* 404, 1089-1101

426. Farrah, T., Deutsch, E. W., Omenn, G. S., Campbell, D. S., Sun, Z., Bletz, J. A., Mallick, P., Katz, J. E., Malmstrom, J., Ossola, R., Watts, J. D., Lin, B., Zhang, H., Moritz, R. L., and Aebersold, R. (2011) A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Mol Cell Proteomics* 10, M110 006353

427. Udby, L., Cowland, J. B., Johnsen, A. H., Sorensen, O. E., Borregaard, N., and Kjeldsen, L. (2002) An ELISA for SGP28/CRISP-3, a cysteine-rich secretory protein in human neutrophils, plasma, and exocrine secretions. *J Immunol Methods* 263, 43-55

References

428. Nielsen, M. J., Petersen, S. V., Jacobsen, C., Oxvig, C., Rees, D., Moller, H. J., and Moestrup, S. K. (2006) Haptoglobin-related protein is a high-affinity hemoglobin-binding plasma protein. *Blood* 108, 2846-2849

429. Ishikawa, Y., Akishima-Fukasawa, Y., Ito, K., Akasaka, Y., Tanaka, M., Shimokawa, R., Kimura-Matsumoto, M., Morita, H., Sato, S., Kamata, I., and Ishii, T. (2007) Lymphangiogenesis in myocardial remodelling after infarction. *Histopathology* 51, 345-353

430. Park, J. H., Yoon, J. Y., Ko, S. M., Jin, S. A., Kim, J. H., Cho, C. H., Kim, J. M., Lee, J. H., Choi, S. W., Seong, I. W., and Jeong, J. O. (2011) Endothelial progenitor cell transplantation decreases lymphangiogenesis and adverse myocardial remodeling in a mouse model of acute myocardial infarction. *Exp Mol Med* 43, 479-485

431. Mayne, J., Starr, A. E., Ning, Z., Chen, R., Chiang, C. K., and Figeys, D. (2014) Fine tuning of proteomic technologies to improve biological findings: advancements in 2011-2013. *Anal Chem* 86, 176-195

432. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* 4, 44-57

433. Tonack, S., Aspinall-O'Dea, M., Neoptolemos, J. P., and Costello, E. (2009) Pancreatic cancer: proteomic approaches to a challenging disease. *Pancreatology* 9, 567-576

434. Fanayan, S., Hincapie, M., and Hancock, W. S. (2012) Using lectins to harvest the plasma/serum glycoproteome. *Electrophoresis* 33, 1746-1754

435. Hao, P., Guo, T., and Sze, S. K. (2011) Simultaneous analysis of proteome, phospho- and glycoproteome of rat kidney tissue with electrostatic repulsion hydrophilic interaction chromatography. *PLoS One* 6, e16884

436. Hao, P., Qian, J., Ren, Y., and Sze, S. K. (2011) Electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus strong cation exchange (SCX) for fractionation of iTRAQ-labeled peptides. *Journal of proteome research* 10, 5568-5574

References

437. Percy, A. J., Chambers, A. G., Smith, D. S., and Borchers, C. H. (2012) Standardized protocols for quality control of MRM-based plasma proteomic workflows. *Journal of proteome research* 12, 222-233

438. Farrah, T., Deutsch, E. W., Omenn, G. S., Campbell, D. S., Sun, Z., Bletz, J. A., Mallick, P., Katz, J. E., Malmström, J., and Ossola, R. (2011) A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Molecular & Cellular Proteomics* 10, M110. 006353

439. Percy, A. J., Chambers, A. G., Yang, J., and Borchers, C. H. (2013) Multiplexed MRM-based quantitation of candidate cancer biomarker proteins in undepleted and non-enriched human plasma. *Proteomics* 13, 2202-2215

440. Fang, R., Elias, D. A., Monroe, M. E., Shen, Y., Mcintosh, M., Wang, P., Goddard, C. D., Callister, S. J., Moore, R. J., and Gorby, Y. A. (2006) Differential label-free quantitative proteomic analysis of Shewanella oneidensis cultured under aerobic and suboxic conditions by accurate mass and time tag approach. *Molecular & Cellular Proteomics* 5, 714-725

