DISSERTATION

INVESTIGATION OF MOLECULAR EFFECTS OF THE SOY-DERIVED PHYTOESTROGEN GENISTEIN ON CARDIOMYOCYTES BY PROTEOMIC ANALYSIS

Submitted by

Zeyu Sun

Department of Chemical and Biological Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2011

Doctoral Committee:

Advisor: Kenneth Reardon
Co-Advisor: Karyn Hamilton

Christopher Orton
Bradley Reisfeld

ABSTRACT

INVESTIGATION OF MOLECULAR EFFECTS OF THE SOY-DERIVED PHYTOESTROGEN GENISTEIN ON CARDIOMYOCYTES BY PROTEOMIC ANALYSIS

The soy-derived phytoestrogen genistein (GEN) has received attention for its potential to benefit the cardiovascular system by providing protection to cardiomyocytes against pathophysiological stresses. Although GEN is a well-known estrogen receptor (ER) agonist and a non-specific tyrosine kinase inhibitor, current understanding of the complex cellular and molecular effects of GEN in cardiomyocytes is still incomplete. The overall goal of this dissertation is to use high throughput proteomics methodologies to better understand the molecular action of GEN in cardiomyocytes and to identify proteins and pathways that respond to GEN treatment. The first study of this project focused on the concentration-dependent proteome changes in cultured HL-1 cardiomyocytes due to GEN treatments. Proteins from HL-1 cardiomyocytes treated with 1 μM and 50 μM GEN were prefractionated into hydrophilic and hydrophobic protein fractions and were analyzed by two-dimensional electrophoresis followed by protein identification using tandem mass spectrometry (MS). In total, 25 and 62 differential expressed proteins were identified in response to 1 μM and 50 μM of GEN treatment, respectively. These results suggest that 1 μM GEN enhanced the expression of heat shock proteins and anti-apoptotic proteins, while 50 μM GEN down-regulated glycolytic and antioxidant enzymes, potentially making cardiomyocytes more susceptible to energy depletion and apoptosis. The second study, employing a two-dimensional liquid chromatography and tandem MS shotgun proteomics workflow, was carried out to dissect the cellular functions changed in cardiomyocytes by ER-dependent or ER-independent actions of GEN. In this study,

primary cardiomyocytes isolated from male adult SD rats were treated with 10 μM GEN without or with 10 μM ER antagonist ICI 182,780 (ERA) before proteomics comparison. A total of 14 and 15 proteins were found differentially expressed in response to the GEN, and the GEN+ERA treatment, respectively. Cellular functions such as glucose and fatty acid metabolism and cardioprotection were found to be modulated by GEN in an ER-dependent fashion, while proteins involved with steroidogenesis and estrogen signaling were identified as novel effectors of GEN via ER-independent actions. In this study, a consensus-iterative searching strategy was also developed to enhance the sensitivity of the shotgun proteomic approach. In the last study, an attempt to explore the response to a GEN stimulus in the signaling pathways, we developed a phosphopeptide enrichment method to assist the detection of protein phosphorylation in a complex peptide mixture. The quantitative performance of a sequential immobilized metal affinity chromatography (SIMAC) protocol was evaluated. We further conducted a preliminary application of this protocol in a large-scale, quantitative, label-free phosphoproteomics study to explore the alterations of protein phosphorylation patterns due to ER-independent GEN action in the SD rat cardiomyocytes. This project demonstrates the usefulness of proteomics methodologies to screen novel molecular targets influenced by GEN in cardiomyocytes. This is also the first investigation of the complex cellular impact of this soy-derived phytoestrogen in cardiomyocytes via a systems biology perspective.

ACKNOWLEDGEMENT

I would like to express my gratitude to all my committee members, Dr. Kenneth Reardon, Dr. Karyn Hamilton, Dr. Christopher Orton and Dr. Bradley Reisfeld who gave me the possibility to complete this dissertation and all the assistance along the journey. I owe my deepest gratitude to my advisor Dr. Reardon and co-advisor Dr. Hamilton for being extraordinary mentors for me by sharing their academic expertise and experience with me, and by providing guidance and encouragement throughout my Ph.D. work. It is an honor for me to work with them these years.

I like to show my gratitude to all helping hands of all past and present members in Dr.Reardon's lab and Dr. Hamilton's lab. I also want to thank the following people who provided critical technical assistance and constructive advices: Dr. Carla Lacerda, Dr. Nichole Reisdorph, Dr. Jessica Prenni, Dr. Ann Hess, Dr. Andrey Ptitsyn, Dr. William Claycomb, Brian Cranmer, Delian Yang, Laurie Biela and Reuland Nellie. This work cannot be accomplished without their generous assistance and valuable input. Two outstanding undergraduate students have actively participated in my research: Kathryn Knopinski and Caitlin Mitchel. I want to use this opportunity to thank their hard working and their contributions to my research progress.

Finally, I would like to say thank you to all my family and friends who have encouraged me during these years. Especially, I would like to give my special thanks to my mother Jiali Zhou, father Weiqiang Sun, and my fiancée Bing Chi for their patient support behind me to complete this work.

Zeyu Sun

At Dept. of Chemical and Biological Engineering

Sep-11-2011

TABLE OF CONTENTS

LIST OF APPENDICES

CHAPTER 1

BACKGROUND AND OBJECTIVES

1. Introduction

1.1. Soy phytoestrogens and cardiovascular health

Cardiovascular diseases (CVDs) are the leading cause of death of both men and women in the United States and many other western industrialized countries. Nearly 30 percent of U.S. mortality is related to CVDs such as heart failure, myocardial infarction, coronary heart disease, and stroke [1-3]. Lifestyle-related risk factors have been shown to be the major reasons of high incidence of CVD in the U.S. Considering the wide-spread reality of high CVD incidence and high mortality, special interest and research priority have been given to the early prevention of CVD and the pharmacological intervention providing cardiac protection to the population with high CVD risks.

The connection between sex hormones and CVD has long been documented. In particular, the presence of endogenous estrogens was identified as beneficial factor preventing the occurrence of CVD. Epidemiological evidence has shown that the pre-menopausal women have much lower CVD incidence than comparably aged men and post-menopausal women [4, 5]. Hormone replacement therapy or HRT [6] has been suggested for postmenopausal women and has been shown, in some cases, to lower the incidence of CVD among other benefits [7-10]. However, HRT has also been received controversial opinions over its adverse effects such as increasing risk in breast cancer and even slightly

high CVD incidence in women who received HRT [11-13]. In recent years, alternative approaches adopting plant phytoestrogens such as genistein (GEN) have been advocated for their potential preventive role against CVD [14-16]. The use of these phytoestrogens, which have weak estrogenic effects as well as anti-cancer properties, has been suggested to be equally effective but also safer than the conventional estrogen-based HRT. Such speculation was further augmented by epidemiological evidence that the intake of phytoestrogens, in particular the soy-derived isoflavones including GEN, is inversely associated with the risk of CVD [17-21]. Previous studies have explored extensively the effects of soy protein and soy–derived phytoestrogens to lower blood lipids and LDL cholesterol level, or effects to promote favorable lipoprotein profiles [22-24]. However, a later meta-analysis study provided contradictory conclusion that soy phytoestrogens have no effect on the blood cholesterol and lipid profiles [25, 26]. In 2006, the American Heart Association (AHA) concluded that the hypolipidemic effects of soy-derived phytoestrogens were non-significant [27]. Thus, other mechanisms still remain to be explored in order to fully understand why soy-derived phytoestrogens afford protection against CVD. Among the several phytoestrogens present in soy products, the primary form of isoflavones, GEN in particular is considered to be the most efficacious. Recently, several studies have suggested that isoflavones, and in particular GEN, provide direct protection to cardiac myocytes against damaging pathophysiological stresses [28-33]. However the underlying mechanism by which the GEN confers cardioprotection remains elusive.

1.2. Cardiomyocytes and cardioprotection

Cardiac muscle is a type of striated muscle mostly found in the ventricle walls of the heart, and it composes 90% of the total mass of a heart. The myocardiocyteal muscle cells or cardiomyocytes (CM) are mononuclear myocytes that comprised the basic contractile functional unit to generate coordinated force to pump blood to the circulation. Individual myocytes are joined by intercellular junctions permitting the continuation of mechanical force and electrical conduction. As the functional unit of the most 'heavy duty' organ, CM almost completely rely on aerobic metabolism majorly fueled by fatty acid beta-oxidation with mitochondria makes up to 25% of cell volume. Therefore, ischemic condition typically as a result of insufficient coronary blood supply can pose a great threat to CM survival and the sustenance of cardiac function.

Ischemic injury can cause CM damages which lead to oncosis manifested by the myocardial infarction. Prolonged infarction causes cell death by necrosis. During sever acute energy depletion, the ionic channels in the cell membrane fail to function hence causes osmotic imbalance, cellular swelling, sarcolemmal disruption, and release of cytoplasmic and internucleosomal contents. Necrotic cell death leads to myocardium inflammation which happens within 12-24 hours after injury. Cardiac fibrosis and cardiac geometry remodeling may follow as the result of collagen deposition leading to reduce the wall thickness and stress [34, 35]. While less damaged CM may die prematurely by apoptosis [36, 37]. The apoptosis is a programmed cell death procedure in which numerous cellular signaling pathways are activated. The end effectors of those pro-apoptotic pathways usually are caspases that cleaves several programmed death

substrates and activates endonuclease, leading to the characteristic fragmentation of DNA [35]. CM apoptosis contributes to the pathogenesis of post-ischemic left ventricular dysfunction, arrhythmias and congestive heart failure.

To our knowledge, there is no drug or therapy to completely reverse the damages caused by myocardial ischemic infarctions. Moreover, if blood supply is resumed to ischemic areas, reperfusion injury due to sudden introduction of oxidative stress can cause even more tissue damages. Myocardial reperfusion can cause apoptosis as a result of the production of oxygen-free radicals. Patients survived from acute infarction will still develop chronic ischemic heart failure. Therefore, means of cardioprotection, in particular against ischemic-reperfusion injuries, are of critical importance to lower mortality and morbidity of CVDs with great medical and social implications. Procedures like ischemic preconditioning [38] by subjecting heart to sublethal cycles of short-term ischemia and reperfusion have been proved effective to ameliorate the severity of a subsequent lethal cardiac infarction by conferring myocytes with tolerance against ischemic-reperfusion [39, 40].

The molecular mechanism of cardioprotection rendered by IPC is probably multi-faceted. Most studies have put attention on connecting the ATP depletion under ischemic conditions and the adenosine receptor/phospholipases-PKC signaling pathway which mediate multiple downstream cellular events to provide cardioprotection [41-46]. Mitogen-activated protein kinases (MAPK), in particular the c-Jun and p38 were identified as ischemic responsive in the heart [47, 48]. Tyrosine kinases were documented

as a key downstream step of PKC signaling to afford cardioprotection by IPC against ischemic injuries in a rabbit model [49]. However, cardioprotection by activation of tyrosine kinases in parallel with PKC was observed in pig and rat heart [50-52]. Additionally, researchers have shown that heart adaptation via IPC approach involves the increasing activity of anti-apoptotic NF kappa B [53] and has been shown to be regulated by tyrosine phosphorylation [54]. Recent studies also demonstrated that the recruitment of anti-apopotic phosphatidylinositol 3-kinase-Akt pathway, namely the RISK pathway, contribute to the cardioprotection afford by IPC [55] and the newly described ischemic postconditioning procedure [56-58].

The role of cytoskeleton remodeling was also characterized to be related to preconditioning triggered cardioprotection. The heat shock protein 27 (HSP27) was found to be over-expressed as the downstream target of p38 MAPK under the cardiac ischemic stress [59]. HSP27 involved with actin proliferation and myofilaments stabilization has been shown to take role in maintaining the cytoskeleton integrity and contractile function in prolonged ischemia [60, 61]. On the other hand, pharmacological evidences have suggested that the mitochondrial ATP-dependent potassium (KATP) channels might be a tenable end effector of cardioprotective preconditioning [62, 63].

The unraveling of many cardioprotective pathways has raised the speculation that pharmacologic activation of those pathways may enable one to harness similar medical benefits. Compare to ischemic pre-/post-conditioning procedure, novel therapeutic strategies based on pharmacological intervention have emerged as more applicable

alternatives with tenable clinical values. Estrogen was one of the examples among others administrated to reduce the infarct size in various in vivo models against ischemia insult [64-70]. The mechanistic explanations of cardioprotection afforded by estrogen were largely focused on the modulation of mitochondrial KATP channel and PI3K/Akt pathway [65, 66, 71-75]. However, despite the ability to demonstrate enhanced cardioprotection using estrogen in animal models, the translation to the human clinical application in the form of HRT has been disappointing along with other unwanted side effects such as high risk of developing breast cancer. Therefore, alternative approaches using plant isoflavones have been promoted for their potential cardioprotective values.

1.3. Genistein, chemical and biological properties

1.3.1. Introduction

Phytoestrogens are a diverse group of naturally synthesized chemicals found in various plants and have estrogenic and/or anti-estrogenic properties due to their similar structure to human estrogens. Classification of phytoestrogens based on structure properties results in two groups of compounds, the flavonoids, and the non-flavonoids. The flavonoinds can further be subcategorized into isoflavones, coumestans and prenyl flavonoids [76]. Genistein (GEN) or 4', 5, 7-trihydroxyisoflavone is a isoflavone that occurs in high concentration in leguminous plants like soybean, lupins, kudzu [77]. Whole soybean and processed soy foods and beverages are the major sources of isoflavone consumption in daily diet. Genistein with another less potent form the daidzein, constitute up to 90% of the isoflavone content in soy. However, it is noteworthy that the contents of GEN and other phytoestrogen differ according to species variety, location and season of harvest,

and also how the soy food is processed [78].

Genistein is majorly present as 7-O-beta-D-glucoside form with limited bioactivity in soy and soy products. However, the genistein glucoside can be hydrolyzed by gut bacterial community prior to absorption. Some fraction of GEN will be modified by intestinal microflora or liver to other form of isoflavones. Similarly, other isoflavones like biochanin A can also be metabolized to generate GEN. Aglucone form of GEN enters blood stream and later either excreted in urine or bile. Consider the complexity of the adsorption and metabolism of isoflavones, blood stream concentration of isoflavones after the consumption of soy meal may vary from case to case. One report suggest that the intake of modest portion of soy food with 45 g soy protein containing 80 mg of isoflavones, results in a 20- to 40-fold increase in blood isoflavone concentrations up to 0.5 μM [79]. In other study, dietary intake of soy foods containing 60 mg/d of isoflavones for 12 weeks can increase the serum isoflavones to 141.6 ng/mL or approximately 0.5 μM in postmenopausal women [80].

Isoflavones all have similar molecular weight to that of estradiol and contain a phenolic ring indispensable for binding to estrogen receptors. Genistein in particular present the optimal pattern of hydroxylation that the distance between aromatic hydroxyl groups is almost identical to that on the estradiol. Therefore, genistein was documented as the most estrogenic component compare to other isoflavones such as daidzein, biochanin A, glycitein, etc. [76]. Additionally, genistein differ from other isoflavones as it also affords broad spectrum tyrosine kinase inhibitory effects.

1.3.2. Concentration-dependent effects

Genistein has been report to exert cellular impact on numerous tissue and cell types in a dose-dependent fashion. Interestingly, cellular responses to GEN are not linearly proportional to the concentration of treatment. In most cases, distinctive arrays of molecular effectors were found response to GEN treatment at different concentrations. Such biphasic phenomenon was largely attributable to the estrogenic effects of GEN at low concentration and tyrosine kinase inhibitory effects at high concentration. In adipocytes, GEN was found to inhibit adipogenesis at low concentration but stimulate adipogenesis at high concentration [81, 82]. The same group found that in osteoblasts cells, GEN stimulates osteogenesis at low concentration and inhibit osteogenesis at high concentration [82, 83]. Similar biphasic phenomenon was also found in prostate cancer model, in which low-dose GEN (0.5 μM) decreased cell proliferation, invasion which were inversely promoted by 50 μM of GEN [84]. In myocardium, low dose GEN treatment tends to protect the CM but high dose treatment induces more myocyte death during ischemic infarction [85] as being discussed later in this chapter.

1.3.3. Estrogenic effects

The biological effects of estrogens and other estrogenic agents are mostly mediated by two estrogen receptor (ER) isoforms, ERα and ERβ. Both isoforms are expressed in neonatal [86] and adult [87] mouse CM. Interestingly, there is no significant difference in the ER abundance and localization between male and female [87, 88]. It was also shown

that ERα isoform located more on caveolae [89] and T-tubular membrane [90], suggest that ERα mediates most non-genomic estrogenic signaling, while the ERβ is heavily localized in nucleus and cytosol mediating most gene transcriptional regulation [87].

The acute estrogenic effects are typically mediated via a non-genomic mechanism in which the membrane-bound ERs upon activation and G-protein signaling which can further activate the phosphatidylinositol 3-kinase (PI3K) pathway [28, 91, 92]. The ERα based acute activation of estrogenic signaling was reported to provide cardioprotection against I/R injuries [69, 93]. Similar to estrogen, GEN was also reported to modulate cAMP-PKA signaling in a non-genomic fashion [94, 95]. On the other hand, the estrogenic genomic-wide changes are triggered by the recruitment of ERs in particular the ERβ isoforms located in nuclear membrane. Upon the binding of estrogenic compounds, ERs dimerize and translocate to the nucleus where they bind to estrogen response elements (ERE) on DNA and work as transcription factors. As the result, multiple estrogenic responsive genes including some cardioprotective genes such as nitric oxide synthase, heat shock proteins, and antioxidant enzymes (AOEs) are up-regulated [96, 97].

Although the binding affinity of GEN to both 1, 2 are 10-100 lower than endogenous hormone counterpart 17β-estradiol (E2) [76], the physiological concentration of GEN in the circulation after a typical soy-based meal can reach up to the range around micromolar, approximately 1000 times higher than that of endogenous estrogen. Genistein at such concentration (1-10 μM) has been shown to be effective for binding to

both ER subtypes [98-100]. However, there have very limited knowledge on whether the GEN at a physiological relevant concentration can cause similar signaling pathway or up-regulate cardioprotective genes in CM.

1.3.4. Tyrosine kinase inhibitory effects

Despite the fact that tyrosine phosphorylation is less widespread compared to serine and threonine phosphorylation, tyrosine kinase signaling nevertheless regulates many key cellular functions. There are two major groups of tyrosine kinases: receptor tyrosine kinases, such as EGFR, insulin receptor, and cytosolic tyrosine kinases, such as Src, Erk, Jak. The former group plays a major role on external signaling transmitting across the cell membrane, while the later group controls a wide array of signal transduction cascades and transcriptional regulations. Genistein have shown the capability to compete for the ATP-binding site of tyrsosine kinases [101]. Methodologically, GEN is being used as a non-specific protein tyrosine kinase (TK) inhibitor in many molecular biology experiments, typically at concentrations >10 μM. This concentration can be reached in the circulation if GEN is administrated as a form of pharmacological intervention. This property is unique to GEN, as compared to estrogen or other form of plant isoflavones. The TK inhibitory effects were also used to explain the anti-estrogenic or ER-independent effects of GEN in other cell types such as endothelium cells [102, 103], adipocytes [81, 104], breast cancer [105], and prostate cancer cell lines [84].

In cardiomyocytes, GEN was also widely used as a signaling diagnostic tool to understand the involvement of tyrosine kinase signaling in cardioprotection [106-109].

Baines et al [49] and Fryer et al [52] have shown that inhibition of tyrosine kinase by GEN can effectively block the downstream signaling of PKC pathway triggered by IPC. Similarly, genistein were also used to block cardioprotective PKC pathway stimulated by brief alcohol exposure [110], St Thomas' solution [111] and octreotide treatment [112]. Additionally, tyrosine kinase signaling was also shown as the downstream cascade step of cardioprotective PI3K/Akt pathway, thus can be blocked by GEN incubation [113, 114]. Shikrut et al found 50 μM of GEN used as tyrosine kinase inhibitor can effective block pro-apoptosis Fas-signaling triggered by hypoxia in murine ventricular myocytes [115]. Other signaling pathways involved with tyrosine phosphorylation, including the JAK/STAT pathway [116-118], were also shown to play a critical role in cardioprotection during ischemic pre- and/or post-conditioning. But no study has shown how GEN affects those complex pathway networks in a systematic approach.

On the other hand, tyrosine kinase was also a key regulator of multiple ion channels essential for myocyte contractile function. Genistein as a PTK inhibitor attenuates the L-type Ca^{2+} current in rat [119] and guinea pig [120-122] ventricular myocytes. Genistein at 50-80 μM also prevents activation of the swelling-activated Cl^- current in canine myocytes in a PTK dependent mechanism [123]. However, other studies have documented the activation the cardiac cAMP-dependent Cl channel by GEN in guinea pig CM [120, 124, 125]. The opening of triphosphate-sensitive potassium (K(ATP)) channels are key contributor of ischemic or pharmacological preconditioning. There were also few studies investigating PTKs inhibition of K(ATP) channels using GEN. GEN can elicit K(ATP) current by inhibit the PTKs [126] while its inactive analog daidzein did

not have the same effect [127]. Gao et al also found that the voltage-dependent potassium channels in rat CM was regulated PTK-dependent fashion and can be inhibited by GEN [128].

1.3.5. Antioxidant property

Genistein, like many other plant-derived isoflavones, is considered as an antioxidant [129, 130]. Early researches have been concentrated on the direct antioxidative effects of isoflavones in particular to reduce the susceptibility of low-density lipoprotein to oxidation [131, 132]. However, such mechanism was later determined as ineffective *in vivo* possibly due to the fact that physiological concentration of isoflavones in circulation is too low to afford any significant antioxidative benefits. Recently, attention has been drawn to the ability of isoflavones to stimulate antioxidant enzymes (AOEs) in cardiovascular system [133-136]. Genistein in particular has been shown to up-regulate endothelial NO synthase (eNOS) in rat vascular endothelial cells[137], glutathione peroxidase in human prostate cancer cells [138], MnSOD in human mammary gland tumor cell [139].

The impacts of isoflavones on AOE system are mostly considered as a part of their estrogenic effects. Studies have shown ERs upon binding of endogenous estrogen induces numerous antioxidant genes with antioxidant response element including MnSOD, HO-2, thioredoxins, and phase II detoxification enzymes like GST and NQO1 as summarized by Siow et al [133]. Such activation of antioxidative genes is likely through the activation of PI3K-Nrf2 pathway [133, 140] and ERK1/2-NFκB pathway [139]. In another study,

E2 and phytoestrogens including equol, GEN and daidzein was found to acutely up-regulate eNOS phosphorylation via ERK1/2-PI3K pathway probably independent of ER binding [141]. However, genistein triggered phosphorylation of eNOS possibly via PKA signaling independent of ERK and PI3K activation was also observed [142]. Recently, direct binding of ER dimmers to the ARE promoter region was discovered by Bianco et al suggesting the estrogenic agents can up-regulate antioxidant genes independent of Nrf2 activation [143]. In summary, it is very likely that isoflavones can affect the cellular redox balance in multiple mechanisms working in convergence.

It is notable that most of abovementioned investigations on the antioxidative action of isoflavones were mainly carried out in endothelial cells. Whether similar antioxidant benefits of GEN can apply to CMs are still in obscure.

1.3.6. Is genistein cardioprotective?

Numerous animal studies and epidemiological evidence advocate that soy phytoestrogens are beneficial for cardiovascular system. However, too much topics still left for debate on what mechanism is behind this cardioprotection effects in both physiology and molecular biology level. Genistein in particular have drawn attentions for its direct protection on CM against ischemic injuries [28, 29, 31, 33]. However, given the fact that most cellular effects of GEN is dose-dependent, it is logical to ask whether the cardioprotection of GEN was rendered only at physiological relevant concentration or can GEN at pharmacological concentration can be used in preconditioning or postconditioning intervention procedures effectively and safely. It is especially important to point out that

GEN at higher concentration has been reported to induce more CM death and block the cardioprotective effects of ICP or other pharmacological preconditioning [49, 106, 108-114, 144-147]. To date only few studies have systematically document the dose-dependent molecular targets of GEN in CM. Giving that a huge pool of molecules that can be affect by GEN have already been identified and possibly with more still remain unknown, oversimplified explanation from one single angle may compromise our comprehensive appreciation on the whole picture of the molecular basis of GEN's action in CM.

1.4. Proteomics and phosphoproteomics

Proteome was defined as the protein complement of an organism's genome in a particular physiological state. The assessment of whole proteome, i.e. the proteomics attempts to capture snapshots of a comprehensive protein profile of an ever-changing system, i.e., any organism, tissue or cell in response to either internal or external signals/perturbations. The idea of –omics type global analysis was first initiated by microarray based mRNA expression survey (transcriptomics). This popular technique has being used in biomedical researches, among many other fields, to study disease development and drug effects. However, evidence shown the mRNA level does not correlate with the actual protein level. There are significant differences in dynamic range and rates of production and degradation between proteins and mRNAs molecules. Regulation on translation and post-translational events such as alternative splicing and protein modifications make the proteome more complex and dynamic than the transcriptome. Different from transcriptomics, the proteomics directly measure the presence and absence of proteins

that execute the actual molecular functions which ultimately determine the state of a biological system.

With the completion of the genome of the target organism, identification of literally any protein in a mixture without the use of antibody can be achieved by interpreting the peptide fragmentation patterns in a highly accurate mass spectrometry. Large-scale separation techniques like two-dimensional electrophoresis (2DE) or multidimensional liquid chromatography (MDLC) allow us to fractionate the complex proteome into much simpler protein/peptide subsets that can be easily characterized by mass spectrometry, often quantitatively. Nowadays, a typical differential proteomics study can quantify the relative abundance of many hundreds or even thousands of proteins across multiple samples.

In recent years, proteomic approaches are increasingly employed to unravel the complexity of isoflavones effects on numerous cell and tissue types. Fuch et al demonstrated the usefulness of proteomics to investigate the effect of GEN on the human endothelial cells stressed by pro-atherogenic stimuli such as oxidized LDL [148] and homocysteine [149, 150]. In another case, Zhang et al carried out a dose-dependent and time–dependent proteomics study to discover key tumorigenesis pathways modulated by of GEN on human leukemia cells [151]. Sotoca et al use both transcriptomic and proteomic approach to shown that GEN exerted similar estrogenic effects as estradiol on breast cancer cells [152]. Similarly, Wang et al [153] and Rowell et al [154] performed two 2DE based proteomics studies to identify key proteins involved in cell proliferation

related pathways to understand mechanisms of action in breast cancer chemoprevention afforded by GEN. In all these studies, proteomics not only allow researchers to observe multiple up- or down-stream gene products regulating or being regulated by the treatment that are known, it also unveils the molecular events that were not expected with and that would otherwise not be discovered via conventional methods.

In recent years, efforts have been made to use new mass spectrometry techniques to characterize the post-translational modifications (PTMs) of proteins, which can never be assessed by transcriptomics. PTMs are vital for the proteins to function correctly and are related to regulations in many cellular processes. Among 200 different forms of PTMs, protein phosphorylation at serine, threonine and tyrosine residues, are the most widely studied. GEN can trigger non-genomic signaling cascade via the binding of membrane ERs and also serves as a non-specific tyrosine kinase modulator in particular at high concentration. Both scenarios involve changes in phosphorylation status of multiple proteins that can be surveyed by proteomics techniques specialized to characterize protein phosphorylation: the phosphoproteomics. Like other branches of proteomics, phosphoproteomics employ high throughput separation techniques and mass spectrometry to provide protein phosphorylation profile in a global scale. For detailed review of phosphoproteomics techniques and applications on cardiovascular research, please refer to chapter 2.

1.5. Models

Both HL-1 cultured CM and primary CM isolated from male Sprague Dawley (SD) rat

will be used as in vitro models in the studies. Compare to in vivo animal model, the use of HL-1 cells or primary cells enable us to focus on the molecular effects of GEN only on CM in a relatively clear-cut experimental setting where the concentration of the treatment can be easily controlled and maintained during the whole experiment. Such in vitro model also avoids the protein contamination from other cell type and abundant proteins from muscle connective tissue which tend to dominate over low abundant but important proteins such as transcription factors during proteomics analysis. Moreover, as the proteomics is inherently sensitive to biological background noise, the low sample variability provide by homogenous cell population enhances the proteomic discovery power by avoiding identifying false positive expression level changes.

HL-1 cell line represents a well-characterized contractile CM culture that can divide continuously as well as maintain a differentiated cardiac phenotype. These cells are derived from mouse atrial CM and have highly organized sarcomere structures and intracellular ANF granules similar to adult mouse atrial CM [155]. Moreover, HL-1 cells maintain the gene expression profile similar to that of adult mouse CM and, thus, they represent a reliable *in vitro* model to study the cellular metabolism of CM [156]. However, inexplicable results on HL-1 cells may expected due to the cellular resources shift towards other functions such as proliferation which may interfere with the interpretation of GEN's effects on CM. Such drawback can be compensated by employing the unproliferable primary cell which is considered to be more biological relevant to *in vivo* model. Nonetheless, due to the lack of tissue architecture and cell-cell interaction with other cell types, it is inevitable that both primary and HL-1 cell models abolish some

cellular functions based on tissue context.

In this dissertation, I will use HL-1 cells as preliminary model to test if the GEN have different impact on cardiac proteome at different concentrations, as HL-1 are generally less sensitive to external stimulus and can withstand GEN treatment with large concentration span, so we can isolate differential expressed gene products at two concentrations with distinctive effects. I will further use primary CM isolated from adult female SD rats as the main model to study GEN's effects on cardiac proteome and phosphoproteome via ER-dependent and ER-independent mechanism. Moreover, one shall keep in mind that simplified cell culture or primary cell milieu cannot reflect the complexity and dynamics of physiological environment of CM, in particular it cannot be used to assess the bioavailability and dynamics of the presence of GEN, nor it can assess hundreds of other possible synergistic/antagonistic effects provided by other *in vivo* stimulus sources such as endogenous hormones. Those discoveries of molecular targets altered by GEN treatment from this cell-based proteomics study certainly need to be further validated by *in vivo* experiments.

2. Objectives and Contributions of This Dissertation

The overall goal of this dissertation is to use high throughput proteomics and phosphoproteomics approaches to investigate the global effects of GEN treatments on CM that are mainly due to estrogenic and non-estrogenic effects. Based on previous knowledge that it is mainly the concentration that determines whether GEN works as an estrogenic or PTK inhibitory agent, we first hypothesize that GEN may cause CM

proteome alteration in a concentration-dependent fashion. Through the analysis of proteins activation or inhibition, key pathways and cellular functions related to concentration-dependent cardioprotection of GEN can be identified.

Further, we hypothesize that the under the pharmacological relevant concentration, the GEN influence CM most likely by both estrogenic and non-estrogenic effects concurrently. In order to measure proteome changes by GEN's via each mechanism separately, estrogen receptor antagonist can be used to block proteome changes due to estrogenic effects of GEN. Further, I hypothesize that non-estrogenic effect mainly due to the PTK inhibitory property of GEN can be reflected by the alteration in CM singling pathways especially in the form of widespread changes in the pattern of protein phosphorylation. The use of phosphoproteomics methodology can enable me to identify key signaling pathway changes that undergirding the molecular action of GEN on CM.

Overall, with both high throughput proteomics and phosphoproteomics data, we can better assess the extent of molecular targets altered by GEN treatments and by which pathway they were targeted. By setting reductionist-style thinking aside and to look at the system as an integration of many complex compartments interacting with each other, we gain a much-expanded view of GEN's effects on CM, which is required for the rationale of new hypotheses in next steps. However, one should always bear in mind that at least by itself, proteomics or any type of -omics study was not meant to be a tool to prove any detailed hypothesis in mechanistic fashion, rather, it provide us a comprehensive way to observe the global changes in a biological system so we can better rationale our tentative

hypothesis about the mechanism, which indeed still need to be addressed by a reductionist approach. Hence, all the proteomics discoveries included in this dissertation need to be viewed as no more than faithful observations. Further, all biological interpretations discreetly drawn from our proteomics and phosphoproteomics data will only be considered as cornerstone to establish future hypotheses which still need to be validated by other methods.

To our knowledge, no published research exists investigating the influences of soy-derived phytoestrogens on the proteome or phosphoproteome of CM. We expect that these data-driven proteomics studies will provide vital clues and a rational foundation for future hypothesis-driven mechanistic studies of cardiac protection afford by GEN and other isoflavones in general. Eventually we hope the molecular detail discovered in those proteomic studies can further help physiologists to provide recommendations if soy or a soy-based diet can be considered as beneficial for cardiovascular system and whether soy-derived phytoestrogens like GEN can be used as medical intervention to prevent or treat CVD.

REFERENCES

[1] X.J. Kochanek KD, Murphy SL, Minino AM, Kung HC. , Preliminary Data for 2009 *National Vital Statistics Reports* **59** (4) (2011), pp.

[2] Morbidity and Mortality: 2009 Chart Book on Cardiovascular and Lung Diseases, National Institute of Health, National Heart, Lung, and Blood Institute, (2009), pp.

[3] D. Lloyd-Jones, R. Adams, M. Carnethon, G. De Simone, T.B. Ferguson, K. Flegal *et al.*, Heart disease and stroke statistics--2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, *Circulation* **119** (3) (2009), pp. 480-486.

[4] W.B. KannelandP.W.F. Wilson, Risk-Factors That Attenuate the Female Coronary-Disease Advantage, *Archives of Internal Medicine* **155** (1) (1995), pp. 57-61.

[5] M.J. Stampfer, G.A. ColditzandW.C. Willett, Menopause and heart disease. A review, *Ann N Y Acad Sci* **592** (1990), pp. 193-203; discussion 257-162.

[6] M.M. Corcoran, O. Rasool, Y. Liu, A. Iyengar, D. Grander, R.E. Ibbotson *et al.*, Detailed molecular delineation of 13q14.3 loss in B-cell chronic lymphocytic leukemia, *Blood* **91** (4) (1998), pp. 1382-1390.

[7] F. Grodstein, M.J. Stampfer, J.E. Manson, G.A. Colditz, W.C. Willett, B. Rosner *et al.*, Postmenopausal estrogen and progestin use and the risk of cardiovascular disease, *New England Journal of Medicine* **335** (7) (1996), pp. 453-461.

[8] G.M. Rosano, P.M. Sarrel, P.A. Poole-WilsonandP. Collins, Beneficial effect of oestrogen on exercise-induced myocardial ischaemia in women with coronary artery disease, *Lancet* **342** (8864) (1993), pp. 133-136.

[9] M.J. Stampfer, G.A. Colditz, W.C. Willett, J.E. Manson, B. Rosner, F.E. Speizer *et al.*, Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from

the nurses' health study, *N Engl J Med* **325** (11) (1991), pp. 756-762.

[10] M.J. StampferandG.A. Colditz, Estrogen Replacement Therapy and Coronary Heart-Disease - a Quantitative Assessment of the Epidemiologic Evidence, *Preventive Medicine* **20** (1) (1991), pp. 47-63.

[11] J.E. Rossouw, G.L. Anderson, R.L. Prentice, A.Z. LaCroix, C. Kooperberg, M.L. Stefanick *et al.*, Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial, *JAMA* **288** (3) (2002), pp. 321-333.

[12] G.L. Anderson, M. Limacher, A.R. Assaf, T. Bassford, S.A. Beresford, H. Black *et al.*, Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial, *JAMA* **291** (14) (2004), pp. 1701-1712.

[13] J.E. Manson, J. Hsia, K.C. Johnson, J.E. Rossouw, A.R. Assaf, N.L. Lasser *et al.*, Estrogen plus progestin and the risk of coronary heart disease, *N Engl J Med* **349** (6) (2003), pp. 523-534.

[14] D. Altavilla, A. Crisafulli, H. Marini, M. Esposito, R. D'Anna, F. Corrado *et al.*, Cardiovascular effects of the phytoestrogen genistein, *Current Medicinal Chemistry Cardiovascular and Hematological Agents* **2** (2) (2004), pp. 179-186.

[15] M. Atteritano, H. Marini, L. Minutoli, F. Polito, A. Bitto, D. Altavilla *et al.*, Effects of the phytoestrogen genistein on some predictors of cardiovascular risk in osteopenic, postmenopausal women: a two-year randomized, double-blind, placebo-controlled study, *J Clin Endocrinol Metab* **92** (8) (2007), pp. 3068-3075.

[16] A. Crisafulli, D. Altavilla, H. Marini, A. Bitto, D. Cucinotta, N. Frisina *et al.*, Effects

of the phytoestrogen genistein on cardiovascular risk factors in postmenopausal women, *Menopause* **12** (2) (2005), pp. 186-192.

[17] A. Menotti, A. Keys, H. Blackburn, D. Kromhout, M. Karvonen, A. Nissinen *et al.*, Comparison of multivariate predictive power of major risk factors for coronary heart diseases in different countries: results from eight nations of the Seven Countries Study, 25-year follow-up, *J Cardiovasc Risk* **3** (1) (1996), pp. 69-75.

[18] X. Zhang, X.O. Shu, Y.-T. Gao, G. Yang, Q. Li, H. Li *et al.*, Soy food consumption is associated with lower risk of coronary heart disease in Chinese women, *The Journal of Nutrition* **133** (9) (2003), pp. 2874-2878.

[19] K.D. SetchellandA. Cassidy, Dietary isoflavones: biological effects and relevance to human health, *J Nutr* **129** (3) (1999), pp. 758S-767S.

[20] M.G. Hertog, D. Kromhout, C. Aravanis, H. Blackburn, R. Buzina, F. Fidanza *et al.*, Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study, *Arch Intern Med* **155** (4) (1995), pp. 381-386.

[21] A. Keys, A. Menotti, C. Aravanis, H. Blackburn, B.S. Djordevic, R. Buzina *et al.*, The seven countries study: 2,289 deaths in 15 years, *Prev Med* **13** (2) (1984), pp. 141-154.

[22] D. Park, T. HuangandW.H. Frishman, Phytoestrogens as cardioprotective agents, *Cardiology in Review* **13** (1) (2005), pp. 13-17.

[23] J.W. Anderson, B.M. JohnstoneandM.E. Cook-Newell, Meta-analysis of the effects of soy protein intake on serum lipids, *N Engl J Med* **333** (5) (1995), pp. 276-282.

[24] M.S. Anthony, T.B. ClarksonandJ.K. Williams, Effects of soy isoflavones on atherosclerosis: potential mechanisms, *Am J Clin Nutr* **68** (6 Suppl) (1998), pp. 1390S-1393S.

[25] R.M. WeggemansandE.A. Trautwein, Relation between soy-associated isoflavones and LDL and HDL cholesterol concentrations in humans: a meta-analysis, *European Journal of Clinical Nutrition* **57** (8) (2003), pp. 940-946.

[26] D.R. Rios, E.T. Rodrigues, A.P. Cardoso, M.B. Montes, S.A. FranceschiniandM.R. Toloi, Lack of effects of isoflavones on the lipid profile of Brazilian postmenopausal women, *Nutrition* **24** (11-12) (2008), pp. 1153-1158.

[27] F.M. Sacks, A. Lichtenstein, L. Van Horn, W. Harris, P. Kris-EthertonandM. Winston, Soy protein, isoflavones, and cardiovascular health: an American Heart Association Science Advisory for professionals from the Nutrition Committee, *Circulation* **113** (7) (2006), pp. 1034-1044.

[28] R. Tissier, X. Waintraub, N. Couvreur, M. Gervais, P. Bruneval, C. Mandet *et al.*, Pharmacological postconditioning with the phytoestrogen genistein, *J Mol Cell Cardiol* **42** (1) (2007), pp. 79-87.

[29] P. Zhai, T.E. Eurell, R.P. Cotthaus, E.H. Jeffery, J.M. BahrandD.R. Gross, Effects of dietary phytoestrogen on global myocardial ischemia-reperfusion injury in isolated female rat hearts, *Am J Physiol Heart Circ Physiol* **281** (3) (2001), pp. H1223-1232.

[30] N. Couvreur, R. Tissier, S. Pons, M. Chenoune, X. Waintraub, A. Berdeaux *et al.*, The ceiling effect of pharmacological postconditioning with the phytoestrogen genistein is reversed by the GSK3beta inhibitor SB 216763 [3-(2,4-dichlorophenyl)-4(1-methyl-1H-indol-3-yl)-1H-pyrrole-2,5-dione] through mitochondrial ATP-dependent potassium channel opening, *J Pharmacol Exp Ther* **329** (3) (2009), pp. 1134-1141.

[31] L. Al-Nakkash, B. Markus, K. Bowden, L.M. Batia, W.C. ProzialeckandT.L. Broderick, Effects of acute and 2-day genistein treatment on cardiac function and

ischemic tolerance in ovariectomized rats, *Gend Med* **6** (3) (2009), pp. 488-497.

[32] E. Souzeau, S. Belanger, S. PicardandC.F. Deschepper, Dietary isoflavones during pregnancy and lactation provide cardioprotection to offspring rats in adulthood, *Am J Physiol Heart Circ Physiol* **289** (2) (2005), pp. H715-721.

[33] B. Deodato, D. Altavilla, G. Squadrito, G.M. Campo, M. Arlotta, L. Minutoli *et al.*, Cardioprotection by the phytoestrogen genistein in experimental myocardial ischaemia-reperfusion injury, *Br J Pharmacol* **128** (8) (1999), pp. 1683-1690.

[34] A. Maes, W. Flameng, J. Nuyts, M. Borgers, B. Shivalkar, J. Ausma *et al.*, Histological Alterations in Chronically Hypoperfused Myocardium - Correlation with Pet Findings, *Circulation* **90** (2) (1994), pp. 735-745.

[35] D. MillerandS. Herrmann. Myocardial Infarction. In: Runge M, Patterson C, editors. Principles of Molecular Cardiology: Humana Press; 2005. p. 219-237.

[36] B. Maisch, How cardiac cells die--necrosis, oncosis and apoptosis, *Herz* **24** (3) (1999), pp. 181-188.

[37] G. MajnoandI. Joris, Apoptosis, oncosis, and necrosis. An overview of cell death, *Am J Pathol* **146** (1) (1995), pp. 3-15.

[38] C.E. Murry, R.B. JenningsandK.A. Reimer, Preconditioning with ischemia: a delay of lethal cell injury in ischemic myocardium, *Circulation* **74** (5) (1986), pp. 1124-1136.

[39] C.S. LawsonandJ.M. Downey, Preconditioning: state of the art myocardial protection, *Cardiovasc Res* **27** (4) (1993), pp. 542-550.

[40] G.S. Liu, J. Thornton, D.M. Vanwinkle, A.W.H. Stanley, R.A. OlssonandJ.M. Downey, Protection against Infarction Afforded by Preconditioning Is Mediated by A1 Adenosine Receptors in Rabbit Heart, *Circulation* **84** (1) (1991), pp. 350-356.

[41] G.S. Liu, S.C. Richards, R.A. Olsson, K.H. Mullane, R.S. WalshandJ.M. Downey, Evidence That the Adenosine a(3) Receptor May Mediate the Protection Afforded by Preconditioning in the Isolated Rabbit Heart, *Cardiovascular Research* **28** (7) (1994), pp. 1057-1061.

[42] J.D. Thornton, G.S. Liu, R.A. OlssonandJ.M. Downey, Intravenous pretreatment with A1-selective adenosine analogues protects the heart against infarction, *Circulation* **85** (2) (1992), pp. 659-665.

[43] J.A. Auchampach, A. Rizvi, Y. Qiu, X.L. Tang, C. Maldonado, S. Teschner *et al.*, Selective activation of A3 adenosine receptors with N6-(3-iodobenzyl)adenosine-5'-N-methyluronamide protects against myocardial stunning and infarction without hemodynamic changes in conscious rabbits, *Circ Res* **80** (6) (1997), pp. 800-809.

[44] K. Przyklenk, M.A. Sussman, B.Z. SimkhovichandR.A. Kloner, Does ischemic preconditioning trigger translocation of protein kinase C in the canine model?, *Circulation* **92** (6) (1995), pp. 1546-1557.

[45] M.B. Mitchell, X.Z. Meng, L.H. Ao, J.M. Brown, A.H. HarkenandA. Banerjee, Preconditioning of Isolated Rat-Heart Is Mediated by Protein-Kinase-C, *Circulation Research* **76** (1) (1995), pp. 73-81.

[46] P.H. Sugden, S.J. Fuller, A. MichaelandA. Clerk, Activation of mitogen-activated protein kinase subfamilies by oxidative stress in the perfused rat heart, *Biochemical Society Transactions* **25** (4) (1997), pp. S565-S565.

[47] P.H. SugdenandA. Clerk, "Stress-responsive" mitogen-activated protein kinases (c-Jun N-terminal kinases and p38 mitogen-activated protein kinases) in the myocardium, *Circulation Research* **83** (4) (1998), pp. 345-352.

[48] C.P. Baines, L. Wang, M.V. CohenandJ.M. Downey, Protein tyrosine kinase is downstream of protein kinase C for ischemic preconditioning's anti-infarct effect in the rabbit heart, *J Mol Cell Cardiol* **30** (2) (1998), pp. 383-392.

[49] C. Vahlhaus, R. Schulz, H. Post, J. RoseandU. Backenkohler, Prevention of ischemic preconditioning only by combined inhibition of protein kinase C and protein tyrosine kinase, *Circulation* **96** (8) (1997), pp. 3207-3207.

[50] C. Vahlhaus, R. Schulz, H. Post, R. OnallahandG. Heusch, No prevention of ischemic preconditioning by the protein kinase C inhibitor staurosporine in swine, *Circulation Research* **79** (3) (1996), pp. 407-414.

[51] R.M. Fryer, J.E. Schultz, A.K. HsuandG.J. Gross, Importance of PKC and tyrosine kinase in single or multiple cycles of preconditioning in rat hearts, *American Journal of Physiology-Heart and Circulatory Physiology* **276** (4) (1999), pp. H1229-H1235.

[52] N. Maulik, H. SasakiandN. Galang, Differential regulation of apoptosis by ischemia-reperfusion and ischemic adaptation, *Ann N Y Acad Sci* **874** (1999), pp. 401-411.

[53] J. Zhang, P. Ping, T.M. Vondriska, X.L. Tang, G.W. Wang, E.M. Cardwell *et al.*, Cardioprotection involves activation of NF-kappa B via PKC-dependent tyrosine and serine phosphorylation of I kappa B-alpha, *Am J Physiol Heart Circ Physiol* **285** (4) (2003), pp. H1753-1758.

[54] D.J. Hausenloy, A. Tsang, M.M. MocanuandD.M. Yellon, Ischemic preconditioning protects by activating prosurvival kinases at reperfusion, *American Journal of Physiology-Heart and Circulatory Physiology* **288** (2) (2005), pp. H971-H976.

[55] M. Zhu, J. Feng, E. Lucchinetti, G. Fischer, L. Xu, T. Pedrazzini *et al.*, Ischemic postconditioning protects remodeled myocardium via the PI3K-PKB/Akt reperfusion

injury salvage kinase pathway, *Cardiovasc Res* **72** (1) (2006), pp. 152-162.

[56] A. Tsang, D.J. Hausenloy, M.M. MocanuandD.M. Yellon, Postconditioning: a form of "modified reperfusion" protects the myocardium by activating the phosphatidylinositol 3-kinase-Akt pathway, *Circ Res* **95** (3) (2004), pp. 230-232.

[57] D.J. Hausenloy, A. TsangandD.M. Yellon, The reperfusion injury salvage kinase pathway: a common target for both ischemic preconditioning and postconditioning, *Trends Cardiovasc Med* **15** (2) (2005), pp. 69-75.

[58] J.L. Martin, R. Mestril, R. Hilal-Dandan, L.L. BruntonandW.H. Dillmann, Small heat shock proteins and protection against ischemic injury in cardiac myocytes, *Circulation* **96** (12) (1997), pp. 4343-4348.

[59] H. Wei, W. CampbellandR.S. Vander Heide, Heat shock-induced cardioprotection activates cytoskeletal-based cell survival pathways, *Am J Physiol Heart Circ Physiol* **291** (2) (2006), pp. H638-647.

[60] X.Y. Lu, L. Chen, X.L. CaiandH.T. Yang, Overexpression of heat shock protein 27 protects against ischaemia/reperfusion-induced cardiac dysfunction via stabilization of troponin I and T, *Cardiovasc Res* **79** (3) (2008), pp. 500-508.

[61] T. Sato, N. Sasaki, J. Seharaseyon, B. O'RourkeandE. Marban, Selective pharmacological agents implicate mitochondrial but not sarcolemmal K(ATP) channels in ischemic cardioprotection, *Circulation* **101** (20) (2000), pp. 2418-2423.

[62] Y. Liu, T. Sato, B. O'RourkeandE. Marban, Mitochondrial ATP-dependent potassium channels: novel effectors of cardioprotection?, *Circulation* **97** (24) (1998), pp. 2463-2469.

[63] M. van Eickels, R.D. Patten, M.J. Aronovitz, A. Alsheikh-Ali, K. Gostyla, F. Celestin *et al.*, 17-beta-estradiol increases cardiac remodeling and mortality in mice with

myocardial infarction, *J Am Coll Cardiol* **41** (11) (2003), pp. 2084-2092.

[64] C.H. Tsai, S.F. Su, T.F. ChouandT.M. Lee, Differential effects of sarcolemmal and mitochondrial K(ATP) channels activated by 17 beta-estradiol on reperfusion arrhythmias and infarct sizes in canine hearts, *J Pharmacol Exp Ther* **301** (1) (2002), pp. 234-240.

[65] T.M. Lee, S.F. Su, C.C. Tsai, Y.T. LeeandC.H. Tsai, Cardioprotective effects of 17 beta-estradiol produced by activation ofmitochondrial ATP-sensitive K(+)Channels in canine hearts, *J Mol Cell Cardiol* **32** (7) (2000), pp. 1147-1158.

[66] T.M. Lee, M.S. Lin, T.F. Chou, C.H. TsaiandN.C. Chang, Adjunctive 17beta-estradiol administration reduces infarct size by altered expression of canine myocardial connexin43 protein, *Cardiovasc Res* **63** (1) (2004), pp. 109-117.

[67] S.L. Hale, Y. BirnbaumandR.A. Kloner, Estradiol, Administered Acutely, Protects Ischemic Myocardium in Both Female and Male Rabbits, *J Cardiovasc Pharmacol Ther* **2** (1) (1997), pp. 47-52.

[68] E.A. Booth, N.R. ObeidandB.R. Lucchesi, Activation of estrogen receptor-alpha protects the in vivo rabbit heart from ischemia-reperfusion injury, *Am J Physiol Heart Circ Physiol* **289** (5) (2005), pp. H2039-2047.

[69] E.A. Booth, M. Marchesi, E.J. KilbourneandB.R. Lucchesi, 17 beta-estradiol as a receptor-mediated cardioprotective agent, *Journal of Pharmacology and Experimental Therapeutics* **307** (1) (2003), pp. 395-401.

[70] X.M. Yang, J.B. Proctor, L. Cui, T. Krieg, J.M. DowneyandM.V. Cohen, Multiple, brief coronary occlusions during early reperfusion protect rabbit hearts by targeting cell signaling pathways, *J Am Coll Cardiol* **44** (5) (2004), pp. 1103-1110.

[71] T.M. Lee, S.F. Su, T.F. ChouandC.H. Tsai, Pharmacologic preconditioning of

estrogen by activation of the myocardial adenosine triphosphate-sensitive potassium channel in patients undergoing coronary angioplasty, *J Am Coll Cardiol* **39** (5) (2002), pp. 871-877.

[72] T.M. Lee, T.F. ChouandC.H. Tsai, Differential role of K(ATP) channels activated by conjugated estrogens in the regulation of myocardial and coronary protective effects, *Circulation* **107** (1) (2003), pp. 49-54.

[73] R.D. PattenandR.H. Karas, Estrogen replacement and cardiomyocyte protection, *Trends Cardiovasc Med* **16** (3) (2006), pp. 69-75.

[74] R.D. Patten, I. Pourati, M.J. Aronovitz, J. Baur, F. Celestin, X. Chen *et al.*, 17beta-estradiol reduces cardiomyocyte apoptosis in vivo and in vitro via activation of phospho-inositide-3 kinase/Akt signaling, *Circ Res* **95** (7) (2004), pp. 692-699.

[75] F. Yildiz. Phytoestrogens in Functional Foods. Boca Raton, FL: CRC Press; 2006.

[76] P.B. Kaufman, J.A. Duke, H. Brielmann, J. BoikandJ.E. Hoyt, A comparative survey of leguminous plants as sources of the isoflavones, genistein and daidzein: implications for human nutrition and health, *J Altern Complement Med* **3** (1) (1997), pp. 7-12.

[77] B.D. Oomah. Pytoestrogens. Boca Raton, FL: CRC Press; 2002.

[78] M.S. Morton, G. Wilcox, M.L. WahlqvistandK. Griffiths, Determination of lignans and isoflavonoids in human female plasma following dietary supplementation, *J Endocrinol* **142** (2) (1994), pp. 251-259.