Appendix A

Appendix A: Publications

1. **Cheow, E. S. H.**, Sim, K. H., de Kleijn, D., Lee, C. N., Sorokin, V., and Sze, S. K. (2015) Simultaneous enrichment of plasma soluble and extracellular vesicular glycoproteins using prolonged ultracentrifugation-ERLIC approach. *Molecular & Cellular Proteomics*, mcp. O114. 046391

2. Dutta, B., Ren, Y., Hao, P., Sim, K. H., **Cheow, E.**, Adav, S., Tam, J. P., and Sze, S. K. (2014) Profiling of the chromatin-associated proteome identifies HP1BP3 as a novel regulator of cell cycle progression. *Molecular & Cellular Proteomics* 13, 2183-2197

3. Ren, Y., Hao, P., Dutta, B., **Cheow, E. S. H.**, Sim, K. H., Gan, C. S., Lim, S. K., and Sze, S. K. (2013) Hypoxia modulates A431 cellular pathways association to tumor radioresistance and enhanced migration revealed by comprehensive proteomic and functional studies. *Molecular & Cellular Proteomics* 12, 485-498

4. Hao, P., Qian, J., Dutta, B., **Cheow, E. S. H.**, Sim, K. H., Meng, W., Adav, S. S., Alpert, A., and Sze, S. K. (2012) Enhanced separation and characterization of deamidated peptides with RP-ERLIC-based multidimensional chromatography coupled with tandem mass spectrometry. *Journal of proteome research* 11, 1804-1811

5. Adav, S. S., Ravindran, A., **Cheow, E. S. H.**, and Sze, S. K. (2012) Quantitative proteomic analysis of secretome of microbial consortium during saw dust utilization. *Journal of proteomics* 75, 5590-5603

6. Adav, S. S., **Cheow, E. S. H.**, Ravindran, A., Dutta, B., and Sze, S. K. (2012) Label free quantitative proteomic analysis of secretome by Thermobifida fusca on different lignocellulosic biomass. *Journal of proteomics* 75, 3694-3706

Appendix B

Appendix B: Conference Presentation

Poster

1. 62nd ASMS Conference on Mass Spectrometry & Allied Topics 2014, June 15 – 19, 2014

Baltimore Convention Center, One West Pratt Street, Baltimore, Maryland 21201

Title: Cheow, E. S. H., P.L Hao, Sze, S. K. Ultracentrifugation-based glycoproteomic: Approach for discovery of plasma glycoprotein markers.

2. 2nd SOCRATES Scientific Meeting 2014, November 3 – 4, 2014

Matrix, A*STAR Biopolis, Singapore.

Title: Cheow, E. S. H., P.L Hao, Sze, S. K. Ultracentrifugation-based glycoproteomic: Approach for discovery of plasma glycoprotein markers

3. Inaugural SOCRATES Scientific Meeting 2013, October 24 – 25, 2013

Matrix, A*STAR Biopolis, Singapore.

Title: Cheow, E. S. H., and Sze, S. K. Cancer Biomarker Discovery from Sera of Cancer Patients

4. International Conference on Natural Products and Health 2013, September 5 – 7, 2013.

Lecture Theatre 7, North Spine, Nanyang Technological University (NTU), Singapore.

Title: Cheow, E. S. H., and Sze, S. K. Proteomics Discovery of Novel Biomarkers for Cardiovascular Diseases.

Invited talk

1. Singapore-Bilkent Joint Workshop on Molecular Biology, May 25 – 26, 2015

Bilkent University, Mithat Çoruh Amphitheatre, Bilkent, 6800 Ankara, Turkey

Title: Cheow, E. S. H., Sim, K. H., de Kleijn, D., Lee, C. N., Sorokin, V., and Sze, S. K. Simultaneous enrichment of plasma soluble and extracellular vesicular glycoproteins using prolonged ultracentrifugation-ERLIC approach.