[79] M.D. Scheiber, J.H. Liu, M.T. Subbiah, R.W. RebarandK.D. Setchell, Dietary inclusion of whole soy foods results in significant reductions in clinical risk factors for osteoporosis and cardiovascular disease in normal postmenopausal women, *Menopause* **8** (5) (2001), pp. 384-392.

[80] Z.C. Dang, Dose-dependent effects of soy phyto-oestrogen genistein on adipocytes: mechanisms of action, *Obes Rev* **10** (3) (2009), pp. 342-349.

[81] Z.C. Dang, V. Audinot, S.E. Papapoulos, J.A. BoutinandC.W. Lowik, Peroxisome proliferator-activated receptor gamma (PPARgamma) as a molecular target for the soy phytoestrogen genistein, *J Biol Chem* **278** (2) (2003), pp. 962-967.

[82] Z.C. DangandC. Lowik, Dose-dependent effects of phytoestrogens on bone, *Trends Endocrinol Metab* **16** (5) (2005), pp. 207-213.

[83] L.H. El TounyandP.P. Banerjee, Identification of a biphasic role for genistein in the regulation of prostate cancer growth and metastasis, *Cancer Res* **69** (8) (2009), pp. 3695-3703.

[84] R. Liew, J.K. Williams, P. CollinsandK.T. MacLeod, Soy-derived isoflavones exert opposing actions on Guinea pig ventricular myocytes, *J Pharmacol Exp Ther* **304** (3) (2003), pp. 985-993.

[85] C. Grohe, S. Kahlert, K. Lobbert, M. Stimpel, R.H. Karas, H. Vetter *et al.*, Cardiac myocytes and fibroblasts contain functional estrogen receptors, *FEBS Lett* **416** (1) (1997), pp. 107-112.

[86] E. Lizotte, S.A. Grandy, A. Tremblay, B.G. AllenandC. Fiset, Expression, distribution and regulation of sex steroid hormone receptors in mouse heart, *Cell Physiol Biochem* **23** (1-3) (2009), pp. 75-86.

[87] A.M. DeschampsandE. Murphy, Activation of a novel estrogen receptor, GPER, is cardioprotective in male and female rats, *Am J Physiol Heart Circ Physiol* **297** (5) (2009), pp. H1806-1813.

[88] T.H. Chung, S.M. Wang, J.Y. Liang, S.H. YangandJ.C. Wu, The interaction of

estrogen receptor alpha and caveolin-3 regulates connexin43 phosphorylation in metabolic inhibition-treated rat cardiomyocytes, *Int J Biochem Cell Biol* **41** (11) (2009), pp. 2323-2333.

[89] A.B. Ropero, M. Eghbali, T.Y. Minosyan, G. Tang, L. ToroandE. Stefani, Heart estrogen receptor alpha: distinct membrane and nuclear distribution patterns and regulation by estrogen, *J Mol Cell Cardiol* **41** (3) (2006), pp. 496-510.

[90] E.J. Filardo, J.A. Quinn, K.I. BlandandA.R. Frackelton, Jr., Estrogen-induced activation of Erk-1 and Erk-2 requires the G protein-coupled receptor homolog, GPR30, and occurs via trans-activation of the epidermal growth factor receptor through release of HB-EGF, *Mol Endocrinol* **14** (10) (2000), pp. 1649-1660.

[91] C.M. Revankar, D.F. Cimino, L.A. Sklar, J.B. ArterburnandE.R. Prossnitz, A transmembrane intracellular estrogen receptor mediates rapid cell signaling, *Science* **307** (5715) (2005), pp. 1625-1630.

[92] J.L. Novotny, A.M. Simpson, N.J. Tomicek, T.S. LancasterandD.H. Korzick, Rapid estrogen receptor-alpha activation improves ischemic tolerance in aged female rats through a novel protein kinase C epsilon-dependent mechanism, *Endocrinology* **150** (2) (2009), pp. 889-896.

[93] W.W. Ng, W. Keung, Y.C. Xu, K.F. Ng, G.P. Leung, P.M. Vanhoutte *et al.*, Genistein potentiates protein kinase A activity in porcine coronary artery, *Mol Cell Biochem* **311** (1-2) (2008), pp. 37-44.

[94] H. TeohandR.Y. Man, Enhanced relaxation of porcine coronary arteries after acute exposure to a physiological level of 17beta-estradiol involves non-genomic mechanisms and the cyclic AMP cascade, *British Journal of Pharmacology* **129** (8) (2000), pp. 1739-

1747.

[95] I. Nikolic, D. Liu, J.A. Bell, J. Collins, C. SteenbergenandE. Murphy, Treatment with an estrogen receptor-beta-selective agonist is cardioprotective, *J Mol Cell Cardiol* **42** (4) (2007), pp. 769-780.

[96] S. Nuedling, R.H. Karas, M.E. Mendelsohn, J.A. Katzenellenbogen, B.S. Katzenellenbogen, R. Meyer *et al.*, Activation of estrogen receptor beta is a prerequisite for estrogen-dependent upregulation of nitric oxide synthases in neonatal rat cardiac myocytes, *FEBS Lett* **502** (3) (2001), pp. 103-108.

[97] M. Maggiolini, D. Bonofiglio, S. Marsico, M.L. Panno, B. Cenni, D. Picard *et al.*, Estrogen receptor alpha mediates the proliferative but not the cytotoxic dose-dependent effects of two major phytoestrogens on human breast cancer cells, *Mol Pharmacol* **60** (3) (2001), pp. 595-602.

[98] G.G. Kuiper, J.G. Lemmen, B. Carlsson, J.C. Corton, S.H. Safe, P.T. van der Saag *et al.*, Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta, *Endocrinology* **139** (10) (1998), pp. 4252-4263.

[99] M.W. Carter, W.W. Smart, Jr.andG. Matrone, Estimation of estrogenic activity of genistein obtained from soybean meal, *Proc Soc Exp Biol Med* **84** (2) (1953), pp. 506-508.

[100] T. Akiyama, J. Ishida, S. Nakagawa, H. Ogawara, S. Watanabe, N. Itoh *et al.*, Genistein, a specific inhibitor of tyrosine-specific protein kinases, *J Biol Chem* **262** (12) (1987), pp. 5592-5595.

[101] A.H. Lin, G.P. Leung, S.W. Leung, P.M. VanhoutteandR.Y. Man, Genistein enhances relaxation of the spontaneously hypertensive rat aorta by transactivation of epidermal growth factor receptor following binding to membrane estrogen receptors-

alpha and activation of a G protein-coupled, endothelial nitric oxide synthase-dependent pathway, *Pharmacol Res* **63** (3) pp. 181-189.

[102] R. Vera, M. Galisteo, I.C. Villar, M. Sanchez, A. Zarzuelo, F. Perez-Vizcaino *et al.*, Soy isoflavones improve endothelial function in spontaneously hypertensive rats in an estrogen-independent manner: role of nitric-oxide synthase, superoxide, and cyclooxygenase metabolites, *J Pharmacol Exp Ther* **314** (3) (2005), pp. 1300-1309.

[103] Q.C. Liao, Y.L. Li, Y.F. Qin, L.D. Quarles, K.K. Xu, R. Li *et al.*, Inhibition of adipocyte differentiation by phytoestrogen genistein through a potential downregulation of extracellular signal-regulated kinases 1/2 activity, *J Cell Biochem* **104** (5) (2008), pp. 1853-1864.

[104] Z.M. Shao, Z.Z. Shen, J.A. FontanaandS.H. Barsky, Genistein's "ER-dependent and independent" actions are mediated through ER pathways in ER-positive breast carcinoma cell lines, *Anticancer Res* **20** (4) (2000), pp. 2409-2416.

[105] I.F. Benter, J.S. Juggi, I. Khan, M.H. Yousif, H. CanatanandS. Akhtar, Signal transduction mechanisms involved in cardiac preconditioning: role of Ras-GTPase, Ca2+/calmodulin-dependent protein kinase II and epidermal growth factor receptor, *Mol Cell Biochem* **268** (1-2) (2005), pp. 175-183.

[106] R.M. Fryer, J.E.J. Schultz, A.K. HsuandG.J. Gross, Pretreatment with tyrosine kinase inhibitors partially attenuates ischemic preconditioning in rat hearts, *American Journal of Physiology-Heart and Circulatory Physiology* **44** (6) (1998), pp. H2009-H2015.

[107] S. Okubo, Y. Tanabe, K. Takeda, M. Kitayama, S. Kanemitsu, R.C. Kukreja *et al.*, Pretreatment with tyrosine kinase inhibitor attenuates the reduction of apoptosis 24 h

after ischemic preconditioning, *Jpn J Physiol* **54** (2) (2004), pp. 143-151.

[108] P. Htun, W.D. Ito, I.E. Hoefer, J. SchaperandW. Schaper, Intramyocardial infusion of FGF-1 mimics ischemic preconditioning in pig myocardium, *J Mol Cell Cardiol* **30** (4) (1998), pp. 867-877.

[109] M. Krenz, C.P. Baines, G. Heusch, J.M. DowneyandM.V. Cohen, Acute alcohol-induced protection against infarction in rabbit hearts: differences from and similarities to ischemic preconditioning, *J Mol Cell Cardiol* **33** (11) (2001), pp. 2015-2022.

[110] N. Hedayati, S.J. Schomisch, J.L. Carino, J. Timothy Sherwood, E.J. LesnefskyandB.L. Cmolik, Cardioprotection by St Thomas' solution is mediated by protein kinase C and tyrosine kinase, *J Surg Res* **113** (1) (2003), pp. 121-127.

[111] T.L. Wang, Y.H. HuangandH. Chang, Somatostatin analogue mimics acute ischemic preconditioning in a rat model of myocardial infarction, *J Cardiovasc Pharmacol* **45** (4) (2005), pp. 327-332.

[112] R. Germack, M. GriffinandJ.M. Dickenson, Activation of protein kinase B by adenosine A1 and A3 receptors in newborn rat cardiomyocytes, *J Mol Cell Cardiol* **37** (5) (2004), pp. 989-999.

[113] S. Lee, G. Chanoit, R. McIntosh, D.A. ZvaraandZ. Xu, Molecular mechanism underlying Akt activation in zinc-induced cardioprotection, *Am J Physiol Heart Circ Physiol* **297** (2) (2009), pp. H569-575.

[114] M. Shilkrut, G. Yaniv, R. Asleh, A.P. Levy, S. LarischandO. Binah, Tyrosine kinases inhibitors block Fas-mediated deleterious effects in normoxic and hypoxic ventricular myocytes, *J Mol Cell Cardiol* **35** (10) (2003), pp. 1229-1240.

[115] S. Negoro, K. Kunisada, E. Tone, M. Funamoto, H. Oh, T. Kishimoto *et al.*,

Activation of JAK/STAT pathway transduces cytoprotective signal in rat acute myocardial infarction, *Cardiovasc Res* **47** (4) (2000), pp. 797-805.

[116] K. Boengler, D. Hilfiker-Kleiner, H. Drexler, G. HeuschandR. Schulz, The myocardial JAK/STAT pathway: from protection to failure, *Pharmacol Ther* **120** (2) (2008), pp. 172-185.

[117] Y.T. Xuan, Y. Guo, H. Han, Y. ZhuandR. Bolli, An essential role of the JAK-STAT pathway in ischemic preconditioning, *Proc Natl Acad Sci U S A* **98** (16) (2001), pp. 9050-9055.

[118] Y. Katsube, H. Yokoshiki, L. Nguyen, M. YamamotoandN. Sperelakis, Inhibition of Ca2+ current in neonatal and adult rat ventricular myocytes by the tyrosine kinase inhibitor, genistein, *Eur J Pharmacol* **345** (3) (1998), pp. 309-314.

[119] C.E. Chiang, S.A. Chen, M.S. Chang, C.I. LinandH.N. Luk, Genistein directly inhibits L-type calcium currents but potentiates cAMP-dependent chloride currents in cardiomyocytes, *Biochem Biophys Res Commun* **223** (3) (1996), pp. 598-603.

[120] T. Ogura, L.M. ShubaandT.F. McDonald, L-type Ca2+ current in guinea pig ventricular myocytes treated with modulators of tyrosine phosphorylation, *Am J Physiol* **276** (5 Pt 2) (1999), pp. H1724-1733.

[121] C. Sims, J. ChiuandR.D. Harvey, Tyrosine phosphatase inhibitors selectively antagonize beta-adrenergic receptor-dependent regulation of cardiac ion channels, *Mol Pharmacol* **58** (6) (2000), pp. 1213-1221.

[122] S. Sorota, Tyrosine protein kinase inhibitors prevent activation of cardiac swelling-induced chloride current, *Pflugers Arch* **431** (2) (1995), pp. 178-185.

[123] L.M. Shuba, T. Asai, S. PelzerandT.F. McDonald, Activation of cardiac chloride

conductance by the tyrosine kinase inhibitor, genistein, *Br J Pharmacol* **119** (2) (1996), pp. 335-345.

[124] L.C. Hool, L.M. MiddletonandR.D. Harvey, Genistein increases the sensitivity of cardiac ion channels to beta-adrenergic receptor stimulation, *Circ Res* **83** (1) (1998), pp. 33-42.

[125] M. Nishio, Y. Habuchi, H. Tanaka, J. Morikawa, T. OkanoueandK. Kashima, Tyrosine kinase-dependent modulation by interferon-alpha of the ATP-sensitive K+ current in rabbit ventricular myocytes, *FEBS Lett* **445** (1) (1999), pp. 87-91.

[126] A. Stadnicka, W.M. Kwok, D.C. WarltierandZ.J. Bosnjak, Protein tyrosine kinase-dependent modulation of isoflurane effects on cardiac sarcolemmal K(ATP) channel, *Anesthesiology* **97** (5) (2002), pp. 1198-1208.

[127] Z. Gao, C.P. Lau, T.M. WongandG.R. Li, Protein tyrosine kinase-dependent modulation of voltage-dependent potassium channels by genistein in rat cardiac ventricular myocytes, *Cell Signal* **16** (3) (2004), pp. 333-341.

[128] A. Arora, M.G. NairandG.M. Strasburg, Antioxidant activities of isoflavones and their biological metabolites in a liposomal system, *Arch Biochem Biophys* **356** (2) (1998), pp. 133-141.

[129] R.M. Han, Y.X. Tian, Y. Liu, C.H. Chen, X.C. Ai, J.P. Zhang *et al.*, Comparison of flavonoids and isoflavonoids as antioxidants, *J Agric Food Chem* **57** (9) (2009), pp. 3780-3785.

[130] M.J. TikkanenandH. Adlercreutz, Dietary soy-derived isoflavone phytoestrogens. Could they have a role in coronary heart disease prevention?, *Biochemical Pharmacology* **60** (1) (2000), pp. 1-5.

[131] Q.H. Meng, P. Lewis, K. Wahala, H. AdlercreutzandM.J. Tikkanen, Incorporation of esterified soybean isoflavones with antioxidant activity into low density lipoprotein, *Biochim Biophys Acta* **1438** (3) (1999), pp. 369-376.

[132] R.C. Siow, F.Y. Li, D.J. Rowlands, P. de WinterandG.E. Mann, Cardiovascular targets for estrogens and phytoestrogens: transcriptional regulation of nitric oxide synthase and antioxidant defense genes, *Free Radic Biol Med* **42** (7) (2007), pp. 909-925.

[133] G.E. Mann, B. Bonacasa, T. IshiiandR.C. Siow, Targeting the redox sensitive Nrf2-Keap1 defense pathway in cardiovascular disease: protection afforded by dietary isoflavones, *Current Opinion in Pharmacology* **9** (2) (2009), pp. 139-145.

[134] G.E. Mann, D.J. Rowlands, F.Y. Li, P. de WinterandR.C. Siow, Activation of endothelial nitric oxide synthase by dietary isoflavones: role of NO in Nrf2-mediated antioxidant gene expression, *Cardiovasc Res* **75** (2) (2007), pp. 261-274.

[135] K. Mahn, C. Borras, G.A. Knock, P. Taylor, I.Y. Khan, D. Sugden *et al.*, Dietary soy isoflavone induced increases in antioxidant and eNOS gene expression lead to improved endothelial function and reduced blood pressure in vivo, *FASEB J* **19** (12) (2005), pp. 1755-1757.

[136] H. SiandD. Liu, Genistein, a soy phytoestrogen, upregulates the expression of human endothelial nitric oxide synthase and lowers blood pressure in spontaneously hypertensive rats, *J Nutr* **138** (2) (2008), pp. 297-304.

[137] K. Suzuki, H. Koike, H. Matsui, Y. Ono, M. Hasumi, H. Nakazato *et al.*, Genistein, a soy isoflavone, induces glutathione peroxidase in the human prostate cancer cell lines LNCaP and PC-3, *Int J Cancer* **99** (6) (2002), pp. 846-852.

[138] C. Borras, J. Gambini, M.C. Gomez-Cabrera, J. Sastre, F.V. Pallardo, G.E. Mann *et*

al., Genistein, a soy isoflavone, up-regulates expression of antioxidant genes: involvement of estrogen receptors, ERK1/2, and NFkappaB, *FASEB J* **20** (12) (2006), pp. 2136-2138.

[139] K.W. Kang, S.H. ChoiandS.G. Kim, Peroxynitrite activates NF-E2-related factor 2/antioxidant response element through the pathway of phosphatidylinositol 3-kinase: the role of nitric oxide synthase in rat glutathione S-transferase A2 induction, *Nitric Oxide* **7** (4) (2002), pp. 244-253.

[140] S. Joy, R.C. Siow, D.J. Rowlands, M. Becker, A.W. Wyatt, P.I. Aaronson *et al.*, The isoflavone Equol mediates rapid vascular relaxation: Ca2+-independent activation of endothelial nitric-oxide synthase/Hsp90 involving ERK1/2 and Akt phosphorylation in human endothelial cells, *J Biol Chem* **281** (37) (2006), pp. 27335-27345.

[141] D. Liu, L.L. HomanandJ.S. Dillon, Genistein acutely stimulates nitric oxide synthesis in vascular endothelial cells by a cyclic adenosine 5'-monophosphate-dependent mechanism, *Endocrinology* **145** (12) (2004), pp. 5532-5539.

[142] N.R. Bianco, L.J. ChaplinandM.M. Montano, Differential induction of quinone reductase by phytoestrogens and protection against oestrogen-induced DNA damage, *Biochem J* **385** (Pt 1) (2005), pp. 279-287.

[143] R.M. Fryer, J.E. Schultz, A.K. HsuandG.J. Gross, Pretreatment with tyrosine kinase inhibitors partially attenuates ischemic preconditioning in rat hearts, *Am J Physiol* **275** (6 Pt 2) (1998), pp. H2009-2015.

[144] R.M. Fryer, J.E. Schultz, A.K. HsuandG.J. Gross, Importance of PKC and tyrosine kinase in single or multiple cycles of preconditioning in rat hearts, *Am J Physiol* **276** (4 Pt 2) (1999), pp. H1229-1235.

[145] R.M. Fryer, Y. Wang, A.K. Hsu, H. NagaseandG.J. Gross, Dependence of delta1-opioid receptor-induced cardioprotection on a tyrosine kinase-dependent but not a Src-dependent pathway, *J Pharmacol Exp Ther* **299** (2) (2001), pp. 477-482.

[146] D.A. Liem, C.C. Gho, B.C. Gho, S. Kazim, O.C. Manintveld, P.D. Verdouw *et al.*, The tyrosine phosphatase inhibitor bis(maltolato)oxovanadium attenuates myocardial reperfusion injury by opening ATP-sensitive potassium channels, *J Pharmacol Exp Ther* **309** (3) (2004), pp. 1256-1262.

[147] D. Fuchs, P. Erhard, R. Turner, G. Rimbach, H. DanielandU. Wenzel, Genistein reverses changes of the proteome induced by oxidized-LDL in EA.hy 926 human endothelial cells, *J Proteome Res* **4** (2) (2005), pp. 369-376.

[148] D. Fuchs, B. Dirscherl, J.H. Schroot, H. DanielandU. Wenzel, Soy extract has different effects compared with the isolated isoflavones on the proteome of homocysteine-stressed endothelial cells, *Mol Nutr Food Res* **50** (1) (2006), pp. 58-69.

[149] D. Fuchs, P. Erhard, G. Rimbach, H. DanielandU. Wenzel, Genistein blocks homocysteine-induced alterations in the proteome of human endothelial cells, *Proteomics* **5** (11) (2005), pp. 2808-2818.

[150] D.H. Zhang, Y.C. Tai, C.H.S. Wong, L.K. Tai, E.S.C. KoayandC.S. Chen, Molecular response of leukemia HL-60 cells to genistein treatment, a proteomics study, *Leukemia Research* **31** (1) (2007), pp. 75-82.

[151] A.M. Sotoca, M.D. Sollewijn Gelpke, S. Boeren, A. Strom, J.A. Gustafsson, A.J. Murk *et al.*, Quantitative proteomics and transcriptomics addressing the estrogen receptor subtype-mediated effects in T47D breast cancer cells exposed to the phytoestrogen genistein, *Mol Cell Proteomics* pp.

[152] J. Wang, A.M. Betancourt, J.A. MobleyandC.A. Lamartiniere, Proteomic discovery of genistein action in the rat mammary gland, *J Proteome Res* **10** (4) pp. 1621-1631.

[153] C. Rowell, D.M. CarpenterandC.A. Lamartiniere, Chemoprevention of breast cancer, proteomic discovery of genistein action in the rat mammary gland, *J Nutr* **135** (12 Suppl) (2005), pp. 2953S-2959S.

[154] S.M. White, P.E. ConstantinandW.C. Claycomb, Cardiac physiology at the cellular level: use of cultured HL-1 cardiomyocytes for studies of cardiac muscle cell structure and function, *Am J Physiol Heart Circ Physiol* **286** (3) (2004), pp. H823-829.

[155] W.C. Claycomb, N.A. Lanson, Jr., B.S. Stallworth, D.B. Egeland, J.B. Delcarpio, A. Bahinski *et al.*, HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte, *Proc Natl Acad Sci U S A* **95** (6) (1998), pp. 2979-2984.

CHAPTER 2

PHOSPHOPROTEOMICS AND MOLECULAR CARDIOLOGY: TECHNIQUES, APPLICATIONS AND CHALLENGES

1. Introduction

The reversible, covalent binding of phosphate groups to proteins is the most common post-translational modification involved in many cellular processes. It has been estimated that close to 50% of the proteins in the mammalian proteome can potentially be phosphorylated. There are more than 100,000 predicted phosphorylation sites in the human proteome of which fewer than 2000 are documented [1, 2]. O-phosphorylation on serine (~90%), threonine (~10%) and tyrosine (~<1%) [3] are the most frequent form of protein phosphorylation, whereas N/S- phosphorylations on histidine and cysteine [4-6] exist in much lower quantities.

Phosphorylation usually causes a protein conformational change for specific functions, or forms motifs recognized by other molecules for interactions including the assembly and dissociation of protein complexes [7-9]. Protein phosphorylation is regulated by kinases and phosphatases, which work in an opposing manner to maintain a particular phosphorylation state. When an internal or external stimulus occurs, many kinases/phosphatases themselves are activated via phosphorylation and cause the phosphorylation of other proteins on a timescale of seconds to minutes. Such signaling

can further be amplified, followed by propagation to downstream kinases, or by cross-talking to many other pathways, which eventually leads to changes in cellular functions such as enzyme activity modulation, cytoskeleton remodeling, and gene expression regulation. Disruptions of these signaling cascades have been linked to several cardiovascular diseases such as ischemia/reperfusion injury and cardiac hypertrophy [10, 11].

Phosphoproteomics, characterizing protein phosphorylation in large-scale with modern high-throughput protein separation and mass spectrometry technologies, has played a significant role in our attempt to unveil the complexity of protein phosphorylation networks and how they link to cardiac pathological processes. Compared to standard proteomics approaches, phosphoproteomics poses some additional challenges as it tries to identify, and in some cases to quantify, the dynamic and reversible phosphorylation process with low *in vivo* occurrence. Phosphoproteomics is a multifaceted and open-ended field where a tremendous amount of new technologies and applications have been introduced every year to broaden the coverage of the phosphoproteome, to measure or quantify phosphorylation stoichiometry and dynamics more accurately, and to extract biological relevance and significance from large-scale phosphoproteomics data. Fig 2.1 summarizes some of the most important technical breakthroughs in the field of MS-based phosphoproteomics. Fig 2.2 summarizes a sequential scheme of typical phosphoproteomic workflows from sample preparation to data analysis. Given the sizeable variety of phosphoproteomic techniques and applications, here we focus on the most mainstream techniques and strategies in this rapidly evolving endeavor with

illustrative phosphoproteomic applications in the field of molecular cardiology.

2. General Sample Preparation Strategies

As is the case in a standard proteomics workflow, a phosphoproteomics workflow usually starts with protein extraction from tissue samples, cell cultures, or subcellular components by sonication, liquid nitrogen grinding, or homogenization with a lysis buffer that can dissolve and often denature proteins. However, there are two critical aspects that need to be considered during the initial sample preparation in order to detect protein phosphorylation using MS. First, as phosphoproteins only represent a very tiny portion of the total proteome, phosphoproteomic analysis usually requires a large amount of starting material, typically in the range of 1-10 mg of protein. However, successful detection of protein phosphorylation is still possible using a few hundred micrograms of protein if an appropriate enrichment strategy is used and the mass spectrometer has superior sensitivity [12]. Another key aspect of sample preparation is the choice of lysis buffer. Protein extraction should be accomplished at low temperatures in the presence of both protease and phosphatase inhibitor cocktails to prevent proteolysis and dephosphorylation during protein extraction. Typical protease inhibitors are phenylmethylsulfonyl fluoride (PMSF), aminoethyl benzylsulfonyl fluoride (AEBSF), ethylene diamine tetraacetic acid (EDTA), pepstatin, benzamidine, leupeptin, and aprotinin [13, 14]. Phosphatase inhibitor cocktails contain components such as sodium ortho-vanadate, imidazole, sodium tartrate, EDTA, okadaic acid, β-glycerophosphoric acid, and sodium pyrophosphate [14, 15]. Chaotropic agents such as HCl-guanidine or urea, and detergents such as SDS are included in lysis buffers to dissolve hydrophobic

proteins and to denature proteases and phosphatases released during the lysis procedure. It is also suggested that metal parts should be avoided during sample preparation chromatographic workflow to minimize unwanted sample loss due to affinity adsorption of phosphoproteins to metals [14].

3. Subcellular Fractionation

Over the last few years, most phosphoproteomic studies have been carried out to document the extent and dynamics of protein phosphorylation in subproteomes of cardiomyocytes. As in eukaryotes, the proteome are very complex with over 10,000 proteins - a number that is beyond the resolving power of most proteomic techniques. Additionally**,** many regulatory kinases and phosphatases have an uneven spatial distribution, and many signaling pathways are regulated in a spatial manner. Many proteins that are phosphorylated or dephosphorylated in different cell compartments carry out vital molecular functions such as protein translocation and protein complex scaffolding. Thus, to portray the spatial occurrence of protein phosphorylation and its corresponding biological significance, it is essential to study the phosphoproteome with spatial-temporal resolution by incorporating subcellular fractionation procedures into the workflow. Trost and coworkers have provided a thorough review on the applications of subcellular phosphoproteomics [15].

Differential density centrifugation is currently the most popular method for organelle enrichment. The easy setup of centrifugation makes it an ideal fit into a proteomics

workflow. Briefly, after homogenization, nuclei, other organelles, and cytosol are fractionated by gradient centrifugation in a solution such as sucrose. Proteins that are extracted from each fraction are representative of the different cellular compartments [16]. Other techniques such as electromigration, free-flow electrophoresis, recombinant-protein pull down, coimmunoprecipitation, epitope-tagged proteins and tandem affinity tags have also been successfully applied for organelle isolation. A comprehensive overview of the application of these different subcellular fractionation methods has been reviewed previously [16-21].

The major challenge in subcellular phosphoproteomics is sample availability for successful phosphopeptide enrichment. The total amount of protein from each subcellular fraction is far less than that in the total lysate, and this is further complicated by the fact that some phosphoproteins are usually present in very low amounts in individual subcellular fractions. The enrichment of phosphopeptides for a typical multidimensional LC-MS analysis often requires at least 1-30 mg of total lysate, a quantity that is difficult to obtain for many subcellular structures. Thus, sample scale-up must consider the relative proportion of target organelle to cell mass, which can range from less than 1% (peroxisome) to 10% (the nucleus) [16].

Compared to other cellular organelles, the phosphoproteome of mitochondria has been studied the most [22-29], as the mitochondria play vital roles in energy metabolism, oxidative radical species production, apoptosis, ion balance, calcium signaling, and myofibril organization, as well as metabolism of amino acids, lipids, and iron [30-32].

Links between mitochondria physiology and cardiac diseases such as ischemia/reperfusion injury and cardiomyopathy have been well documented [31]. The proteasome is also of interest. Several studies of the cardiac proteasome phosphoproteome have shown the importance of phosphorylation of certain components in the dysregulation of protein degradation and accumulation of toxic aggregates associated with many cardiac diseases [33-35]. Although most subcellular proteomics and phosphoproteomics studies focus mainly on cardiac mitochondria and proteasome, it is also worthwhile to investigate other cardiac cellular components such as the plasma membrane, lipid rafts, golgi apparatus, and endoplasmic reticulum that have significant impact on cardiac functions.

4. 2DE Workflow

Complex protein mixtures can be mapped and characterized by high-resolution two-dimensional electrophoresis (2DE) gel analysis. In a 2DE workflow, the proteins are first separated according to isoelectric point by immobilized pH gradient (IPG)-gel based isoelectric focusing (IEF) and then separated by molecular weight via regular SDS-PAGE gel [36]. Proteins separated by 2DE can be visualized by a variety of staining protocols that have been summarized by Steinberg [37]. Phosphoproteins can be visualized by ^{32}P / ^{33}P autoradiography, phosphospecific stains or by western blotting techniques. Protein phosphorylation/dephosphorylation usually lead to changes in pI, thus cause horizontal shift in spot position. After gel image analysis, spots of interest can be excised from the polyacrylamide gel, and digested by proteases such as trypsin. The proteolytic peptides extracted from gel plugs are then identified by MALDI-MS/MS or LC-ESI-/MS/MS. An

optional enrichment procedure can be used prior to MS/MS analysis to increases chances for the detection of phosphopeptides and determination of phosphosites.

Although 2DE remains a most popular tool in the field of proteomics and phosphoproteomics, there are multiple limitations associated with this method. First, the inherently limited resolving power of 2DE only allows separation of proteins within certain isoelectric point and molecular weight boundaries. Second, 2DE is generally biased against hydrophobic proteins, *e.g.* membrane proteins and nuclear proteins [38]. Third, none of the gel-based methods can be multiplexed to quantify protein phosphorylation from more than two samples. Finally, 2DE is labor intensive and hence difficult to automate, such that significant variance and artifacts are generated during the procedure. Gel-free LC-based techniques coupled to MS overcome some of these problems as discussed later in this review.

4.1. Autoradiography

Despite the tediousness of the procedure, *in vivo* or *in vitro* autoradioactive $^{32}P/^{33}P$ labeling of phosphate groups is the most sensitive method for detecting phosphoprotein on gels [39-41]. Moreover, this unbiased labeling technique gives the direct evidence of the presence of phosphor group in all kinds of amino acid residues and sequence motifs [40, 42]. Employing autoradiography, Chu et al. [43] resolved 120 phosphoproteins from 300 µg of protein extracted from ^{32}P-labeled mouse cardiomyocytes. The authors also suggest the ^{32}P- autoradiography demonstrates sensitivity superior to silver staining for some phosphoproteins [43].

4.2. Phosphoprotein stains

Compared to autoradiography, phosphor-specific fluorescent stains are easy to use and can detect phosphoproteins in any cellular state as they do not require protein turnover for the incorporation of the radioactive phosphate. The MS-compatible fluorescent dye, Pro-Q Diamond (Molecular Probes) [44, 45], binds specifically to phosphoproteins with sensitivity at the nanogram level. Hopper et al. [26] applied both Pro-Q Diamond stain and ^{32}P-labeling to detect phosphoproteins from 500 µg of porcine cardiac mitochondria extract. Over 200 phosphoproteins were resolved via fluorescent staining, demonstrating comparable sensitivity to that of radioactive labeling. However, the author also reported notable differences in the stain patterns between two methods. A similar conclusion was drawn in another cardiac mitochondria phosphoproteomics study from the same research group [22] comparing the commercial fluorescent stain Phos-Tag 540 (Perkin-Elmer) [46] with ^{32}P-labeling.

4.3. Immunoblotting

Immunoblotting can also be used for detection of phosphoproteins in a 2D gel. General procedures and antibody choices has been reviewed in detail [47, 48]. Most western-blotting based phosphoproteomic studies use only anti-p-Tyr antibodies, which are more specific than anti-p-Ser/Thr antibodies. However, global analysis of all Ser/Thr/Tyr phosphosites by western blotting is achievable. Feng et al. [25] successfully detected and quantified 61 rat cardiac mitochondrial phosphoproteins using monoclonal anti-p-Ser/Thr/Tyr antibodies (LuBioScience) on native BN-PAGE gel. In another study by

Zong et al. [34], three different anti-p-Ser, anti-p-Thr and anti-p-Tyr antibodies were used to survey the phosphorylation pattern of the cardiac 20S proteasome complex. Of 14 subunits, two were detected by all three antibodies, three were only detected by anti-p-Ser antibodies and one only by Pro-Q phosphostain. Despite successes in small-scale organelle phosphoproteomics studies, it is apparent that the major drawback of antibody-based approaches is the variation in sensitivities and specificities of antibodies, which limits the use of western blotting as a high-throughput quantitative approach [41]. Perhaps the true merit of western blotting-based visualization comes when one desires to study the phosphorylation targets of a known kinase or kinase group. Antibodies against specific phosphorylation sites or motifs within the substrate sequence of a certain kinase can unveil highly specific targeted information relevant to the architecture of the signaling network. Using anti-p-Tyr antibody blotting, Schwertz et al. [29] demonstrated the link between p38 MAP kinase inhibition and tyrosine phosphorylation reduction of troponin T, VDAC-1 and HSP73 which contribute to the cardioprotection of rabbit cardiomyocytes from I/R injury.

5. Liquid Chromatographic Methods

The multidimensional liquid chromatography (MDLC) method was first introduced as a separation method alternative to 2DE [49], and has now been widely utilized in almost every branch of proteomics including phosphoproteomics. Compared to a 2DE workflow, LC-based methods have a greater ability to separate peptides with similar chemical properties compare to proteins, thus making them a more high-throughput method

compared to gel-based approaches. Other advantages of MDLC workflows include easy automation, minimum sample loss, and high reproducibility [50-53]. Most MDLC-based proteomic workflows, or 'shotgun' workflows [50], consist of two dimensions of LC separations that ideally are orthogonal to one another to enhance the overall resolution and minimize MS instrumental undersampling. Briefly, a protein mixture is first digested by a protease or two, in some cases. The resulting peptide mixture is fractionated, typically by strong cation exchange (SCX) chromatography. Individual fractions of the peptide mixture can be collected off-line prior to a second LC separation coupled with ESI-MS/MS. Alternatively, peptides eluted from the first LC separation can be directly fed into a 2^{nd} LC system in a MudPIT fashion [51]. As the last LC separation is usually coupled with ESI-MS/MS using reverse phase (RP) HPLC and thus differentiates peptides on the basis of hydrophobicity, current MDLC protocols for fractionation of phosphopeptides use a first dimensional chromatography that capitalizes on the one of two distinctive properties of phosphopeptides: the additional negative charges and higher hydrophilicity of phosphopeptides due to the presence of the phosphate group. Strong cation exchange (SCX), strong anion exchange (SAX), hydrophilic interaction liquid chromatography (HILIC), and electrostatic repulsion hydrophilic interaction chromatography (ERLIC) have been selected for phosphoproteomics applications. The goal of developing a successful first dimension LC separation is two-fold: to concentrate peptides, including phosphopeptides, into less complex mixture fractions for downstream procedures like phosphopeptides enrichment [53, 54]; and to partially isolate phosphopeptides from non-phosphorylated peptides. The reader is referred to excellent, comprehensive reviews of MDLC workflows in phosphoproteomics [39, 40, 55].

5.1. Strong cation exchange liquid chromatography

Strong cation exchange (SCX) fractionation has been used extensively for the fractionation of phosphopeptides [15, 56-58]. A typical SCX procedure separates peptides according to their electrostatic properties by increasing ionic strength in the mobile phase. At low pH (<2.6), the majority of tryptic peptides have at least two positive charges because all amino groups from the N-terminus and Arg/Lys side chains will be protonated, and all acidic carboxyl groups will also be protonated. Protonated peptides will be strongly retained by the SCX stationary phase surface, which displays negative charges. Phosphopeptides bearing negatively charged phosphate groups are either 1+, neutral, or negatively charged at pH 2.6, and therefore will elute out from SCX earlier than other species [57, 58]. Theoretically, by carefully selecting the pH of the mobile phase, SCX can isolate phosphopeptides from the abundant acidic peptides that would otherwise compete with phosphopeptides in a metal-affinity enrichment step. However, since phosphopeptides have been identified in all SCX fractions in most phosphoproteomics studies, SCX fractionation is usually coupled with other enrichment methods, such as metal oxide affinity chromatography (MOAC) [59-63]. This is possibly due to the fact that tryptic phosphopeptides may have 2^+ charges if they contain His or multiple Arg/Lys groups due to partial digestion, which is more common for phosphopeptides [58]. Despite the claims of success of using other chromatographic method for separating phosphopeptides such as HILIC or ERLIC, SCX still shown superior performance to cover large amount of phosphopeptides in a recent comparison

study paralleling SCX-TiO_2, HILIC-TiO_2, ERLIC-TiO_2 to enrich phosphopeptides from 4 mg of HeLa protein [64].

5.2. Strong anion exchange liquid chromatography

As phosphopeptides tend to be more negatively charged than other peptides, strong anion exchange (SAX) chromatography is a natural choice for their fractionation. Theoretically, phosphopeptides have stronger retention in an SAX column than other unmodified peptides which interact weakly with SAX materials. This means that SAX alone can isolate phosphopeptides from a mixture, and fractionate these compounds under gradient elution [65]. Dai, et al. developed a fully automatic SAX/RP-LC-MS/MS MudPIT procedure using an MS-compatible mobile phase utilizing a pH gradient to enrich and separate phosphopeptides from 0.5 mg HeLa cell total lysate, from which 1561 phosphopeptides were identified [66]. In another automatic SAX/RP-LC-MS/MS MudPIT system, Wang, et al. identified 1554 unique phosphopeptides from 1.5 mg of human liver tissue extract. Recently, different groups have described protocols combining SAX separation combined with other enrichment methods such as MOAC or IMAC [67, 68]. Nie, et at., developed a strategy based on SAX fractionation followed by flow-through enrichment by TiO_2 (AFET), by which 2466 unique phosphopeptides were identified from only 0.5 mg of protein digest from HeLa cell [68]. In another interesting case describing a 'Yin-Yang MDLC' procedure, Dai et al. combined the isolation effect of SCX and the separation power of SAX by loading a phosphopeptide enriched flow-through fraction from SCX into an SAX for separation. In this protocol, over 800

phosphopeptides were identified from one mg of mouse liver tryptic digest without any further IMAC/MOAC enrichment [69]. Similarly, Motoyama et al. described a SAX/SCX MudPIT method to enhance the phosphopeptide identification using a salt step elution [70].

5.3. Hydrophilic interaction liquid chromatography

Previously used for the enrichment of glycoproteins and small polar metabolites, hydrophilic interaction liquid chromatography (HILIC) [71] is now drawing attention due to its applications in separating phosphopeptides. HILIC takes advantage of the greater hydrophilic nature if phosphopeptides compared to other peptides [23, 72-79]. In the HILIC procedure, as opposed to RPLC, peptides are introduced in a mobile phase that is relatively high in organic content interacting with the neutral hydrophilic stationary phase via hydrogen bonding. By gradually decreasing the organic contents in the mobile phase, peptides elute in order of increasing hydrophilicity [71, 72]. HILIC is considered to be truly orthogonal to reverse-phase chromatography and to offer much higher separation resolution than ion-exchange based LC methods [80]. Such resolution is critical in the HILIC application for separating multiphosphorylated peptide isomers as they are more likely to coelute in RPLC resulting in ambiguous MS/MS identifications [76]. Therefore, a HILIC-RP system can potentially yield much higher overall peak capacity and resolution than other 2DLC procedures. Unlike ion exchange methods, HILIC uses a mobile phase system based on an organic solvent gradient, and thus it can minimize sample loss by avoiding a desalting step. In fact, HILIC is directly compatible with ESI-

MS/MS [76]. Although HILIC provides good separations for phosphopeptides, alone it cannot separate phosphopeptides from other species. HILIC requires a pre- or post-enrichment step such as IMAC or MOAC in order to isolate phosphopeptides from complex backgrounds [23, 72-75, 77]. Using HILIC prefractionation, McNulty, et al. achieved IMAC selectivity over 95% [75]. Also, when using an IMAC-HILIC protocol, Wu, et al. identified 2857 unique phosphorylation sites in 1338 phosphoproteins from one mg of cell lysate [73].

5.4. Electrostatic repulsion hydrophilic interaction chromatography

Electrostatic repulsion hydrophilic interaction chromatography (ERLIC) is a relatively new LC method showing good applications in phosphoproteomics [81-83]. ERLIC utilizes both hydrophilic interactions and electrostatic repulsion as the separation mechanisms, and thus can potentially enrich and fractionate the phosphopeptides in one step [83]. ERLIC is achieved by operates a weak anion exchange (WAX) column with a high organic content (70% ACN) mobile phase at a low pH. Under such conditions, the anionic phosphopeptides can be selectively retained in the WAX column, while non-ionized peptides will be washed out by the organic solvent, and peptides with protonated carboxyl groups are electrostatically repulsed by the column [83]. Gan, et al., compared ERLIC to SCX-IMAC for the enrichment and fractionation of phosphopeptides [82]. In their study, ERLIC alone and SCX/IMAC coupled with the MS^3 identification strategy detected 926 and 1,315 unique phosphopeptides, respectively, from a total of 10 mg of human epithelial carcinoma cell lysate. Interestingly, ERLIC identified a higher number

of multiple phosphorylated peptides when compared to the SCX/IMAC procedure. The results suggest that both methods are complementary to each other, as there was only 12% overlap of unique phosphopeptides identified by both methods [82]. This phenomenon was reported in another paper suggesting a ERLIC-TiO_2 procedure can be used to enrich the flow-through fraction from SCX to reach better coverage for multiphosphorylated peptides [64].

6. Affinity Enrichment Strategies

6.1. General considerations of using enrichment strategy

As researchers have gained experience with strategies to identify protein phosphorylations from a complex protein mixture, such as mammalian cell or tissue total lysate, there has been a trend to implement some form of enrichment technique(s) in order to separate phosphopeptides or phosphoproteins from un-modified peptide species before MS analysis. This is mainly attributed to the following reasons:

- Most key phosphoproteins to the biological process in question, such as signaling molecules and transcriptional factors, are predominantly in low levels;
- Often with any given biological state, very few protein species containing potential phosphorylation sites are actually phosphorylated, or are phosphorylated in a low stoichiometry. Further, not all phosphosites are modified on proteins containing multiple phosphorylation sites. The technical difficulty resulting from low stoichiometry of phosphorylation is obvious: the proteolytic digest of phosphoprotein

generates more non-phosphorylated peptide species than phosphorylated species, and other peptides are generated from unphosphorylated proteins;

- Phosphopeptides are generally difficult to ionize using mass spectrometry, as they can easily be ion-suppressed by non-phosphorylated counterparts under positive-mode ionization, thus resulting in low signal intensities in a complex mixture.

The methods for phosphoprotein or phosphopeptide enrichment fall into three major categories, namely, antibody-based approaches, affinity-based approaches, and chemical derivatization approaches [84]. In general, the following suggestions have been given for development of a practical enrichment protocol for phosphoproteomics:

- Since different methods tend to exhibit bias towards certain peptide populations, and each only reveals a portion of the whole phosphoproteome, a combination of different methods should be used in order to expand the coverage of phosphopeptide species;
- For complex samples such as mammalian cells or tissue lysates, a pre-fractionation step such as 1D-SDS-PAGE, SCX or HILIC chromatography is recommended prior to any enrichment step, as samples with high complexity may introduce severe non-specific binding or /reactions during enrichment steps;
- Use large amounts of starting material to ensure sensitivity within sample and budgetary constraints;
- Limit sample-handling steps to minimize sample loss. Compromise between these various tradeoffs must be made in order to balance the sensitivity, selectivity, recovery rate and practicality of each protocol.

6.2. Immunoaffinity method

Antibody-based methods, *e.g.* immunoprecipitation and immunoaffinity chromatography, are targeted to certain types of phosphorylation events, such as the anti-P-Tyr antibody approach [85-88], or to known motifs such as the substrate sequences of a certain group of kinases. Generally phosphoproteins or phosphopeptides bind to antibodies by sequence specific interactions at normal physiological environment, and such interaction can be disrupted by high level of detergent with phosphor analog such as phenylphosphate or by heat for elution. It is noteworthy that anti-P-Tyr antibody provides very highly specific solution for targeting tyrosine kinase signaling over other nonspecific enrichment methods. This is because p-Tyr only occurs in extremely low frequency and is hard to detect if analyzed with large amount of p-Ser/Thr phosphopeptides in a LC-MS/MS experiment. On the other hand, one should be cautious when using anti-p-Ser and anti-p-Thr immunoaffinity chromatography due to their limited specificity [89].

6.3. Immobilized metal affinity chromatography

Immobilized metal affinity chromatography (IMAC) is a universal enrichment technique capable of non-specific capture of phosphopeptides, regardless of the phosphorylation type and the local sequence. Immobilized ferric ion (III) was first utilized to purify phosphoproteins and phosphopeptides [90, 91]. Other multivalent metal cations, such as gallium (III) [92], aluminum (III) [93], zirconium (II) [94], nickel (II), and titanium (IV) [95] were found to have similar binding affinities with protein-bound phosphates at low pH. Typically, metal cations are immobilized to a supporting matrix via chelation with either nitriloacetic acid (NTA) or iminodiacetic acid (IDA). Phosphopeptides generally

bind to immobilized metal cations through electrostatic interactions at low pH (<3), and dissociate at a higher pH (>7). Currently, a variety of different IMAC resins are commercially available in LC-column, SPE column, spin column, LC-tip and magnetic bead formats, comprising various types of immobilized metal ions, all of which demonstrate different enrichment efficiencies and specificities. However, a significant challenge associated with the IMAC method is non-specific binding of peptides rich in acidic amino acids to the IMAC. One way to alleviate this problem is O-methyl esterification of all carboxylate groups prior to IMAC enrichment [96]. This procedure has not been adopted by many researchers as the reaction conditions are difficult to control and incomplete esterification, unwanted byproducts and significant samples losses have been reported [97]. Another way to circumvent the non-specific binding of acidic peptides is to use endoproteinase glu-C rather than trypsin for protein digestion. As glu-C cleaves the protein at the C-terminus of Glu and Asp residues, the resulting proteolytic peptides only carry one acidic amino acid residue. Another challenge reported recently is that the IMAC is less efficient for enrichment of mono-phosphorylated peptides than for multiple phosphorylated species [98, 99]. In order to obtain the complete set of phosphopeptides from a complex sample, Ndassa et al. developed an improved IMAC protocol giving much higher recovery rates for peptides of single and multiple phosphorylations by modifying the binding and washing buffer conditions (0.1% acetic acid in 1:1:1 acetonitrile/methanol/water) [100]. In another case, Thinghole et al. reported a protocol using IMAC in conjunction with a TiO_2-MOAC procedure, which is biased towards the mono-phosphorylated peptides [101]. Using this sequential IMAC (SIMAC) technique, they were able to enrich 186 multiple and 306 single phosphorylated

peptides from a human mesenchymal stem cell lysate digest.

6.4. Metal oxide affinity chromatography

Recently, much attention has been drawn to the use of metal oxide affinity chromatography (MOAC) for phosphopeptide enrichment from large complex samples due to its high recovery rate and selectivity [102-111]. A typical MOAC material includes a multivalent metal oxide such as titanium dioxide (TiO_2) [112], zirconium dioxide (ZrO_2) [113], aluminum oxide (Al_2O_3) [114], aluminum hydroxide ($Al(OH)_3$) [110] and niobium oxide (Nb_2O_5) [115]. Similar to the IMAC procedure, positively charged metal oxides can selectively capture phosphopeptides at low pH (<3) and release them at higher pH (>7). Like IMAC, MOAC also experiences non-specific binding of acidic peptides. However, because the phosphopeptides interact with metal oxides via a "bridging bidenatate" mechanism, a quenching agent like salicylic acid or 2,5-dihydroxybenzoic acid (DHB) must be used to competitively occupy binding sites on the MOAC resin to prevent adsorption of nonphosphorylated peptides [107, 109]. Using this quenching effect of DHB, Christensen, et al. successfully characterized over 10,000 phosphopeptides from digested cell extracts pre-fractionated by SCX prior to TiO_2-MOAC enrichment [116]. However, it is noteworthy that in a LC-ESI-MS experiment, residual DHB or salicylic acid co-elutes with phosphopeptides and may cause clogging in the RP column and may introduce ion suppression in the ESI source. Sugiyama et al. also reported that aliphatic hydroxy acids, such as lactic acid, β-hydroxypropanoic acid (HPA), phthalic acid, gallic acid, and glycolic acid also serve as "nonphosphopeptide excluders" by forming a cyclic chelate complex with the metal oxide [117, 118]. They suggest that

chelation between aliphatic hydroxy acids and metal oxides are weaker than that between phosphates and metal oxides but stronger than between the carboxylic group of amino acids and metal oxides. Because these acids are more hydrophilic than DHB, they can be removed by a C18 desalting procedure. Using lactic acid as an excluder in loading and washing steps, Sugiyama et al. identified over 1,100 phosphopeptides from less than one mg of HeLa cell cytoplasmic extracts with no additional LC pre-fractionation [117].

6.5. Chemical derivatization methods

Several chemical derivatization procedures have been introduced to assist in phosphopeptide enrichment [119], amongst which the three most widely used approaches are O-methyl-esterification of carboxylate groups, β-elimination of phosphorylated Ser/Thr, and phosphoramidate chemistry (PAC) [39, 40]. O-methyl esterification [96] was introduced as a method to alleviate non-specific binding of acidic peptides during IMAC/MOAC enrichment as discussed in section 6.3.

Using β-elimination at high pH, labile phosphoester bonds of modified Ser/Thr residues can be cleaved to form dehydroalanine and β-methyldehydroalanine respectively. By nucleophilic addition of free sulfhydryl groups to the unsaturated bonds via Michael's addition, the side chain of Ser/Thr can form stable thiol-based derivatives that can further be cross-linked to a biotin tag [120]. Once the phosphate group has been replaced by biotinylated moieties, the formerly phosphorylated proteins or peptides can be enriched by immobilized avidin affinity chromatography [120]. However, sample losses have been reported as biotin−avidin interaction is very strong and modification of peptides may cause changes in fragmentation patterns in the tandem MS that are difficult to interpret

[121]. Alternatively, McLachlin *et al*. demonstrated that after β-elimination and Michael's addition, the derivitized thiol group itself can be captured by thiol–sepharose affinity resins [121]. In another interesting application of β-elimination using cysteamine for Michael's addition, Knight *et al*. converted phosphorylated Ser/Thr residues into aminoethylcysteine and β-methylaminoethylcysteine, respectively. Both modified residues are lysine analogs that can be cleaved by lysine-specific proteases, *e.g*. trypsin and Lys-C [122]. In this way, peptides cleaved at modified Ser/Thr sites can be used to map sites of phosphorylation. However, due to the formation of diastereomeric aminoethylcysteine, only 50% of modified sites will be cleaved as trypsin only cleaves the R stereo-isoform of aminoethylcysteine. Moreover, when developing β-elimination based protocols, caution should be paid to side reactions that occur on glycopeptides, sulfopeptides, alkylated cysteine residues and some unmodified serine residues under the same conditions [123-125]. It is also noteworthy that all β-elimination based methods only apply to peptides with p-Ser/The, but not to peptides with p-Tyr.

In contrast, phosphoramidate chemistry (PAC) enables the derivatization of all types of phosphorylated amino acids. Instead of transforming phosphate groups to thiol moieties, the PAC approach directly forms a reactive phosphoramidate that can be coupled to glass beads [126] or to a dendrimer [127]. Phosphate groups are eventually regenerated by hydrolysis with an acid such as TFA. Bodenmiller *et al*. suggested multiple phosphopeptide enrichment strategies including PAC, IMAC, and TiO_2-MOAC approaches are complementary to each other as they all reproducibly enriched disparate and partially overlapping segments of phosphopeptide populations from a *D.*

melanogaster Kc167 cells lysate [128].

7. Identification of Phosphopeptides by Tandem Mass Spectrometry

7.1. General considerations

Application of cutting-edge tandem MS technologies with high sampling speed and high sensitivity have enabled researchers to characterize thousands of peptides from a wide variety of biological contexts including enriched phosphopeptide mixtures that would be impractical using other sequencing techniques such as Edman degradation.

In a common shotgun-based tandem MS phosphoproteomics experiment, a phosphopeptide-enriched mixture is first separated according to hydrophobicity in a nano-LC column filled with a reverse-phase (C18) material. In some cases, the LC separation is carried out in an automated LC-on-chip system which incorporates sample loading, peptide trapping, and LC separation in one mechanical unit [129-131]. The LC eluent is directly introduced into the mass spectrometer by electrospray ionization (ESI), which immediately disperses the sample, evaporates the solvent, and protonates peptides at multiple sites [132]. However, simple phosphopeptide mixtures from digests of single or small numbers of proteins can also be analyzed by matrix-assisted laser desorption/ionization (MALDI) [133] in which peptides are protonated via a matrix agent i.e. DHB, sinapic acid, excited by laser beams [134, 135]. The positively charged peptide ions then enter the mass analyzer where the mass-to-charge ratio (m/z) and intensity of the intact peptide precursor are first recorded by a full MS scan. Then, different precursors with specific m/z values are automatically selected for fragmentation by

MS/MS, usually in a data dependent acquisition (DDA) mode. However, some coeluting peptides might not be selected in the DDA mode as a result of undersampling. One study has demonstrated that multiple injections of the same enrichment fraction can significantly increase the number of unique phosphopeptide IDs [128, 136]. The resulting fragment spectrum can further be interpreted and matched to peptide sequences by a protein database search engine. However, loss of labile phosphate groups and poor backbone fragmentation has made the sequencing of phosphopeptides and the assignment of the phosphorylation sites very challenging. Among many MS/MS fragmentation strategies made to characterize phosphopetpides [119, 137-139], two major fragmentation methods have been used widely in the field of phosphoproteomics: neutral loss with collision-induced dissociation MS/MS/MS, and electron transfer dissociation. A more thorough discussion of MS aspects of phosphoproteomics can be found in reviews [119, 136-140].

7.2. Collision-induced dissociation with neutral loss scan

Collision-induced dissociation (CID) is the most widely used collision mode in commercial mass spectrometers for peptide sequencing [141-143]. Peptide cations are accelerated and selected to be bombarded by neutral gas molecules such as helium or argon, which causes the cleavage on peptide bonds. Peptide ions then dissociate into a series of b- and y-type ion fragments, forming the MS/MS spectra that can be further interpreted to obtain the peptide sequence [142, 144, 145]. An example of using CID MS/MS to characterize cardiac phosphopeptides was made by Ruse, et al [146]. However, the CID of phosphopeptides usually causes the loss of a phosphate group (H_3PO_4, −98 Da)

on Ser and Thr and has shown lower efficiency for breaking the peptide backbone [147, 148]. The phosphate group on Tyr is more stable than those on Ser and Thr, but partial neutral losses (HPO_3, -80Da) have also been reported [135, 149, 150]. Thus, most MS/MS spectra of phosphopeptides contain dominant neutral loss peaks and a relatively low abundance of y- and b- fragment peaks, which make the interpretation of backbone sequence and the phosphosite assignment difficult. Nonetheless, some strategies have been developed to take advantage of this distinctive fragmentation behavior of phosphopeptides, namely, by use of mass spectrometers capable of neutral loss scanning. In a quadrupole ion trap (IT) instrument, the neutral loss precursor ion originating from the loss of phosphate group during MS/MS can be selected for further fragmentation by MS^3 in order to provide additional sequence information [58]. Palmisano *et al*. identified six novel phosphorylation sites from bovine heart mitochondrial oxidative phosphorylation complex I using MS^3 fragmentation with neutral loss scanning mode for 98, 49 or 32.7 Da on a linear IT-Fourier transform (LIT-FT) mass spectrometer [27]. In another strategy implemented on quadrupole IT or linear IT-Orbitrap instruments, called multi-stage activation (MSA), a neutral loss ion can be simultaneously activated and form a composite MS/MS spectrum with significantly enhanced intensity of backbone fragments peaks [151]. However, both MS^3 and MSA strategies result in an increasing duty cycle and undersample more than the conventional MS/MS approach [136, 152, 153]. Also, the newer generation of mass spectrometers such as hybrid LTQ-Orbitrap is capable of generating enough backbone structure information on MS/MS-only mode to make the MS^3 a less attractive option for high-throughput phosphoproteomics experiments [153, 154].

7.3. Electron transfer dissociation

Recently, electron capture dissociation (ECD) [155, 156] and electron transfer dissociation (ETD) [157, 158] have received increasing attention as promising alternative fragmentation modes to characterize phosphopeptides. ECD, which typically requires an FT-ICR instrument, directly adds a low-energy free electron to the multiprotonated precursor ions. In ETD, which is widely compatible with other IT instruments, a radical gaseous anion such as fluoranthene is used to transfer an electron to the peptide cation. Unlike CID, ECD/ETD methods allow the preservation of labile phosphate groups because fragmentation occurs solely on the phosphopeptide backbone with more thorough and uniform fragmentation patterns in the form of c- and z- ion series [159, 160]. Also, as ETD fragmentation tends to perform better with larger peptides that have >+3 charges, CID and ETD offer coverage over different peptide populations [161]. Interestingly, reports have shown that phosphopeptides are more likely to have miscleavages during proteolytic digestion steps. Thus phosphopeptides tend to have longer sequences and carry more charges [162, 163], which gives ETD further advantages over CID fragmentation for phosphopeptide analysis [164]. Similarly, alternative proteases such as Lys-C [165, 166] or Lys-N [167] which generate longer peptide fragments, are a logical choice for phosphoprotein digestion when teamed with ETD type instruments. Nevertheless, the biased fragmentation preference against +2 charged species also comprises the ability of ETD to unveil the diversity of proteolytic peptide populations [168]. Two strategies have been applied to specifically address the low rate of interpretation of +2 charged peptides in the ETD approach. One way is to pair ETD and CID fragmentation. This strategy can achieved via analyzing the same sample

in separate instruments exemplified by Lu et al. [33], Deng et al. [24] and Zong et al. [35], or using an instrument capable of alternating between CID and ETD modes automatically [164, 169, 170] . It is also noteworthy that CID and ETD generate different series of fragment ions (y+b *vs.* z+c) which can later be combined to get improved coverage of backbone information. However, similar to the MS^3 approach, ETD-CID modes also have longer duty cycles, leading to undersampling issues when handling complex mixtures. Reports also have shown the add-on modifications of the mass spectrometer to perform the CID-ETD mixed mode can increase the detector noise level [171, 172]. The second approach to enhance ETD fragmentation efficiency of small doubly charged peptides was proposed by Swaney et al. [165] to use supplemental low energy collisional activation of +2 charged species (ETcaD). In this method, after electron transfer, non-dissociative ions can be further activated by collisional energy to form useful c- and z-type fragments. However, this might actually additionally cause neutral loss of the labile phosphate group [165]. Wu et al. also introduced another variation to ETcaD, the charge-reduced CID (CRCID), in which a charge-reduced electron transfer species is isolated and fragmented to generate cleaner and easier-to-interpret spectra [173].

8. Bioinformatics for Phosphoproteomics

8.1. General procedure

In general, raw spectra files generated by tandem mass spectrometry are centroided and deisotoped to generate experimental peak lists with m/z and intensities of ions and ion fragments in open formats such as mgf, pkl, dta, mzXML or mzData. When using MS^3 methods, MS^2 and MS^3 spectra from the same phosphopeptide can either be analyzed

separately or merged together for a single search [154, 174]. For ETD raw data, studies have shown that the post-acquisition removal of over-abundant peaks of precursors, charge-reduced precursors, and neutral losses from charge-reduced precursors can make the ETD MS/MS spectra interpretable to a greater extent by common search engines [175].

After peak list extraction, spectra files can be submitted to a search algorithm of choice to compare against amino acid sequences in protein databases for determination of the peptide sequence from observed mass spectra. Despite the fact that common search algorithms yield identifications of phosphopeptides, the localization of phosphosites from these algorithms generally have very low confidence level. Thus, the identification of phosphopeptides from mass spectra should be further augmented by additional scripts to locate phosphosites on the peptide sequences. Only unambiguous identification of phosphosites together with peptide sequences can be used for further motif analysis and kinase-substrate predication in order to generate biological meaningful data.

8.2. Peptide identification

Search algorithms using various scoring models have been developed to assess the likelihood of a mass spectrum match with a particular peptide sequence. Popular search algorithms include Mascot [176], OMSSA [177], Sequest [178], Spectrum Mill (Agilent), Phenyx (GeneBio), ProteinLynx (Waters), InsPecT [179], ProteinProspector [180], and X!Tandem [181]. Reviews of search algorithms can be found elsewhere [182, 183]. As details of each algorithm differ greatly from one another, each search engine has its own

scoring system. Identification selectivity and sensitivity varies among the different search engines as well and are performed under different conditions. As most search engines provides results that are complementary, consensus searching strategies using combinations of multiple search algorithms are used to improve the sensitivity and specificity of peptide identification from complex samples [184, 185]. A typical search process is carried out within constrains set by experimentally and instrumentally specified search parameters such as species, enzyme specificity, fixed and dynamic chemical modifications, fragmentation rules, and mass tolerance on both the MS and MS/MS level. For phosphoproteomics data, either dynamic modification of p-S/T/Y should be specified, or water loss on S/T/Y should be used in neutral loss CID data unless a specific chemical derivatization method was used for transformation phosphate groups into other forms. Since MSA- or ETD-based phosphoproteomic experiments are carried out in IT instruments, Sequest, OMSSA, X!Tandem or Protein Prospector should be used as search engine in order to better interpret low mass accuracy data [136, 186]. However, it is also reported that Mascot and Spectrum Mill perform better for ETD data [187]. One should keep in mind that the performance of search algorithms differ from case to case, and the best choice of database search strategies only come by trial and error. It is not possible to compare search results from different search algorithms because they report the score in different scales and also because conventional score thresholds are generally too stringent for phosphopeptide identification. Instead, the use of the estimated false discovery rate (FDR) as a universal indicator of the specificity of the search strategy has gained popularity in recent years. In this strategy, the search, regardless of which algorithm is used, is repeated using an identical parameter against a random or reversed

database (the decoy) that has the same size and complexity of the forward database in order to predict the number of false positive identifications made during the target database search [188, 189]. The decoy search option is available in most popular database search algorithms including Mascot, X!Tandem, Sequest, and OMSSA [190] and automatically estimates the possible error associated with the target search. Alternatively, only one search will be carried out against a forward-reverse concatenated database to estimate the FDR. *De novo* interpretation of MS/MS spectra has also shown potential applications for phosphoproteomics and may be the best solution for identification of phosphopeptides in the absence of a complete genome/proteome database. Commercial *de novo* sequencing packages like PEAKS® (Bioinformatics Solution Inc.) are now capable of processing both CID and ETD data types from a wide selection of tandem mass spectrometers [191, 192]. However, real application of *de novo* sequencing in phosphoproteomics is still lacking a robust estimation of the value of this approach.

For large-scale shotgun based phosphoproteomics studies, additional scripts like public PhosphoPIC [193] and DTASelect [194] can be used to remove non-phosphorylated peptide IDs and compile them according to predetermined false discovery rates.

8.3. Phosphosite determination

Due to the fact that many phosphopeptides may contain multiple phosphosites, common search engines may generate ambiguous reports about the exact location of phosphorylation during a database search. A number of bioinformatic tools have been developed to re-interpret the tandem MS spectra and to compute confidence scores

associated with the localization of each phosphorylation site. Beausoleil *et al*. published the Ascore algorithm, in which a probability score is calculated to confirm the presence of each possible phosphosite on identified peptides via assessment of the likelihood of finding phosphosite-determining ions within the mass spectrum [195]. This algorithm is available as a stand-alone script (http://ascore.med.harvard.edu) and also is implemented within other search packages such as Sequest Sorcerer (Sage-N Research, CA), Scaffold PTM (Proteome Software, OR) and ArMone (http://bioanalysis.dicp.ac.cn/proteomics/software/ArMone.html) [196]. Balley *et al*, developed the SloMo algorithm based on Ascore for localization of user specified protein modifications from on ETD/ECD data [197]. Ruttenberg *et al*. described another open source tool, PhosphoScore, using a scoring model that localizes phosphosites based on pdMS^n data processed in Sequest environment [198]. A similar package, MSQuant, (http://msquant.sourceforge.net) was developed for confidence assignment of phosphosites from pdMS^n spectra processed by Mascot [107, 199].

8.4. Protein phosphorylation database

Compared to conventional detection methods of protein phosphorylation, current studies using high-throughput shotgun MS-based surveying technologies can generate hundreds or even thousands of protein phosphorylation identifications. To make these data more accessible and easier for data mining and sharing, many web-based depositories now accept phosphoproteomics data directly from the public research community. Examples of such depositories includes Phosida (**www.phosida.com**) [200], Phospho.ELM

(**http://phospho.elm.eu.org**) [201], PhosphoSite (**www.phosphosite.org**) [202], and dbPTM (**http://dbptm.mbc.nctu.edu.tw**) [203]. These databases also incorporate the phosphorylation sites documented in universal protein databases such as UniProt (**www.uniprot.org**) as well as accepting other published phosphoproteomics data.

However, concerns have been raised over the quality of publically submitted datasets to these databases since the confidence of phosphosite identification in each study varies depending on the choice of instrument, search engine, identification criteria, and, to a large extent, the judgment of each investigator in determining which criteria should be used.

8.5. Phosphosites motif analysis and kinase prediction

Although high-throughput phosphoproteomics provides a global view of the phosphorylation status of the proteome, such large-scale ‘dot’ surveys usually do not reflect the ‘dot-to-dot’ signaling context or functional significance of identified protein phosphorylations. It was estimated that in a eukaryote organism there are close to 500 kinases [204, 205] responsible for several thousands of phosphorylations across the entire proteome via the reorganization of specific local or global amino acid sequences. However, large portions of these kinase-substrate relationships still have not been characterized [206]. Hence, when processing datasets that contain large amounts of protein phosphorylation identifications, it is worthwhile to extract phosphorylation motifs overrepresented in the datasets and to predict kinases responsible for the observed phosphorylation pattern via *in silico* approaches. An example software that provide

sequence motif analysis of phosphorylated peptides is Motif-X (http://motif-x.med.harvard.edu/) [207]. Popular kinase prediction tools include Scansite (http://scansite.mit.edu/) [208], Predikin (http://predikin.biosci.uq.edu.au/), NetPhosK (http://www.cbs.dtu.dk/services/NetPhosK/) [209], NetworKin (http://networkin.info) [210], and KinasePhos (http://kinasephos.mbc.nctu.edu.tw/) [211, 212].

9. Quantitative Phosphoproteomics

9.1. General considerations

Quantitative phosphoproteomics usually involves the relative quantitation of protein phosphorylation variance, comparing samples from two or more pathological stages, external stimuli, or pharmacological treatments. Gel-based image analysis is still a popular procedure for surveying changes in the phosphoproteome for many molecular cardiology studies due to its simplicity [25, 26, 29, 43, 213, 214]. However, gel-based methods often solely quantify a small number of phosphoproteins, lack the ability to 'zoom-in' for investigation of the abundance of phosphorylations on specific phosphosites, and are not readily suited for multi-sample comparisons. Large-scale, LC-based phosphoproteomic studies have demonstrated the ability of MS to monitor the temporal dynamics of thousands of protein phosphorylations in response to stimulation. Although mass spectrometry provides an excellent tool for peptide structure characterization, quantitation using only ion abundance in MS is unreliable because ionization efficiency is highly variable from peptide to peptide. Thus, quantitative phosphoproteomic techniques usually refer to quantitation of relative changes in protein phosphorylation between samples. The more challenging issue is elucidation of the

stoichiometry of phosphorylation. Comparison of the abundance of phosphorylated and unphosphorylated peptide isoform in MS is generally not suitable, as they behave differently during LC separation and MS ionization. However, one can still normalize phosphoproteomics data using protein abundance inferred from other peptide fragments derived from the same phosphorylated protein to determine whether or not the change in phosphorylation abundance is due to protein synthesis, degradation or solely to phosphorylation. Several shotgun-based quantitation approaches have been implemented for phosphoproteomics, including popular stable-isotope labeling techniques and some label-free strategies as summarized in Fig 2.3.

9.2. Metabolic labeling

When and how a label is introduced into a given sample varies amongst different labeling techniques. It is generally advisable to incorporate isotopic tags and to pool samples as early as possible such that risks of systematic errors during further sample processing are minimized. Metabolic labeling, such as $^{14}N/^{15}N$ labeling [215], is advantageous, as it generally involves culturing cells in a medium containing stable isotopes such that synthesis of necessary amino acids can incorporate those isotope compounds via protein turnover. A protocol that utilizes stable isotope labeling with amino acids in a cell culture (SILAC) [216, 217] is now the most popular metabolic encoding method employed in phosphoproteomic studies [116, 218-222]. Briefly, cultured cells are propagated in medium containing isotope-encoded arginine (normal Arg, $^{13}C_6$-Arg or $^{13}C_6/^{15}N_4$-Arg) or lysine (normal Lys, $^{2}H_4$-Lys or $^{13}C_6/^{15}N_2$-Lys). Subsequent trypsin mediated digestion theoretically allows for each proteolytic peptide to carry only one SILAC encoded Lys or

Arg. As peptides with isotope variants are chemically identical and not separable chromatographically, they are assayed by MS simultaneously with distinctive mass increments in MS spectra. Relative peptide abundance from different samples can be obtained through comparison of the ion intensities of isotopic doublet or triplet peaks derived from corresponding SILAC tagging. Repeat measurements are made via calculation of the SILAC ratio from consecutive MS scans across the chromatographic peak of a given peptide. Although a maximum of three different samples can be compared in one SILAC experiment, it is still possible to use a common reference sample to bridge several SILAC experiments for further multiplexing. SILAC was initially restricted to cell culture studies, but *in vivo* labeling using a SILAC protocol for rat [223] and mice [224] models was reported. Notably, a variant SILAC protocol using a $^{13}C_9$-tyrosine tag was described by Cantin et al. for quantification of phosphotyrosine-containing peptides [225].

9.3. Chemical labeling

Post-extraction labeling of proteins or peptides allows for the use of much broader sample sources. Most chemical labeling techniques initially introduced to quantitative proteomics have found an application in phosphoproteomics. However, not all proteomic labeling techniques can be successfully transferred. For example, isotope-coded affinity tags or ICAT [226] only label cysteine residues and were developed for protein quantitation. However a particular phosphopeptide may not contain a cysteine group, thus be undetectable. The isobaric tag for relative and absolute quantitation (iTRAQ) technology [227] was also unavailable for quantitative phosphoproteomics due to the

incompatibility of the fragmentation mechanism for iTRAQ quantitation and the phosphorylation identification because iTRAQ was typically implemented on a Q-tof platform, while most IT instruments for phosphoproteomic applications have low-molecular cut-off issues. Fortunately, the new generation LTP-Orbitrap with higher-energy C-trap dissociation [23, 228] and linear IT instruments equipped with the pulsed-Q dissociation (PQD) technique [229] is now used for quantification of iTRAQ labeled peptides. Additionally, ETD instruments have been shown to be compatible with iTRAQ quantitation [230-232]. Although there is still no report on the application of ETD alone towards quantification of iTRAQ labeled phosphopeptides, several studies have taken advantage of CID/ETD hybrid MS/MS modes in which ETD was used to enhance phosphopeptide identification, while PQD or HCD provide improved quantification of iTRAQ-labeled phosphopeptides [233, 234].The iTRAQ reaction solely targets the N-terminus and lysine groups of tryptic peptides. Each isobaric tag contains an isotope-encoded reporter group as well as a counter balancing group in order to ensure all tags have the same molecular weight, such that peptides from different samples labeled with different iTRAQ tags are indistinguishable in MS spectra. Only upon fragmentation, will the reporter ions be released and form reporter isotopic peak clusters in the 113-121 m/z region. Relative peptide abundance may be obtained by comparison of the intensities of the reporter ion peaks within the cluster that has a 1Da mass shift. iTRAQ was developed to accommodate up to four samples and further multiplexing can be achieved by constructing a reference sample for use in several iTRAQ experiments. Recently, 8-plex iTRAQ [235] was introduced, capable of assaying eight samples for further simplification of experimental design. In particular, for time-dependent signaling dynamic studies or

projects with multiple replicates. However, one should be cautious when using iTRAQ for quantification of phosphopeptides, as it often relies on very few replicates of the MS/MS spectrum instead of taking advantage of the whole chromatographic profile as compared to SILAC or label-free quantitation. Data compressing issues associated with iTRAQ, possibly due to the precursor selection window of up to several m/z ratios, has made quantification less sensitive to small changes [236].

Recently, another isobaric amine-reactive labeling technique, the tandem mass tag or TMT [237] was successfully implemented for quantitative phosphoproteomic purposes [238]. O^{18} labeling [239], achieved by incorporation of ^{18}O into the peptides during protease digestion, can also be used for phosphoproteomic applications [239-241].

9.4. Label-free quantitation

To a lesser extent, label-free methodologies have been applied to the quantification of protein phosphorylation, specifically, when isotope tags are not compatible with the study subjects or the MS instrument. One common label-free approach utilizes the positive correlation of the analyte concentration and the MS signal intensity of ESI instruments. Briefly, all samples are first surveyed by LC-MS to document significant changes of chromatographic peak areas of unknown precursors. As each analyte investigated with MS bears a unique m/z ratio and retention time, an accurate mass and retention time (AMT) tag should be assigned to each component. Usually, visualization via two-dimensional images of ion intensities in the span of retention times and m/z ratios from LC-MS are used to assist in data mining of peptide species with significant changes.

Later an AMT list of precursors of interest is generated and fed back to the MS. An additional LC-MS/MS run with exactly the same chromatographic conditions is conducted for characterization of those compounds on the AMT “wanted” list [242]. A variant of this method for quantitative phosphoproteomics employs the direct comparison of selected ion intensities of compounds in LC-MS/MS runs, based on which the identification and quantification of phosphopeptides is achieved simultaneously [243]. Spectral counting is a popular yet less complicated strategy for use in label-free quantitative phosphoproteomics [244]. It makes use of the number of spectra matched to peptides as the surrogate semi-quantitative measurement of the peptide abundances. In the shotgun proteomics setting, several studies have demonstrated good agreement between protein quantification via spectral counting and quantification via MS intensities [245, 246]; however, such agreement has not been proven in the phosphoproteomic setting. Regardless of the strategy used, label-free quantification does not allow a sample pooling step and all samples including biological replicates are necessarily separated by LC and analyzed by MS individually. Thus, label-free quantification data with potential large run-to-run variability must be normalized by overall chromatographic intensity or by spiking of a known amount of protein standard prior to analysis [243, 247-249].

10. Phosphoproteomics in Cardiac Molecular Research

10.1. Protein phosphorylation and cardiac diseases

Protein phosphorylation has been studied widely in the field of molecular cardiology. One aspect is to study the abnormalities of key signaling pathways in respect to their roles in the pathogenesis of cardiovascular diseases. For example, hyperactivation of

Ras/MEK/ERK1/2 has also been linked to the pathogenesis of cardiac hypertrophy and increased contraction [11, 250]. Disregulation of PKA signaling [251, 252] and GPCR-activated protein kinase signaling, in particular the CaMKII pathway [10, 253], has been shown to cause hypertrophic dysfunction and sudden death. A variety of signaling pathways are activated during ischemic-reperfusion (I/R) stress, mainly triggered by the release of endogenous reactive oxygen species (ROS). ROS stimulate Src tyrosine kinases [254, 255] and mitogen-activated protein kinases (MAPKs) such as JNKs, BMK1, ERK1/2, p38-MAPK during I/R insult [254, 256, 257]. Among them, some signaling proteins such as p38 have been described as being valuable therapeutic targets for preservation of cardiac function and inhibition of cytotoxic effects from cytokines released during myocardial reperfusion injury [258-260]. In addition, many cardiac diseases were found to be related to hormone-induced signaling triggered by noradrenalin, endothelin, and angiotensin all of which are mediated by cellular phosphorylation networks [261-264].

Conversely, it is also interesting to investigate signaling pathways that initiate cardioprotective action for their potential therapeutic value. For example, JAK-STAT pathway plays an essential role in the development of ischemic preconditioning to prevent myocardial infarction [265-268]. Protein kinase B/Akt signaling [269-271] and protein kinase C (PKC) signaling [272-280] has been proven to provide cardioprotection against ischemic insult. Recent studies have proposed a comprehensive view of reperfusion injury salvage kinase (RISK) pathway [281] which interconnect prosurvival (PI3K)-Akt and the p42 (Erk1/2) signaling mediating the cardioprotection provided by

I/R and pharmacological preconditioning as reviewed by Hausenloy et al. [282, 283].

Other than the signaling pathway, attention has been drawn to the aberrant regulation of phosphorylation of other key functional components in cardiomyocytes and their roles in cardiomyopathy development. Cardiac contractile machinery contains phosphorylated proteins such as troponin I [284, 285], and myosin light chain 2 [286], myosin-binding protein C [287, 288], with an altered phosphorylation pattern found in many pathophysiological conditions [289]. Functions of major ion channels in cardiomyocytes such as ryanodine-sensitive calcium (Ca^{2+}) channels [290, 291], voltage-gated L-type Ca^{2+} channels [292], and delayed rectifier potassium (K^{+}) channels [293] are also regulated via protein phosphorylation [294].

Given the breadth of protein phosphorylation in many critical cardiac functions as well as the implication of protein phosphorylation in cardiac disease development, kinase modulators have emerged as a main class of cardiovascular drugs [10, 295, 296]. However, despite what is already known, it is likely that only the tip of the iceberg has been found for both the complexity and the dynamics of the cardiac phosphoproteome. Global approaches making the use of systems biology tools have the potential to provide a vital perspective to resolve the complexity of global phosphorylation pattern changes. Perhaps more practically, phosphoproteomics can be applied to selection of putative targets with therapeutic values and to verify the true global impact of novel pharmacological intervention on the phosphoproteome.

10.2. Examples of phosphoproteomics applications in the field of molecular cardiology

In recent decades, the application of functional genomics approach like proteomics has been increasingly used to unveil the molecular events related to many cardiac research topics as showed in Fig. 2.4. Among these studies only few nevertheless targeted on the signaling transduction aspect of cardiac proteome. For example, early quest on prosurvival PKC signaling by Edmondson and others using a standard proteomics have successfully identified 93 proteins constituting the PKC signalosome in ischemic mouse heart [297]. However, the phosphorylation pattern of the PKC protein complex is still missing. Despite the current momentum in the application of phosphoproteomics on other fields, the number of phosphoproteomic studies in cardiac research is stagnantly low. Here we want to review some examples to demonstrate how phosphoproteomics can be helpful to dissect the molecular events underlying some critical cardiac physiological and pathological changes.

In the quest to investigate global impacts of I/R stress on the signaling pathway, Chou et al. [298] implemented a p-Tyr specific phosphoproteomic approach to study the response of H9C2 cardiomyocytes to oxidative stress. Using p-Tyr based immunoprecipitation approach and LC-MS/MS, the authors further identified 23 H_2O_2-induced tyrosine phosphorylated proteins including novel targets like platelet-derived growth factor receptor-β and γ-adducin. Interestingly, many identified phosphoproteins modulated by oxidative stress in this study confirmed the known role of Src kinase to mediate ROS signaling in cardiomyocytes [298]. Fernado et al. [299] employed a novel 2DE-based

kinase assay coupled with mass spectrometry for interrogation of kinase signaling in the biological context of cardiac hypertrophic adaptation dictated by the MKK6-p38 pathway. In this study, the authors first compared the kinase activity map from transgenic mice with congenital cardiac hypertrophy overexpressing MKK6 with that from the wild type. Using a 2DE gel cross-linked with a ubiquitous kinase substrate, they were able to identify new kinase candidates affected by MKK6 in the myopathic signaling cascade, including 5'-AMP activated kinase (AMPK) [300], Rho-associated kinase (RAK) [301], and protein kinase N (PKN) [302]. The authors also surveyed the direct substrates of MKK6 from the total cardiac protein lysate separated by 2DE. Some new candidate proteins previously reported to be associated with cardiac hypotrophy, including α-adducin [303] and semaphorin [304], were identified by MS as novel downstream substrates of MKK6 in this study. Calcineurin signaling pathway has been documented as a positive regulator of cardiac hypertrophy [305]. In an attempt to understand the role of calsarcin-1 a negative regulator in calcineurin mediated cardiac hypertrophy, Paulsson et al. launched a small scale proteomics study to characterize the phosphorylation status of calsarcin-1 during cardiac injury. Using ETD-CID hybrid MS, the authors determined multiple novel phosphorylation sites on calsarcin during hypertrophy which provides more detailed insight into molecular mechanism triggers cardiac hypertrophy [306].

Global phosphoproteomics has also been used to resolve the complexity of hormone-regulated signaling networks. As a member of seven-transmembrane receptors (7TMRs), the angiotensin II type 1 receptor (AT_1R) was documented to exhibit functional selectivity upon binding to different ligands [307]. Using quantitative phosphoproteomics

based on the SILAC metabolic labeling technique, Christensen et al. [116] compared $G\alpha_q$-dependent AT_1R signaling and $G\alpha_q$-independent signaling in cardiomyocytes using full AT_1R agonist angiotensin II (Ang II) and biased agonist SII angiotensin II (SII Ang II). Totally over 10,000 phosphorylation sites were identified using a SCX-TiO_2 shotgun phosphopeptides enrichment protocol combined with a high-resolution LTQ-Orbitrap MS operated under MSA mode. Among this large amount of newly discovered phosphosites, 1183 sites from 500 phosphoproteins were regulated by Ang II or its analogue SII Ang II. As expected, 36% of those sites were only modulated by SII Ang II which proved the hypothesis that the $G\alpha_q$ protein-dependent and -independent pathways activated distinct groups of kinases. Moreover, this conclusion was confirmed by further analysis of phosphorylated sequence motifs. Motifs flank the $G\alpha_q$ protein-dependent and -independent phosphorylated sites shown different types of consensus sequences and hence indicate the involvement of different kinases groups. Moreover, phosphoproteomics data led to the discovery of protein kinase D as a critical intermediator for both $G\alpha_q$ protein-dependent and -independent signaling and was suggested to take an important role on the development of cardiac hypertrophy [116].

Over the last few years, there have been trends to dissect the phosphoproteome within the subcellular compartment from cardiomyocytes, such as mitochondria and 20S proteasome and myofilament. Given the importance of mitochondria in critical cardiac functions such as energy metabolism, apoptosis regulation, and Ca^{2+} homeostasis, great attention has been drawn to the portrayal of signaling dynamics in mitochondria. Three gel-based studies have been carried out to define the extent of the mitochondrial

phosphoproteome [22, 26, 308], in which phosphorylations on several proteins of from respiratory chain complexes and enzymes involved in intermediary metabolism were identified. Quantitative phosphoproteomics using 2DE with ProQ staining allowed Hopper et al. to identify the dynamic changes in phosphorylation patterns in mitochondrial matrix proteins following Ca^{2+} stimulation, most notably the dephosphorylation of PDH, manganese superoxide dismutase, and F_0F_1-ATPase [26]. In another study, Boja et al. isolated mitochondria from a porcine heart and enriched the phosphopeptides by a novel SCX/HILIC-TiO_2 procedure. Using iTRAQ labeling techniques combined with LTQ-Orbitrap ETD-HCD hybrid instrument, the author quantified phosphorylation pattern changes in proteins such as pyruvate dehydrogenase (PDH), branched chain α keto-acid dehydrogenase (BCKDH), ATP synthase and mitofilin upon Ca^{2+} stimulus, de-energization, and treatment with dichloroacetate, a drug candidate for the treatment of genetic mitochondrial diseases by as a PDH complex inhibitor [23]. In another study aiming to reveals regulatory pathways of cardiac mitochondria, Deng et al. identified 210 novel phosphorylation sites on a CID-ETD hybrid MS platform from multiple murine cardiac mitochondrial proteins including components of the electron transport chain (ETC) complexes. Furthermore, the author suggested that calcium overload can compromise ETC activities via phosphorylation modulation, and illustrated that this type of mitochondria injury can be restored by enhanced phosphorylation of ETC using phosphatase inhibitors [309]. Cardiac myofilament contains multiple contractile related proteins that are heavily regulated by phosphorylation. Yin et al. performed a proteomics study to map the protein phosphorylation on myofilemental subproteome from murine cardiomyocytes stimulation with endothelin-1 and

isoproterenol both as inducers of myofilament phosphorylation. Using LTQ-Orbitrap with ETD capability, 5 phosphorylation sites were identified from 3 contractile proteins. Moreover, over 600 additional proteins were identified in the myofilament subproteome including many previously unrecognized kinases and phosphatases indicating a complex network of myofilament regulation via protein phosphorylation [310]. Functional difference between neonatal and adult cardiomyocytes reflect by differential phosphorylation on sarcomeric proteins was known but with limited knowledge on how protein phosphorylation changes during cardiac development. Using a 2DE with Pro-Q Diamond stain protocol, Yuan et al analyzed sarcomeric phosphoproteome difference between neonatal and adult rat hearts [214]. Phosphoprotein from neonatal and adult heart sample were further labeled by O^{16}/O^{18} and enriched by TiO_2 prior to LTQ-FT MS analysis from which several phosphosites on myofilamental proteins like myosin-binding protein and tropomyosin were found to be differentially phosphorylated at two different stages [214]. There are also some applications of phosphoproteomics on cardiac 20S proteasome to explore the molecular basis involved with the regulation of myocardial proteolytic system. Zong et al. [34] employed immunoprecipitation with anti-p-Ser/Thr/Tyr antibodies coupled with Q-TOF MS/MS to survey the phosphorylation pattern of the cardiac 20S proteasome complex. Based on phosphoproteomics data, 2 novel signaling partners associated with cardiac 20S complexes were found: protein phosphatase 2A, and protein kinase A. The authors also found that the peptidase activity of 20S proteasome can be enhanced by site-specific phosphorylation. The role of PKA in governing 20S phosphoproteome was further investigated by Lu et al. [311]. In this study, the authors first identified 20 phosphosites from the murine heart 20S proteasome using

IEF/SDS-PAGE-TiO_2 protocol combined with CID and ETD MS analysis, from which 13 phosphosites were identified. Moreover, the author recorded changes in the phosphorylation pattern and peptidase activity of 20S proteasome upon PKA activation [214].

High-throughput phosphoproteomics has also emerged as a novel screening technique for potential diagnostic biomarkers with clinical implications. Using a fluorescent phosphostain-based 2DE workflow, Dubois et al. [312] successfully detected 69 differentially phosphorylated proteins in sample from the left ventricles of rats as the result of myocardial infarction. Within these altered phosphoproteins, they further demonstrated that the phosphorylation of serine207 on troponin T (TnT) was significantly decreased both in rat LV and plasma samples with myocardial infarction and left ventricle remodeling (LVR). The authors further confirmed that the serine207-phosphorylated TnT/total TnT ratio was also significantly low in human plasma samples from patients with intermediate or high LVR suggesting that phosphorylated troponin T in circulation can be used as a potential biomarker for LVR [312].

11. Challenges

Phosphoproteomics is still facing numerous technical challenges, particularly with samples of high complexity. Low coverage of the phosphoproteome is probably the biggest issue inherent to the MS-based phosphoproteomics technique itself. First, there is no enrichment strategy that can guarantee complete extraction of phosphopeptide species from complex samples with low sample loss and high selectivity against interfering

compounds like acidic peptides. Secondly, since MS is usually used in a DDA mode, some peptide species, especially those in low abundance, are likely to be 'ignored' due to undersampling issues. Third, identification of phosphopeptides from a tandem MS spectrum is still a difficult task, as not all search engines can effectively interpret spectra from phosphoproteomics-oriented instruments, such as MSA or MS^3-CID or ETD. There is still a general dearth of strategies to localize phosphosites, even given successful peptide backbone identification. Last but not least, the field of phosphoproteomics lacks universal identification criteria for reliable data quality and general guidance of how phosphoproteomics data should be processed, stored, and published. This was evidenced by dramatic variations in the number of phosphopeptide IDs identified in different phosphoproteomics studies with similar procedures on similar subjects.

Considerable challenges in the application of phosphoproteomics still remain to be addressed, particularly the data-mining and interpretation of massive amounts of high-throughput information. Current advances in enrichment and MS methodologies for large-scale phosphoproteomic applications can help to reveal large numbers of phosphorylation sites from complex whole cell or tissue extract samples. However, the functional significance of the vast majority of identified phosphosites still remains to be determined and associated with a dynamic signaling network. Despite the emergence of protein databases with functional annotations and new generations of *in silico* pathway analysis packages, most information is still obtained via non-experimental means. In a typical quantitative analysis of phosphoproteome with differing patho-physiological state or pharmaceutical perturbations, the biological causal linkage between biological input

and the observed changes in phosphorylation patterns remains difficult to characterize. Moreover, systematic approaches still lack the power to confirm the relevance of pathways to disease pathology or other downstream phenotypic endpoints. The key question that remains to be answered is whether the changes in those pathways identified from a phosphoproteomic study is the contributing factor or a side-effect of the diseases. For this usually non-global methodology are complementary to proteomic approaches.

Another significant concern in the phosphoproteomics field is sizeable data variability. One inherent disadvantage of MS-based phosphoproteomics is the accuracy of identification and quantitation. Compared to standard proteomics for analysis of protein expression levels, the identification and quantification of protein phosphorylation often relies on individual peptide species with very low numbers of MS spectrum replicates. Also, the number of samples that can be compared is limited, even with multiplexing labeling techniques. Lack of biological or technical replicates tends to introduce systematic noise and to yield measurements with large variability. This restriction also signifies that the researcher must choose the conditions to be compared very carefully, as there is limited room for multiple conditions and time points in the experimental design, which eventually restricts the power of phosphoproteomics for assessment of the dynamics of signaling networks.

However, as an ever-evolving research field, new enrichment methods, new quantitation strategies, new instrumentation, and new bioinformatics tools have been rapidly introduced to the analysis of the phosphoproteome yielding much higher throughput,

accuracy, coverage, and sensitivity for quantitation of changes in protein phosphorylation on a large scale, as well as higher capabilities for illustration of kinase-substrate interactions. Also, more efforts are needed to make phosphoproteomic techniques less complex, amenable, and cost-effective for researchers with limited experience. These advances would be particularly useful in clinical settings to assist diagnosis, prognosis, and treatment of cardiovascular diseases.

12. Concluding Remarks

Distortion in protein phosphorylation patterns can lead to the onset and progression of many cardiac diseases. Technically, phosphoproteomics provides time-efficient, less tedious and high-throughput identification of protein phosphorylation; conceptually, it permits the evaluation of the complexity of protein phosphorylation in a non-biased global perspective with little *a priori* knowledge. This review provides a nonspecialist working in the field of molecular cardiology an introduction to up-to-date phosphoproteomics tools, workflows, and examples of phosphoproteomics applications in a range of cardiac research topics. We foresee that phosphoproteomics will become a vital tool contributing to the mechanistic understanding of the signaling and regulatory aspects of cellular functions under normal or cardiopathological status, as well as identification of pivotal kinases or protein phosphorylation sites, thereby allowing the discovery of novel therapeutic targets.

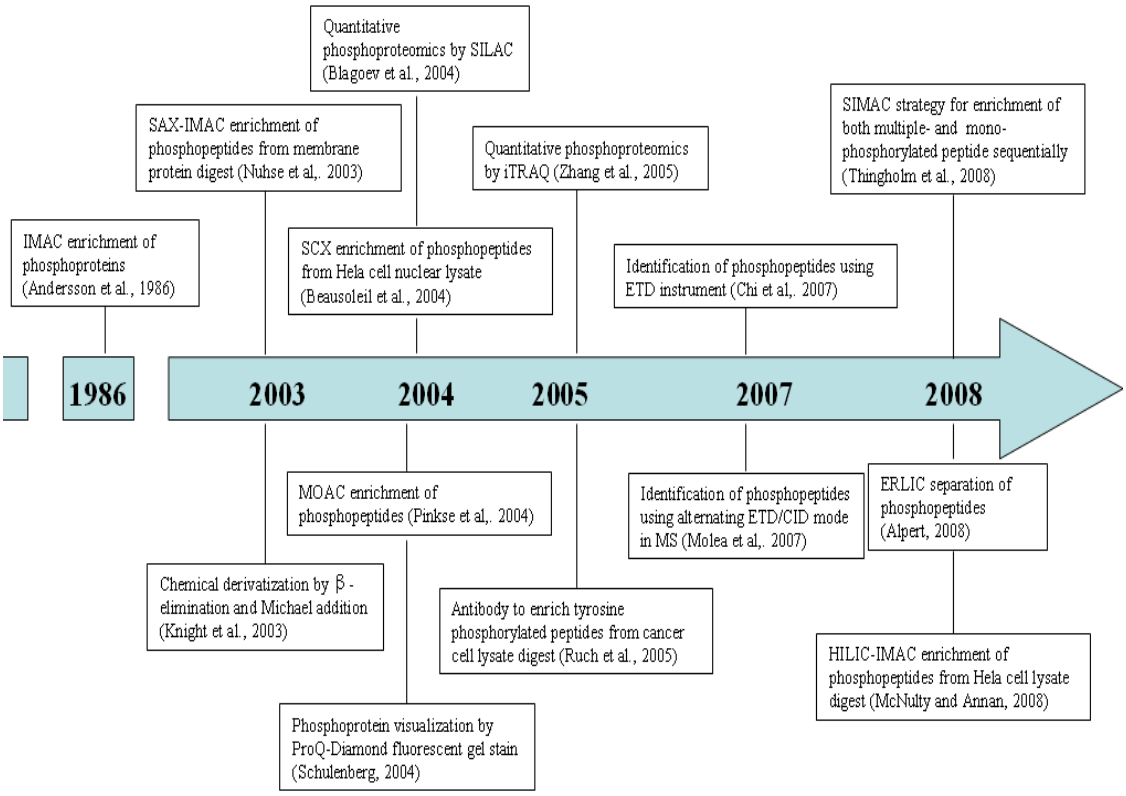


Figure 2.1. The past decade has witnessed the rapid development of large-scale phosphoproteome methodology using enrichment techniques and high accuracy mass spectrometry. Here only key technical milestones in development of mass spectrometry based phosphoproteomics methodology are highlighted.

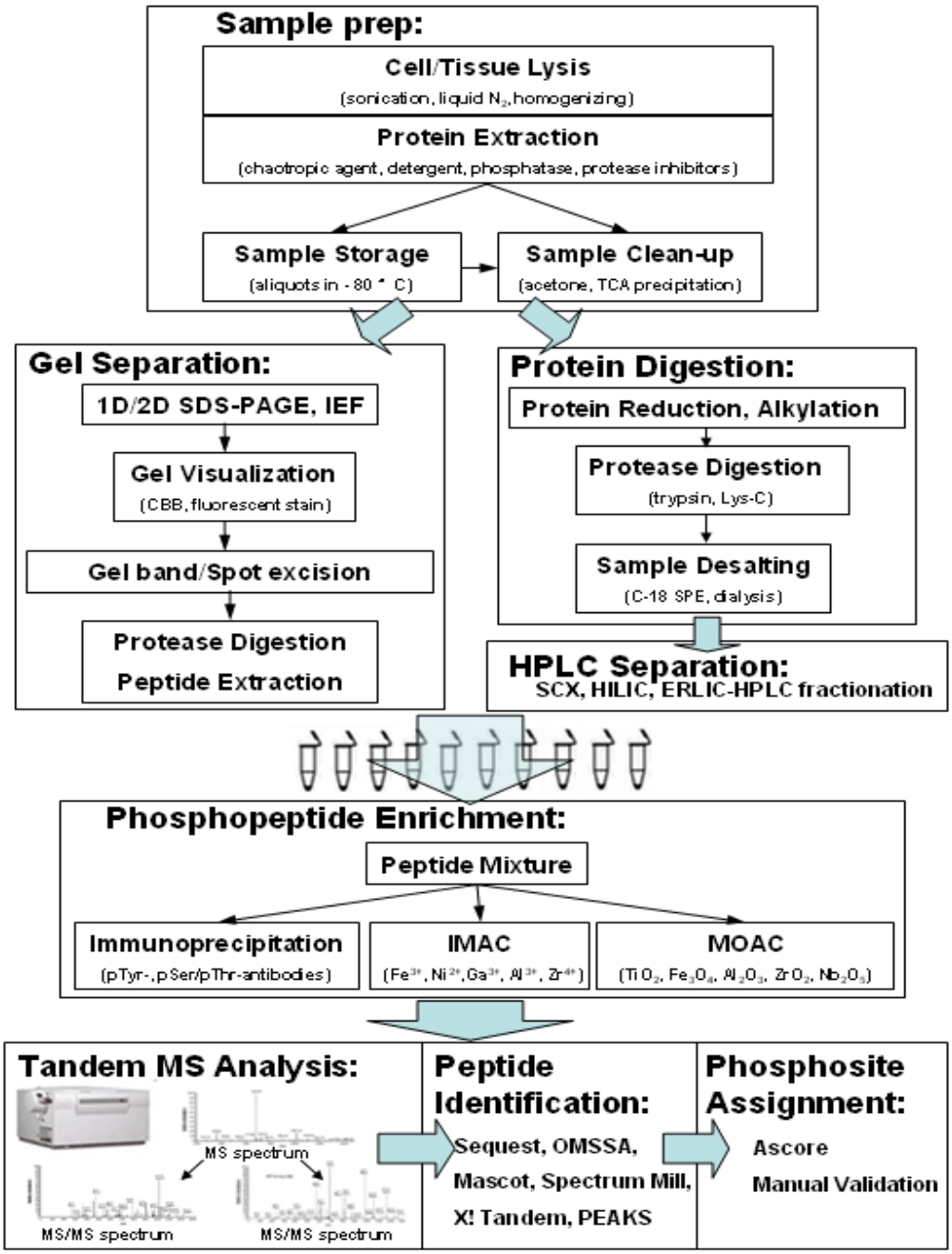


Figure 2.2. An overview of typical large-scale phosphoproteomics workflow. Most current phosphoproteomics procedures consist of at least two-dimensional orthogonal separation of proteins (2DE protocol) before proteolytic digestion or peptides (2DLC protocol) after proteolytic digestion. Typically a step of enrichment procedure is used to isolate phosphopeptides from complex mixture. Simplified fractions containing phosphopeptides are then introduced into mass spectrometry to determine the peptide identity and the presence of phosphor group.

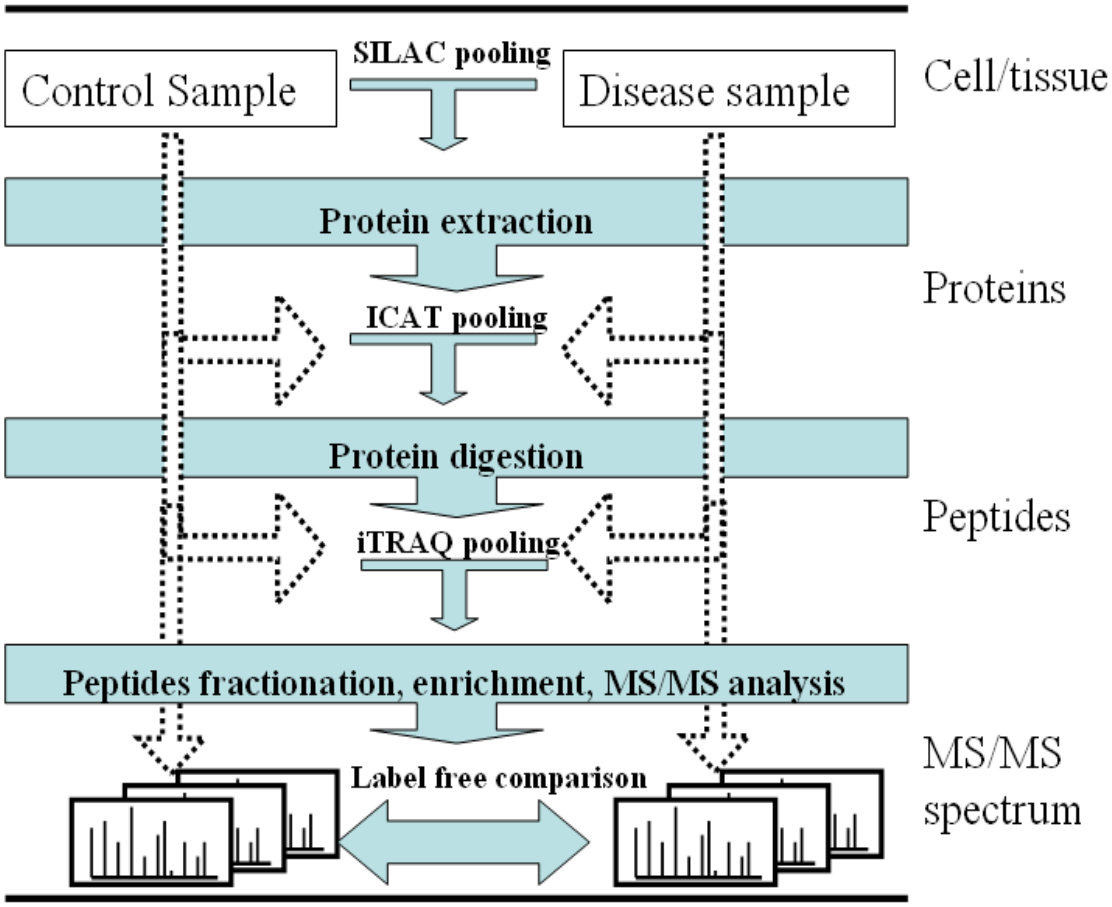


Figure 2.3 Comparison of isotopic labeling and label-free quantitative shotgun workflows. Different labeling strategies pool samples for relative quantitative at different sample processing steps. After pooling, peptides from different samples will be analyzed together within the same MS or MS/MS spectrum. Label-free workflow introduces no sample pooling, comparing MS signal or number of MS/MS spectra from separate LC-MS runs.

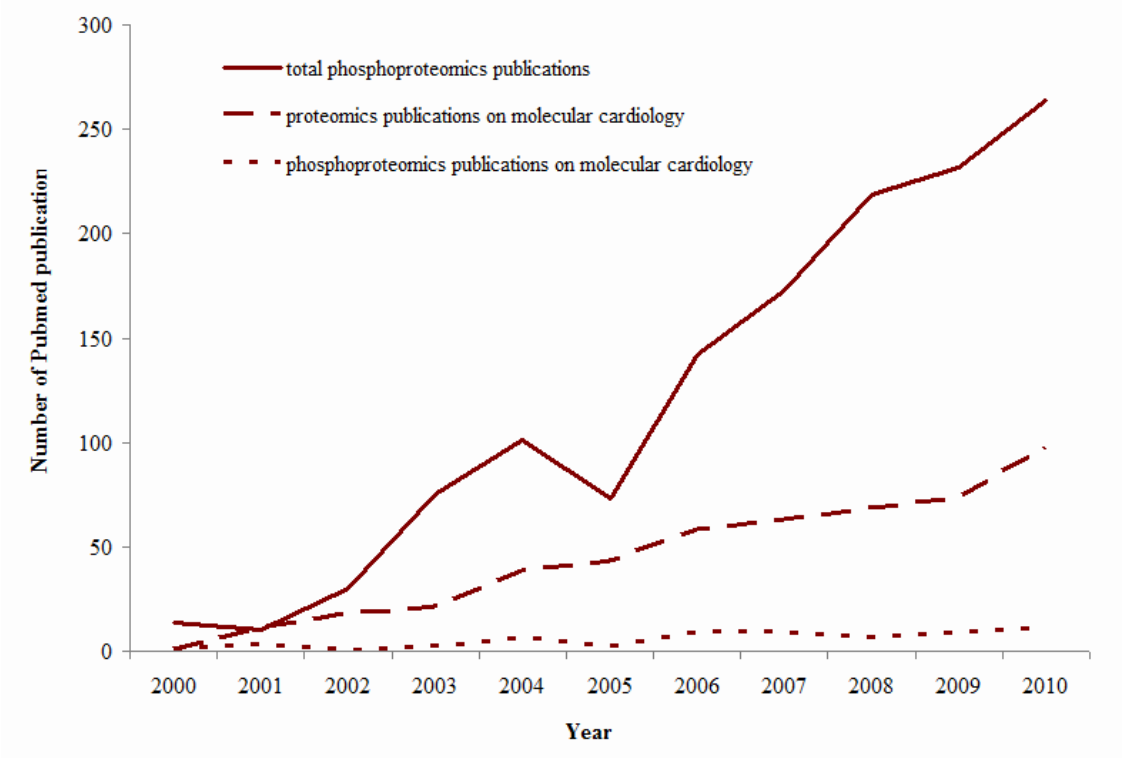


Figure 2.4. Compare to the general increasing trend in phosphoproteomics application found in PubMed on other research topics in the recent decade, there are still very few number of publications using phosphoproteomics for cardiac research purpose. However, the application of standard proteomics in the area of cardiac research shown a steady increasing trend.

REFERENCES

[1] H. Zhang, X.M. Zha, Y. Tan, P.V. Hornbeck, A.J. Mastrangelo, D.R. Alessi *et al.*, Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs, *Journal of Biological Chemistry* **277** (42) (2002), pp. 39379-39387.

[2] D.E. Kalume, H. Molinaand A. Pandey, Tackling the phosphoproteome: tools and strategies, *Current Opinion in Chemical Biology* **7** (1) (2003), pp. 64-69.

[3] B.T. Seet, I. Dikic, M.M. Zhouand T. Pawson, Reading protein modifications with interaction domains, *Nat Rev Mol Cell Biol* **7** (7) (2006), pp. 473-483.

[4] H.R. Matthews, Protein-Kinases and Phosphatases That Act on Histidine, Lysine, or Arginine Residues in Eukaryotic Proteins - a Possible Regulator of the Mitogen-Activated Protein-Kinase Cascade, *Pharmacology & Therapeutics* **67** (3) (1995), pp. 323-350.

[5] J. Puttick, E.N. Bakerand L.T.J. Delbaere, Histidine phosphorylation in biological systems, *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1784** (1) (2008), pp. 100-105.

[6] S. Klumppand J. Krieglstein, Reversible phosphorylation of histidine residues in vertebrate proteins, *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1754** (1-2) (2005), pp. 291-295.

[7] T. Hunter, Signaling - 2000 and beyond, *Cell* **100** (2000), pp. 113-127.

[8] P. Cohen, The regulation of protein function by multisite phosphorylation - a 25 year update, *Trends Biochem Sci* **25** (2000), pp. 596-601.

[9] P. Cohen, Signal integration at the level of protein kinases, protein phophatasesand their substrates, *Trends Biochem Sci* **17** (1992), pp. 408-413.

[10] M.E. Anderson, L.S. Higginsand H. Schulman, Disease mechanisms and emerging therapies: protein kinases and their inhibitors in myocardial disease, *Nature Clinical Practice Cardiovascular Medicine* **3** (8) (2006), pp. 437-445.

[11] K. Lorenz, J.P. Schmitt, M. Vidaland M.J. Lohse, Cardiac hypertrophy: targeting Raf/MEK/ERK1/2-signaling, *Int J Biochem Cell Biol* **41** (12) (2009), pp. 2351-2355.

[12] J.D. Hoffertand M.A. Knepper, Taking aim at shotgun phosphoproteomics, *Analytical Biochemistry* **375** (1) (2008), pp. 1-10.

[13] A. Bodzon-Kulakowska, A. Bierczynska-Krzysik, T. Dylag, A. Drabik, P. Suder, M. Noga *et al.*, Methods for samples preparation in proteomic research, *J Chromatogr B Analyt Technol Biomed Life Sci* **849** (1-2) (2007), pp. 1-31.

[14] J.v. Hagen. Proteomics sample preparation: Wiley-VCH; 2008.

[15] M. Trost, G. Bridon, M. Desjardinsand P. Thibault, Subcellular phosphoproteomics, *Mass Spectrom Rev* **29** (6) pp. 962-990.

[16] L.A. Huber, K. Pfallerand I. Vietor, Organelle proteomics - Implications for subcellular fractionation in proteomics, *Circulation Research* **92** (9) (2003), pp. 962-968.

[17] C. Pasquali, I. Fialkaand L.A. Huber, Subcellular fractionation, electromigration analysis and mapping of organelles, *J Chromatogr B Biomed Sci Appl* **722** (1-2) (1999), pp. 89-102.

[18] J.R. Yates Iii, A. Gilchrist, K.E. Howelland J.J.M. Bergeron, Proteomics of organelles and large cellular structures, *Nat Rev Mol Cell Biol* **6** (9) (2005), pp. 702-714.

[19] D.E. Warnock, E. Fahyand S.W. Taylor, Identification of protein associations in organelles, using mass spectrometry-based proteomics, *Mass Spectrometry Reviews* **23** (4) (2004), pp. 259-280.

[20] M. Dreger, Subcellular proteomics, *Mass Spectrom Rev* **22** (1) (2003), pp. 27-56.

[21] P. Ping, Identification of novel signaling complexes by functional proteomics, *Circ Res* **93** (7) (2003), pp. 595-603.

[22] A.M. Aponte, D. Phillips, R.K. Hopper, D.T. Johnson, R.A. Harris, K. Blinova *et al.*, Use of P-32 To Study Dynamics of the Mitochondrial Phosphoproteome, *Journal of Proteome Research* **8** (6) (2009), pp. 2679-2695.

[23] E.S. Boja, D. Phillips, S.A. French, R.A. Harrisand R.S. Balaban, Quantitative Mitochondrial Phosphoproteomics Using iTRAQ on an LTQ-Orbitrap with High Energy Collision Dissociation, *Journal of Proteome Research* **8** (10) (2009), pp. 4665-4675.

[24] N. Deng, J. Zhang, C. Zong, Y. Wang, H. Lu, P. Yang *et al.*, Phosphoproteome analysis reveals regulatory sites in major pathways of cardiac mitochondria, *Mol Cell Proteomics* pp.

[25] J. Feng, M. Zhu, M.C. Schaub, P. Gehrig, B. Roschitzki, E. Lucchinetti *et al.*, Phosphoproteome analysis of isoflurane-protected heart mitochondria: phosphorylation of adenine nucleotide translocator-1 on Tyr194 regulates mitochondrial function, *Cardiovasc Res* **80** (1) (2008), pp. 20-29.

[26] R.K. Hopper, S. Carroll, A.M. Aponte, D.T. Johnson, S. French, R.F. Shen *et al.*, Mitochondrial matrix phosphoproteome: Effect of extra mitochondrial calcium, *Biochemistry* **45** (8) (2006), pp. 2524-2536.

[27] G. Palmisano, A.M. Sardanelli, A. Signorile, S. Papaand M.R. Larsen, The phosphorylation pattern of bovine heart complex I subunits, *Proteomics* **7** (10) (2007), pp. 1575-1583.

[28] G. Pocsfalvi, Selective Enrichment in Phosphopeptides for the Identification of

Phosphorylated Mitochondrial Proteins, *Methods in Enzymology, Vol 457: Mitochondrial Function, Partb Mitochondrial Protein Kinases, Protein Phosphatases and Mitochondrial Diseases* **457** (2009), pp. 81-96.

[29] H. Schwertz, J.M. Carter, M. Abdudureheman, M. Russ, U. Buerke, A. Schlitt *et al.*, Myocardial ischemia/reperfusion causes VDAC phosphorylation which is reduced by cardioprotection with a p38 MAP kinase inhibitor, *Proteomics* **7** (24) (2007), pp. 4579-4588.

[30] D.D. Newmeyerand S. Ferguson-Miller, Mitochondria: releasing power for life and unleashing the machineries of death, *Cell* **112** (4) (2003), pp. 481-490.

[31] M.A. Frohman, Mitochondria as integrators of signal transduction and energy production in cardiac physiology and disease, *Journal of Molecular Medicine-Jmm* **88** (10) (2010), pp. 967-970.

[32] C. Horbinskiand C.T. Chu, Kinase signaling cascades in the mitochondrion: a matter of life or death, *Free Radic Biol Med* **38** (1) (2005), pp. 2-11.

[33] H.J. Lu, C.G. Zong, Y.J. Wang, G.W. Young, N. Deng, P. Souda *et al.*, Revealing the Dynamics of the 20 S Proteasome Phosphoproteome A COMBINED CID AND ELECTRON TRANSFER DISSOCIATION APPROACH, *Molecular & Cellular Proteomics* **7** (11) (2008), pp. 2073-2089.

[34] C. Zong, A.V. Gomes, O. Drews, X. Li, G.W. Young, B. Berhane *et al.*, Regulation of murine cardiac 20S proteasomes: role of associating partners, *Circ Res* **99** (4) (2006), pp. 372-380.

[35] C. Zong, G.W. Young, Y. Wang, H. Lu, N. Deng, O. Drews *et al.*, Two-dimensional electrophoresis-based characterization of post-translational modifications of mammalian

20S proteasome complexes, *Proteomics* **8** (23-24) (2008), pp. 5025-5037.

[36] P.H. O'Farrell, High resolution two-dimensional electrophoresis of proteins, *J Biol Chem* **250** (10) (1975), pp. 4007-4021.

[37] T.H. Steinberg, Protein gel staining methods: an introduction and overview, *Methods Enzymol* **463** (2009), pp. 541-563.

[38] V. Santoni, M. Molloyand T. Rabilloud, Membrane proteins and proteomics: un amour impossible?, *Electrophoresis* **21** (6) (2000), pp. 1054-1070.

[39] T.E. Thingholm, O.N. Jensenand M.R. Larsen, Analytical strategies for phosphoproteomics, *Proteomics* **9** (6) (2009), pp. 1451-1468.

[40] J. Reindersand A. Sickmann, State-of-the-art in phosphoproteomics, *Proteomics* **5** (16) (2005), pp. 4052-4061.

[41] S. Morandell, T. Stasyk, K. Grosstessner-Hain, E. Roitinger, K. Mechtler, G.K. Bonn *et al.*, Phosphoproteomics strategies for the functional analysis of signal transduction, *Proteomics* **6** (14) (2006), pp. 4047-4056.

[42] A.K. Bendt, A. Burkovski, S. Schaffer, M. Bott, M. Farwickand T. Hermann, Towards a phosphoproteome map of Corynebacterium glutamicum, *Proteomics* **3** (8) (2003), pp. 1637-1646.

[43] G.X. Chu, G.F. Egnaczyk, W. Zhao, S.H. Jo, G.C. Fan, J.E. Maggio *et al.*, Phosphoproteome analysis of cardiomyocytes subjected to beta-adrenergic stimulation - Identification and characterization of a cardiac heat shock protein p20, *Circulation Research* **94** (2) (2004), pp. 184-193.

[44] T.H. Steinberg, B.J. Agnew, K.R. Gee, W.-Y. Leung, T. Goodman, B. Schulenberg *et al.*, Global quantitative phosphoprotein analysis using Multiplexed Proteomics

technology, *Proteomics* **3** (7) (2003), pp. 1128-1144.

[45] B. Schulenberg, T.N. Goodman, R. Aggeler, R.A. Capaldiand W.F. Patton, Characterization of dynamic and steady-state protein phosphorylation using a fluorescent phosphoprotein gel stain and mass spectrometry, *Electrophoresis* **25** (15) (2004), pp. 2526-2532.

[46] E. Kinoshita, E. Kinoshita-Kikuta, K. Takiyamaand T. Koike, Phosphate-binding tag, a new tool to visualize phosphorylated proteins, *Mol Cell Proteomics* **5** (4) (2006), pp. 749-757.

[47] H. Kaufmann, J.E. Baileyand M. Fussenegger, Use of antibodies for detection of phosphorylated proteins separated by two-dimensional gel electrophoresis, *Proteomics* **1** (2) (2001), pp. 194-199.

[48] J.W. Mandell, Phosphorylation state-specific antibodies: applications in investigative and diagnostic pathology, *Am J Pathol* **163** (5) (2003), pp. 1687-1698.

[49] G.J. Opiteck, K.C. Lewis, J.W. Jorgensonand R.J. Anderegg, Comprehensive on-line LC/LC/MS of proteins, *Anal Chem* **69** (8) (1997), pp. 1518-1524.

[50] C.C. Wuand M.J. MacCoss, Shotgun proteomics: tools for the analysis of complex biological systems, *Curr Opin Mol Ther* **4** (3) (2002), pp. 242-250.

[51] M.P. Washburn, D. Woltersand J.R. Yates, 3rd, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat Biotechnol* **19** (3) (2001), pp. 242-247.

[52] A. Motoyamaand J.R. Yates, 3rd, Multidimensional LC separations in shotgun proteomics, *Anal Chem* **80** (19) (2008), pp. 7187-7193.

[53] M.L. Fournier, J.M. Gilmore, S.A. Martin-Brownand M.P. Washburn,

Multidimensional separations-based shotgun proteomics, *Chem Rev* **107** (8) (2007), pp. 3654-3686.

[54] J.M. Davisand J.C. Giddings, Statistical-Theory of Component Overlap in Multicomponent Chromatograms, *Analytical Chemistry* **55** (3) (1983), pp. 418-424.

[55] W.G. Chenand F.M. White, Proteomic analysis of cellular signaling, *Expert Rev Proteomics* **1** (3) (2004), pp. 343-354.

[56] S.A. Beausoleil, M. Jedrychowski, D. Schwartz, J.E. Elias, J. Villén, J. Li *et al.*, Large-scale characterization of HeLa cell nuclear phosphoproteins, *Proceedings of the National Academy of Sciences of the United States of America* **101** (33) (2004), pp. 12130-12135.

[57] K.B. Limand D.B. Kassel, Phosphopeptides enrichment using on-line two-dimensional strong cation exchange followed by reversed-phase liquid chromatography/mass spectrometry, *Anal Biochem* **354** (2) (2006), pp. 213-219.

[58] S.A. Beausoleil, M. Jedrychowski, D. Schwartz, J.E. Elias, J. Villen, J. Li *et al.*, Large-scale characterization of HeLa cell nuclear phosphoproteins, *Proc Natl Acad Sci U S A* **101** (33) (2004), pp. 12130-12135.

[59] M.W.H. Pinkse, S. Mohammed, J.W. Gouw, B. van Breukelen, H.R. Vosand A.J.R. Heck, Highly Robust, Automated, and Sensitive Online TiO 2-Based Phosphoproteomics Applied To Study Endogenous Phosphorylation in Drosophila melanogaster, *Journal of Proteome Research* **7** (2) (2008), pp. 687-697.

[60] J.A. Wu, P. Warren, Q. Shakey, E. Sousa, A. Hill, T.E. Ryan *et al.*, Integrating titania enrichment, iTRAQ labeling, and Orbitrap CID-HCD for global identification and quantitative analysis of phosphopeptides, *Proteomics* **10** (11) (2010), pp. 2224-2234.

[61] J. Villén, S.A. Beausoleil, S.A. Gerberand S.P. Gygi, Large-scale phosphorylation analysis of mouse liver, *Proceedings of the National Academy of Sciences* **104** (5) (2007), pp. 1488-1493.

[62] A. Gruhler, J.V. Olsen, S. Mohammed, P. Mortensen, N.J. Færgeman, M. Mann *et al.*, Quantitative Phosphoproteomics Applied to the Yeast Pheromone Signaling Pathway, *Molecular & Cellular Proteomics* **4** (3) (2005), pp. 310-327.

[63] J. Villenand S.P. Gygi, The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry, *Nat Protoc* **3** (10) (2008), pp. 1630-1638.

[64] M. Zarei, A. Sprenger, F. Metzger, C. Gretzmeierand J. Dengjel, Comparison of ERLIC-TiO(2), HILIC-TiO(2), and SCX-TiO(2) for Global Phosphoproteomics Approaches, *J Proteome Res* pp.

[65] G.H. Han, M.L. Ye, H.J. Zhou, X.N. Jiang, S. Feng, X.G. Jiang *et al.*, Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography, *Proteomics* **8** (7) (2008), pp. 1346-1361.

[66] J. Dai, L.S. Wang, Y.B. Wu, Q.H. Sheng, J.R. Wu, C.H. Shieh *et al.*, Fully automatic separation and identification of phosphopeptides by continuous pH-gradient anion exchange online coupled with reversed-phase liquid chromatography mass spectrometry, *J Proteome Res* **8** (1) (2009), pp. 133-141.

[67] T.S. Nühse, A. Stensballe, O.N. Jensenand S.C. Peck, Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry, *Molecular & Cellular Proteomics: MCP* **2** (11) (2003), pp. 1234-1243.

[68] S. Nie, J. Dai, Z.B. Ning, X.J. Cao, Q.H. Shengand R. Zeng, Comprehensive Profiling of Phosphopeptides Based on Anion Exchange Followed by Flow-Through Enrichment with Titanium Dioxide (AFET), *Journal of Proteome Research* **9** (9) (2010), pp. 4585-4594.

[69] J. Dai, W.H. Jin, Q.H. Sheng, C.H. Shieh, J.R. Wuand R. Zeng, Protein phosphorylation and expression profiling by Yin-yang multidimensional liquid chromatography (Yin-yang MDLC) mass spectrometry, *J Proteome Res* **6** (1) (2007), pp. 250-262.

[70] A. Motoyama, T. Xu, C.I. Ruse, J.A. Wohlschlegeland J.R. Yates, 3rd, Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides, *Anal Chem* **79** (10) (2007), pp. 3623-3634.

[71] A.J. Alpert, Hydrophilic-Interaction Chromatography for the Separation of Peptides, Nucleic-Acids and Other Polar Compounds, *Journal of Chromatography* **499** (1990), pp. 177-196.

[72] D.E. McNultyand R.S. Annan, Hydrophilic Interaction Chromatography Reduces the Complexity of the Phosphoproteome and Improves Global Phosphopeptide Isolation and Detection, *Molecular & Cellular Proteomics* **7** (5) (2008), pp. 971-980.

[73] C.-J. Wu, Y.-W. Chen, J.-H. Taiand S.-H. Chen, Quantitative Phosphoproteomics Studies Using Stable Isotope Dimethyl Labeling Coupled with IMAC-HILIC-nanoLC/MS/MS for Estrogen-Induced Transcriptional Regulation, *Journal of Proteome Research* pp.

[74] R.A. Saleemand J.D. Aitchison, Quantitative phosphoproteomics in fatty acid stimulated Saccharomyces cerevisiae, *Journal of Visualized Experiments: JoVE* (32)

(2009), pp.

[75] D.E. McNultyand R.S. Annan, Hydrophilic interaction chromatography for fractionation and enrichment of the phosphoproteome, *Methods Mol Biol* **527** (2009), pp. 93-105, x.

[76] D. Singer, J. Kuhlmann, M. Muschketand R. Hoffmann, Separation of multiphosphorylated peptide isomers by hydrophilic interaction chromatography on an aminopropyl phase, *Anal Chem* **82** (15) pp. 6409-6414.

[77] R.A. Saleem, R.S. Rogers, A.V. Ratushny, D.J. Dilworth, P.T. Shannon, D. Shteynberg *et al.*, Integrated phosphoproteomics analysis of a signaling network governing nutrient response and peroxisome induction, *Mol Cell Proteomics* **9** (9) pp. 2076-2088.

[78] S.D. Garbis, T.I. Roumeliotis, S.I. Tyritzis, K.M. Zorpas, K. Pavlakisand C.A. Constantinides, A Novel Multidimensional Protein Identification Technology Approach Combining Protein Size Exclusion Prefractionation, Peptide Zwitterion-Ion Hydrophilic Interaction Chromatography, and Nano-Ultraperformance RP Chromatography/nESI-MS(2) for the in-Depth Analysis of the Serum Proteome and Phosphoproteome: Application to Clinical Sera Derived from Humans with Benign Prostate Hyperplasia, *Anal Chem* pp.

[79] C.P. Albuquerque, M.B. Smolka, S.H. Payne, V. Bafna, J. Engand H. Zhou, A multidimensional chromatography technology for in-depth phosphoproteome analysis, *Mol Cell Proteomics* **7** (7) (2008), pp. 1389-1396.

[80] M. Gilar, P. Olivova, A.E. Dalyand J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, *Anal Chem* **77** (19) (2005), pp. 6426-6434.

[81] H. Zhang, T. Guo, X. Li, A. Datta, J.E. Park, J. Yang *et al.*, Simultaneous characterization of glyco- and phosphoproteomes of mouse brain membrane proteome with electrostatic repulsion hydrophilic interaction chromatography, *Molecular & Cellular Proteomics: MCP* **9** (4) pp. 635-647.

[82] C.S. Gan, T. Guo, H. Zhang, S.K. Limand S.K. Sze, A comparative study of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus SCX-IMAC-based methods for phosphopeptide isolation/enrichment, *Journal of Proteome Research* **7** (11) (2008), pp. 4869-4877.

[83] A.J. Alpert, Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides, *Analytical Chemistry* **80** (1) (2008), pp. 62-76.

[84] J.D. Dunn, G.E. Reidand M.L. Bruening, Techniques for phosphopeptide enrichment prior to analysis by mass spectrometry, *Mass Spectrom Rev* **29** (1) pp. 29-54.

[85] G.R. Yan, C.L. Xiao, G.W. He, X.F. Yin, N.P. Chen, Y. Cao *et al.*, Global phosphoproteomic effects of natural tyrosine kinase inhibitor, genistein, on signaling pathways, *Proteomics* **10** (5) pp. 976-986.

[86] H. Steen, B. Kuster, M. Fernandez, A. Pandeyand M. Mann, Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway, *J Biol Chem* **277** (2) (2002), pp. 1031-1039.

[87] J. Rush, A. Moritz, K.A. Lee, A. Guo, V.L. Goss, E.J. Spek *et al.*, Immunoaffinity profiling of tyrosine phosphorylation in cancer cells, *Nat Biotechnol* **23** (1) (2005), pp. 94-101.

[88] J. Villen, S.A. Beausoleil, S.A. Gerberand S.P. Gygi, Large-scale phosphorylation

analysis of mouse liver, *Proc Natl Acad Sci U S A* **104** (5) (2007), pp. 1488-1493.

[89] M.O. Collins, L. Yuand J.S. Choudhary, Analysis of protein phosphorylation on a proteome-scale, *Proteomics* **7** (16) (2007), pp. 2751-2768.

[90] G. Muszynska, G. Dobrowolska, A. Medin, P. Ekmanand J.O. Porath, Model Studies on Iron(Iii) Ion Affinity-Chromatography .2. Interaction of Immobilized Iron(Iii) Ions with Phosphorylated Amino-Acids, Peptides and Proteins, *Journal of Chromatography* **604** (1) (1992), pp. 19-28.

[91] L. Anderssonand J. Porath, Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography, *Anal Biochem* **154** (1) (1986), pp. 250-254.

[92] M.C. Posewitzand P. Tempst, Immobilized gallium(III) affinity chromatography of phosphopeptides, *Anal Chem* **71** (14) (1999), pp. 2883-2892.

[93] M.A. Coletti-Previeroand A. Previero, Alumina-phosphate complexes for immobilization of biomolecules, *Anal Biochem* **180** (1) (1989), pp. 1-10.

[94] M. Feng S Fau - Ye, H. Ye M Fau - Zhou, X. Zhou H Fau - Jiang, X. Jiang X Fau - Jiang, H. Jiang X Fau - Zou, B. Zou H Fau - Gong *et al.*, Immobilized zirconium ion affinity chromatography for specific enrichment of phosphopeptides in phosphoproteome analysis, (1535-9476 (Print)) pp.

[95] Z. Yu, G. Han, S. Sun, X. Jiang, R. Chen, F. Wang *et al.*, Preparation of monodisperse immobilized Ti(4+) affinity chromatography microspheres for specific enrichment of phosphopeptides, *Anal Chim Acta* **636** (1) (2009), pp. 34-41.

[96] S.B. Ficarro, M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz *et al.*, Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae, *Nat Biotechnol* **20** (3) (2002), pp. 301-305.

[97] Stewart, II, T. Thomsonand D. Figeys, 18O labeling: a tool for proteomics, *Rapid Commun Mass Spectrom* **15** (24) (2001), pp. 2456-2465.

[98] M.R. Jensen Ss Fau - Larsenand M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques, (0951-4198 (Print)) pp.

[99] M. Nousiainen, H.H. Sillje, G. Sauer, E.A. Niggand R. Korner, Phosphoproteome analysis of the human mitotic spindle, *Proc Natl Acad Sci U S A* **103** (14) (2006), pp. 5391-5396.

[100] Y.M. Ndassa, C. Orsi, J.A. Marto, S. Chenand M.M. Ross, Improved immobilized metal affinity chromatography for large-scale phosphoproteomics applications, *J Proteome Res* **5** (10) (2006), pp. 2789-2799.

[101] T.E. Thingholm, O.N. Jensen, P.J. Robinsonand M.R. Larsen, SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides, *Mol Cell Proteomics* **7** (4) (2008), pp. 661-671.

[102] A. Sanoand H. Nakamura, Titania as a chemo-affinity support for the column-switching HPLC analysis of phosphopeptides: application to the characterization of phosphorylation sites in proteins by combination with protease digestion and electrospray ionization mass spectrometry, *Anal Sci* **20** (5) (2004), pp. 861-864.

[103] A. Sanoand H. Nakamura, Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides, *Anal Sci* **20** (3) (2004), pp. 565-566.

[104] U.K. Aryaland A.R. Ross, Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and

mass spectrometry, *Rapid Commun Mass Spectrom* **24** (2) pp. 219-231.

[105] M.B. Gates, K.B. Tomerand L.J. Deterding, Comparison of metal and metal oxide media for phosphopeptide enrichment prior to mass spectrometric analyses, *J Am Soc Mass Spectrom* **21** (10) pp. 1649-1659.

[106] A. Leitner, M. Sturm, O. Hudecz, M. Mazanek, J.H. Smatt, M. Linden *et al.*, Probing the phosphoproteome of HeLa cells using nanocast metal oxide microspheres for phosphopeptide enrichment, *Anal Chem* **82** (7) pp. 2726-2733.

[107] J.V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen *et al.*, Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, *Cell* **127** (3) (2006), pp. 635-648.

[108] B. Macek, I. Mijakovic, J.V. Olsen, F. Gnad, C. Kumar, P.R. Jensen *et al.*, The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis, *Mol Cell Proteomics* **6** (4) (2007), pp. 697-707.

[109] M.R. Larsen, T.E. Thingholm, O.N. Jensen, P. Roepstorffand T.J. Jorgensen, Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns, *Mol Cell Proteomics* **4** (7) (2005), pp. 873-886.

[110] F. Wolschin, S. Wienkoopand W. Weckwerth, Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC), *Proteomics* **5** (17) (2005), pp. 4389-4397.

[111] M.W. Pinkse, P.M. Uitto, M.J. Hilhorst, B. Oomsand A.J. Heck, Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns, *Anal Chem* **76** (14) (2004), pp. 3935-3943.

[112] Y. Ikeguchiand H. Nakamura, Determination of organic phosphates by column-

switching high performance anion-exchange chromatography using on-line preconcentration on titania, *Analytical Sciences* **13** (3) (1997), pp. 479-483.

[113] H.J. Zhou, R.J. Tian, M.L. Ye, S.Y. Xu, S. Feng, C.S. Pan *et al.*, Highly specific enrichment of phosphopeptides by zirconium dioxide nanoparticles for phosphoproteome analysis, *Electrophoresis* **28** (13) (2007), pp. 2201-2215.

[114] E.W. Shin, J.S. Han, M. Jang, S.H. Min, J.K. Parkand R.M. Rowell, Phosphate adsorption on aluminum-impregnated mesoporous silicates: Surface structure and behavior of adsorbents, *Environmental Science & Technology* **38** (3) (2004), pp. 912-917.

[115] S.B. Ficarro, J.R. Parikh, N.C. Blankand J.A. Marto, Niobium(V) oxide (Nb2O5): Application to phosphoproteomics, *Analytical Chemistry* **80** (12) (2008), pp. 4606-4613.

[116] G.L. Christensen, C.D. Kelstrup, C. Lyngso, U. Sarwar, R. Bogebo, S.P. Sheikh *et al.*, Quantitative Phosphoproteomics Dissection of Seven-transmembrane Receptor Signaling Using Full and Biased Agonists, *Molecular & Cellular Proteomics* **9** (7) pp. 1540-1553.

[117] N. Sugiyama, T. Masuda, K. Shinoda, A. Nakamura, M. Tomitaand Y. Ishihama, Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications, *Mol Cell Proteomics* **6** (6) (2007), pp. 1103-1109.

[118] S.S. Jensenand M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques, *Rapid Commun Mass Spectrom* **21** (22) (2007), pp. 3635-3645.

[119] E. Salih, Phosphoproteomics by mass spectrometry and classical protein chemistry approaches, *Mass Spectrom Rev* **24** (6) (2005), pp. 828-846.

[120] Y. Oda, T. Nagasuand B.T. Chait, Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome, *Nat Biotech* **19** (4) (2001), pp. 379-382.

[121] D.T. McLachlinand B.T. Chait, Improved β-Elimination-Based Affinity Purification Strategy for Enrichment of Phosphopeptides, *Analytical Chemistry* **75** (24) (2003), pp. 6826-6836.

[122] Z.A. Knight, B. Schilling, R.H. Row, D.M. Kenski, B.W. Gibsonand K.M. Shokat, Phosphospecific proteolysis for mapping sites of protein phosphorylation, *Nat Biotech* **21** (9) (2003), pp. 1047-1054.

[123] G.J.L. Bernardes, J.M. Chalker, J.C. Erreyand B.G. Davis, Facile Conversion of Cysteine and Alkyl Cysteines to Dehydroalanine on Protein Surfaces: Versatile and Switchable Access to Functionalized Proteins, *Journal of the American Chemical Society* **130** (15) (2008), pp. 5052-5053.

[124] B. Herbert, F. Hopwood, D. Oxley, J. McCarthy, M. Laver, J. Grinyer *et al.*, β-elimination: An unexpected artefact in proteome analysis, *Proteomics* **3** (6) (2003), pp. 826-831.

[125] A.M. Taylor, O. Holstand J. Thomas-Oates, Mass spectrometric profiling ofO-linked glycans released directly from glycoproteins in gels using in-gel reductive β-elimination, *Proteomics* **6** (10) (2006), pp. 2936-2946.

[126] H. Zhou, J.D. Wattsand R. Aebersold, A systematic approach to the analysis of protein phosphorylation, *Nat Biotech* **19** (4) (2001), pp. 375-378.

[127] W.A. Tao, B. Wollscheid, R. O'Brien, J.K. Eng, X.-j. Li, B. Bodenmiller *et al.*, Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry, *Nat Meth* **2** (8) (2005), pp. 591-598.

[128] B. Bodenmiller, L.N. Mueller, M. Mueller, B. Domonand R. Aebersold, Reproducible isolation of distinct, overlapping segments of the phosphoproteome, *Nature Methods* **4** (3) (2007), pp. 231-237.

[129] H. Yin, K. Killeen, R. Brennen, D. Sobek, M. Werlichand T. van de Goor, Microfluidic Chip for Peptide Analysis with an Integrated HPLC Column, Sample Enrichment Column, and Nanoelectrospray Tip, *Analytical Chemistry* **77** (2) (2005), pp. 527-533.

[130] H. Yinand K. Killeen, The fundamental aspects and applications of Agilent HPLC-Chip, *Journal of Separation Science* **30** (10) (2007), pp. 1427-1434.

[131] M.-H. Fortier, E. Bonneil, P. Goodleyand P. Thibault, Integrated Microfluidic Device for Mass Spectrometry-Based Proteomics and Its Application to Biomarker Discovery Programs, *Analytical Chemistry* **77** (6) (2005), pp. 1631-1640.

[132] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wongand C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, *Science* **246** (4926) (1989), pp. 64-71.

[133] M. Karas, D. Bachmann, U. Bahrand F. Hillenkamp, Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *International Journal of Mass Spectrometry and Ion Processes* **78** (1987), pp. 53-68.

[134] V. Panchagnula, A. Mikulskis, L. Song, Y. Wang, M. Wang, T. Knubovets *et al.*, Phosphopeptide analysis by directly coupling two-dimensional planar electrochromatography/thin-layer chromatography with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *J Chromatogr A* **1155** (1) (2007), pp. 112-123.

[135] R.S. Annanand S.A. Carr, Phosphopeptide analysis by matrix-assisted laser desorption time-of-flight mass spectrometry, *Anal Chem* **68** (19) (1996), pp. 3413-3421.

[136] C.E. Bakalarski, W. Haas, N.E. Dephoureand S.P. Gygi, The effects of mass accuracy, data acquisition speed, and search algorithm choice on peptide identification rates in phosphoproteomics, *Anal Bioanal Chem* **389** (5) (2007), pp. 1409-1419.

[137] P.A. Grimsrud, D.L. Swaney, C.D. Wenger, N.A. Beaucheneand J.J. Coon, Phosphoproteomics for the masses, *ACS Chem Biol* **5** (1) pp. 105-119.

[138] A. Nita-Lazar, H. Saito-Benzand F.M. White, Quantitative phosphoproteomics by mass spectrometry: past, present, and future, *Proteomics* **8** (21) (2008), pp. 4433-4443.

[139] P.J. Boersema, S. Mohammedand A.J. Heck, Phosphopeptide fragmentation and analysis by mass spectrometry, *J Mass Spectrom* **44** (6) (2009), pp. 861-878.

[140] K. Blackburnand M.B. Goshe, Challenges and strategies for targeted phosphorylation site identification and quantification using mass spectrometry analysis, *Brief Funct Genomic Proteomic* **8** (2) (2009), pp. 90-103.

[141] L. Slenoand D.A. Volmer, Ion activation methods for tandem mass spectrometry, *J Mass Spectrom* **39** (10) (2004), pp. 1091-1112.

[142] J.M. Wellsand S.A. McLuckey, Collision-induced dissociation (CID) of peptides and proteins, *Methods Enzymol* **402** (2005), pp. 148-185.

[143] D.F. Hunt, J.R. Yates, 3rd, J. Shabanowitz, S. Winstonand C.R. Hauer, Protein sequencing by tandem mass spectrometry, *Proc Natl Acad Sci U S A* **83** (17) (1986), pp. 6233-6237.

[144] K.F. Medzihradszky, Peptide sequence analysis, *Methods Enzymol* **402** (2005), pp. 209-244.

[145] H. Steenand M. Mann, The ABC's (and XYZ's) of peptide sequencing, *Nat Rev Mol Cell Biol* **5** (9) (2004), pp. 699-711.

[146] C.I. Ruse, F.L. Tan, M. Kinterand M. Bond, Intregrated analysis of the human cardiac transcriptome, proteome and phosphoproteome, *Proteomics* **4** (5) (2004), pp. 1505-1516.

[147] A.M. Palumbo, J.J. Tepeand G.E. Reid, Mechanistic insights into the multistage gas-phase fragmentation behavior of phosphoserine- and phosphothreonine-containing peptides, *Journal of Proteome Research* **7** (2) (2008), pp. 771-779.

[148] J.P. DeGnoreand J. Qin, Fragmentation of phosphopeptides in an ion trap mass spectrometer, *J Am Soc Mass Spectrom* **9** (11) (1998), pp. 1175-1188.

[149] R.S. Annan, M.J. Huddleston, R. Verma, R.J. Deshaiesand S.A. Carr, A multidimensional electrospray MS-based approach to phosphopeptide mapping, *Anal Chem* **73** (3) (2001), pp. 393-404.

[150] M. Salek, A. Alonso, R. Pipkornand W.D. Lehmann, Analysis of protein tyrosine phosphorylation by nanoelectrospray ionization high-resolution tandem mass spectrometry and tyrosine-targeted product ion scanning, *Anal Chem* **75** (11) (2003), pp. 2724-2729.

[151] M.J. Schroeder, J. Shabanowitz, J.C. Schwartz, D.F. Huntand J.J. Coon, A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry, *Anal Chem* **76** (13) (2004), pp. 3590-3598.

[152] J. Villen, S.A. Beausoleiland S.P. Gygi, Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis, *Proteomics* **8** (21) (2008), pp. 4444-4452.

[153] X. Li, S.A. Gerber, A.D. Rudner, S.A. Beausoleil, W. Haas, J. Villen *et al.*, Large-scale phosphorylation analysis of alpha-factor-arrested Saccharomyces cerevisiae, *Journal of Proteome Research* **6** (3) (2007), pp. 1190-1197.

[154] P.J. Ulintz, A.K. Yocum, B. Bodenmiller, R. Aebersold, P.C. Andrewsand A.I. Nesvizhskii, Comparison of MS2-Only, MSA, and MS2/MS3 Methodologies for Phosphopeptide Identification, *Journal of Proteome Research* **8** (2) (2009), pp. 887-899.

[155] R.A. Zubarev, D.M. Horn, E.K. Fridriksson, N.L. Kelleher, N.A. Kruger, M.A. Lewis *et al.*, Electron capture dissociation for structural characterization of multiply charged protein cations, *Anal Chem* **72** (3) (2000), pp. 563-573.

[156] F.W. McLafferty, D.M. Horn, K. Breuker, Y. Ge, M.A. Lewis, B. Cerda *et al.*, Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance, *J Am Soc Mass Spectrom* **12** (3) (2001), pp. 245-249.

[157] J.E. Syka, J.J. Coon, M.J. Schroeder, J. Shabanowitzand D.F. Hunt, Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proc Natl Acad Sci U S A* **101** (26) (2004), pp. 9528-9533.

[158] L.M. Mikesh, B. Ueberheide, A. Chi, J.J. Coon, J.E. Syka, J. Shabanowitz *et al.*, The utility of ETD mass spectrometry in proteomic analysis, *Biochim Biophys Acta* **1764** (12) (2006), pp. 1811-1822.

[159] A.J. Kleinnijenhuis, F. Kjeldsen, B. Kallipolitis, K.F. Haselmannand O.N. Jensen, Analysis of histidine phosphorylation using tandem MS and ion - Electron reactions, *Analytical Chemistry* **79** (19) (2007), pp. 7450-7456.

[160] A. Stensballe, O.N. Jensen, J.V. Olsen, K.F. Haselmannand R.A. Zubarev, Electron capture dissociation of singly and multiply phosphorylated peptides, *Rapid Commun*

Mass Spectrom **14** (19) (2000), pp. 1793-1800.

[161] S.J. Pitteri, P.A. Chrisman, J.M. Hoganand S.A. McLuckey, Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: reactions of doubly and triply protonated peptides with SO2*, *Anal Chem* **77** (6) (2005), pp. 1831-1839.

[162] A. Schlosser, R. Pipkorn, D. Bossemeyerand W.D. Lehmann, Analysis of protein phosphorylation by a combination of elastase digestion and neutral loss tandem mass spectrometry, *Anal Chem* **73** (2) (2001), pp. 170-176.

[163] M. Benoreparsons, N.G. Seidahand L.P. Wennogle, Substrate Phosphorylation Can Inhibit Proteolysis by Trypsin-Like Enzymes, *Archives of Biochemistry and Biophysics* **272** (2) (1989), pp. 274-280.

[164] H. Molina, D.M. Horn, N. Tang, S. Mathivananand A. Pandey, Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry, *Proceedings of the National Academy of Sciences of the United States of America* **104** (7) (2007), pp. 2199-2204.

[165] D.L. Swaney, G.C. McAlister, M. Wirtala, J.C. Schwartz, J.E.P. Sykaand J.J. Coon, Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors, *Analytical Chemistry* **79** (2) (2007), pp. 477-485.

[166] A. Chi, C. Huttenhower, L.Y. Geer, J.J. Coon, J.E.P. Syka, D.L. Bai *et al.*, Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry, *Proceedings of the National Academy of Sciences of the United States of America* **104** (7) (2007), pp. 2193-2198.

[167] N. Taouatas, M.M. Drugan, A.J.R. Heckand S. Mohammed, Straightforward ladder sequencing of peptides using a Lys-N metalloendopeptidase, *Nature Methods* **5** (5)

(2008), pp. 405-407.

[168] D.M. Good, M. Wirtala, G.C. McAlisterand J.J. Coon, Performance characteristics of electron transfer dissociation mass spectrometry, *Mol Cell Proteomics* **6** (11) (2007), pp. 1942-1951.

[169] D.L. Swaney, G.C. McAlisterand J.J. Coon, Decision tree-driven tandem mass spectrometry for shotgun proteomics, *Nat Methods* **5** (11) (2008), pp. 959-964.

[170] S.M.M. Sweet, C.M. Bailey, D.L. Cunningham, J.K. Heathand H.J. Cooper, Large Scale Localization of Protein Phosphorylation by Use of Electron Capture Dissociation Mass Spectrometry, *Molecular & Cellular Proteomics* **8** (5) (2009), pp. 904-912.

[171] S.J. Pitteri, P.A. Chrismanand S.A. McLuckey, Electron-transfer ion/ion reactions of doubly protonated peptides: effect of elevated bath gas temperature, *Anal Chem* **77** (17) (2005), pp. 5662-5669.

[172] J. Wiesner, T. Premslerand A. Sickmann, Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications, *Proteomics* **8** (21) (2008), pp. 4466-4483.

[173] S.L. Wu, A.F. Huhmer, Z. Haoand B.L. Karger, On-line LC-MS approach combining collision-induced dissociation (CID), electron-transfer dissociation (ETD), and CID of an isolated charge-reduced species for the trace-level characterization of proteins with post-translational modifications, *J Proteome Res* **6** (11) (2007), pp. 4230-4244.

[174] P.J. Ulintz, B. Bodenmiller, P.C. Andrews, R. Aebersoldand A.I. Nesvizhskii, Investigating MS2/MS3 matching statistics: a model for coupling consecutive stage mass spectrometry data for increased peptide identification confidence, *Mol Cell Proteomics* **7**

(1) (2008), pp. 71-87.

[175] D.M. Good, C.D. Wenger, G.C. McAlister, D.L. Bai, D.F. Huntand J.J. Coon, Post-acquisition ETD spectral processing for increased peptide identifications, *J Am Soc Mass Spectrom* **20** (8) (2009), pp. 1435-1440.

[176] D.J. Pappin, P. Hojrupand A.J. Bleasby, Rapid identification of proteins by peptide-mass fingerprinting, *Curr Biol* **3** (6) (1993), pp. 327-332.

[177] L.Y. Geer, S.P. Markey, J.A. Kowalak, L. Wagner, M. Xu, D.M. Maynard *et al.*, Open mass spectrometry search algorithm, *J Proteome Res* **3** (5) (2004), pp. 958-964.

[178] J.K. Eng, A.L. Mccormackand J.R. Yates, An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database, *Journal of the American Society for Mass Spectrometry* **5** (11) (1994), pp. 976-989.

[179] S. Tanner, H. Shu, A. Frank, L.C. Wang, E. Zandi, M. Mumby *et al.*, InsPecT: identification of posttranslationally modified peptides from tandem mass spectra, *Anal Chem* **77** (14) (2005), pp. 4626-4639.

[180] R.J. Chalkley, K.C. Hansenand M.A. Baldwin, Bioinformatic methods to exploit mass spectrometric data for proteomic applications, *Methods Enzymol* **402** (2005), pp. 289-312.

[181] R. Craigand R.C. Beavis, TANDEM: matching proteins with tandem mass spectra, *Bioinformatics* **20** (9) (2004), pp. 1466-1467.

[182] R.G. Sadygov, D. Cociorvaand J.R. Yates, Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book, *Nat Meth* **1** (3) (2004), pp. 195-202.

[183] A.I. Nesvizhskii, O. Vitekand R. Aebersold, Analysis and validation of proteomic

data generated by tandem mass spectrometry, *Nat Methods* **4** (10) (2007), pp. 787-797.

[184] T. Sultana, R. Jordanand J. Lyons-Weiler, Optimization of the Use of Consensus Methods for the Detection and Putative Identification of Peptides via Mass Spectrometry Using Protein Standard Mixtures, *Journal of proteomics & bioinformatics* **2** (6) (2009), pp. 262-273.

[185] R.K. Dagda, T. Sultanaand J. Lyons-Weiler, Evaluation of the Consensus of Four Peptide Identification Algorithms for Tandem Mass Spectrometry Based Proteomics, *Journal of proteomics & bioinformatics* **3** pp. 39-47.

[186] R.J. Chalkley, P.R. Baker, K.F. Medzihradszky, A.J. Lynnand A.L. Burlingame, In-depth analysis of tandem mass spectrometry data from disparate instrument types, *Mol Cell Proteomics* **7** (12) (2008), pp. 2386-2398.

[187] K. Kandasamy, A. Pandeyand H. Molina, Evaluation of Several MS/MS Search Algorithms for Analysis of Spectra Derived from Electron Transfer Dissociation Experiments, *Analytical Chemistry* **81** (17) (2009), pp. 7170-7180.

[188] J.E. Eliasand S.P. Gygi, Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry, *Nat Methods* **4** (3) (2007), pp. 207-214.

[189] J. Peng, J.E. Elias, C.C. Thoreen, L.J. Lickliderand S.P. Gygi, Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome, *J Proteome Res* **2** (1) (2003), pp. 43-50.

[190] B.M. Balgley, T. Laudeman, L. Yang, T. Songand C.S. Lee, Comparative evaluation of tandem MS search algorithms using a target-decoy search strategy, *Molecular &*

Cellular Proteomics **6** (9) (2007), pp. 1599-1608.

[191] B. Ma, K.Z. Zhang, C. Hendrie, C.Z. Liang, M. Li, A. Doherty-Kirby *et al.*, PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry, *Rapid Communications in Mass Spectrometry* **17** (20) (2003), pp. 2337-2342.

[192] X. Liu, B. Shan, L. Xinand B. Ma, Better score function for peptide identification with ETD MS/MS spectra, *BMC Bioinformatics* **11 Suppl 1** pp. S4.

[193] J.D. Hoffert, G. Wang, T. Pisitkun, R.F. Shenand M.A. Knepper, An automated platform for analysis of phosphoproteomic datasets: application to kidney collecting duct phosphoproteins, *J Proteome Res* **6** (9) (2007), pp. 3501-3508.

[194] D.L. Tabb, W.H. McDonaldand J.R. Yates, 3rd, DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics, *J Proteome Res* **1** (1) (2002), pp. 21-26.

[195] S.A. Beausoleil, J. Villen, S.A. Gerber, J. Rushand S.P. Gygi, A probability-based approach for high-throughput protein phosphorylation analysis and site localization, *Nat Biotechnol* **24** (10) (2006), pp. 1285-1292.

[196] X. Jiang, M. Ye, K. Chengand H. Zou, ArMone: a software suite specially designed for processing and analysis of phosphoproteome data, *J Proteome Res* **9** (5) pp. 2743-2751.

[197] C.M. Bailey, S.M. Sweet, D.L. Cunningham, M. Zeller, J.K. Heathand H.J. Cooper, SLoMo: automated site localization of modifications from ETD/ECD mass spectra, *J Proteome Res* **8** (4) (2009), pp. 1965-1971.

[198] B.E. Ruttenberg, T. Pisitkun, M.A. Knepperand J.D. Hoffert, PhosphoScore: an open-source phosphorylation site assignment tool for MSn data, *J Proteome Res* **7** (7)

(2008), pp. 3054-3059.

[199] J.V. Olsenand M. Mann, Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation, *Proceedings of the National Academy of Sciences of the United States of America* **101** (37) (2004), pp. 13417-13422.

[200] F. Gnad, S. Ren, J. Cox, J.V. Olsen, B. Macek, M. Oroshi *et al.*, PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites, *Genome Biol* **8** (11) (2007), pp. R250.

[201] F. Diella, S. Cameron, C. Gemund, R. Linding, A. Via, B. Kuster *et al.*, Phospho.ELM: A database of experimentally verified phosphorylation sites in eukaryotic proteins, *BMC Bioinformatics* **5** (2004), pp. -.

[202] P.V. Hornbeck, I. Chabra, J.M. Kornhauser, E. Skrzypekand B. Zhang, PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation, *Proteomics* **4** (6) (2004), pp. 1551-1561.

[203] T.Y. Lee, H.D. Huang, J.H. Hung, H.Y. Huang, Y.S.O. Yangand T.H. Wang, dbPTM: an information repository of protein post-translational modification, *Nucleic Acids Research* **34** (2006), pp. D622-D627.

[204] G. Manning, D.B. Whyte, R. Martinez, T. Hunterand S. Sudarsanam, The protein kinase complement of the human genome, *Science* **298** (5600) (2002), pp. 1912-1934.

[205] S. Caenepeel, G. Charydczak, S. Sudarsanam, T. Hunterand G. Manning, The mouse kinome: Discovery and comparative genomics of all mouse protein kinases, *Proceedings of the National Academy of Sciences of the United States of America* **101** (32) (2004), pp. 11707-11712.

[206] M. Hjerrild, A. Stensballe, T.E. Rasmussen, C.B. Kofoed, N. Blom, T. Sicheritz-

Ponten *et al.*, Identification of phosphorylation sites in protein kinase A substrates using artificial neural networks and mass spectrometry, *Journal of Proteome Research* **3** (3) (2004), pp. 426-433.

[207] D. Schwartzand S.P. Gygi, An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets, *Nat Biotechnol* **23** (11) (2005), pp. 1391-1398.

[208] J.C. Obenauer, L.C. Cantleyand M.B. Yaffe, Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs, *Nucleic Acids Res* **31** (13) (2003), pp. 3635-3641.

[209] M.L. Millerand N. Blom, Kinase-specific prediction of protein phosphorylation sites, *Methods Mol Biol* **527** (2009), pp. 299-310, x.

[210] R. Linding, L.J. Jensen, A. Pasculescu, M. Olhovsky, K. Colwill, P. Bork *et al.*, NetworKIN: a resource for exploring cellular phosphorylation networks, *Nucleic Acids Res* **36** (Database issue) (2008), pp. D695-699.

[211] Y.H. Wong, T.Y. Lee, H.K. Liang, C.M. Huang, T.Y. Wang, Y.H. Yang *et al.*, KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns, *Nucleic Acids Res* **35** (Web Server issue) (2007), pp. W588-594.

[212] H.D. Huang, T.Y. Lee, S.W. Tzengand J.T. Horng, KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites, *Nucleic Acids Res* **33** (Web Server issue) (2005), pp. W226-229.

[213] D.K. Arrell, I. Neverova, H. Fraser, E. Marbanand J.E. Van Eyk, Proteomic analysis of pharmacologically preconditioned cardiomyocytes reveals novel phosphorylation of

myosin light chain 1, *Circ Res* **89** (6) (2001), pp. 480-487.

[214] C. Yuan, Q. Sheng, H. Tang, Y. Li, R. Zengand R.J. Solaro, Quantitative comparison of sarcomeric phosphoproteomes of neonatal and adult rat hearts, *Am J Physiol Heart Circ Physiol* **295** (2) (2008), pp. H647-656.

[215] Y. Oda, K. Huang, F.R. Cross, D. Cowburnand B.T. Chait, Accurate quantitation of protein expression and site-specific phosphorylation, *Proceedings of the National Academy of Sciences of the United States of America* **96** (12) (1999), pp. 6591-6596.

[216] S.E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Molecular & Cellular Proteomics* **1** (5) (2002), pp. 376-386.

[217] H.N. Zhu, S.Q. Pan, S. Gu, E.M. Bradburyand X. Chen, Amino acid residue specific stable isotope labeling for quantitative proteomics, *Rapid Communications in Mass Spectrometry* **16** (22) (2002), pp. 2115-2123.

[218] N. Ibarrola, D.E. Kalume, M. Gronborg, A. Iwahoriand A. Pandey, A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture, *Analytical Chemistry* **75** (22) (2003), pp. 6043-6049.

[219] C. Pan, J.V. Olsen, H. Dauband M. Mann, Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics, *Mol Cell Proteomics* **8** (12) (2009), pp. 2796-2808.

[220] M. Kruger, I. Kratchmarova, B. Blagoev, Y.H. Tseng, C.R. Kahnand M. Mann, Dissection of the insulin signaling pathway via quantitative phosphoproteomics, *Proc Natl Acad Sci U S A* **105** (7) (2008), pp. 2451-2456.

[221] D.L. Cunningham, S.M. Sweet, H.J. Cooperand J.K. Heath, Differential phosphoproteomics of fibroblast growth factor signaling: identification of Src family kinase-mediated phosphorylation events, *J Proteome Res* **9** (5) pp. 2317-2328.

[222] M. Hilger, T. Bonaldi, F. Gnadand M. Mann, Systems-wide analysis of a phosphatase knock-down by quantitative proteomics and phosphoproteomics, *Mol Cell Proteomics* **8** (8) (2009), pp. 1908-1920.

[223] C.C. Wu, M.J. MacCoss, K.E. Howell, D.E. Matthewsand J.R. Yates, Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis, *Analytical Chemistry* **76** (17) (2004), pp. 4951-4959.

[224] M. Kruger, M. Moser, S. Ussar, I. Thievessen, C.A. Luber, F. Forner *et al.*, SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function, *Cell* **134** (2) (2008), pp. 353-364.

[225] G.T. Cantin, J.D. Venable, D. Cociorvaand J.R. Yates, 3rd, Quantitative phosphoproteomic analysis of the tumor necrosis factor pathway, *J Proteome Res* **5** (1) (2006), pp. 127-134.

[226] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelband R. Aebersold, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nature Biotechnology* **17** (10) (1999), pp. 994-999.

[227] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan *et al.*, Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents, *Mol Cell Proteomics* **3** (12) (2004), pp. 1154-1169.

[228] J.V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horningand M. Mann, Higher-energy C-trap dissociation for peptide modification analysis, *Nat Methods* **4** (9) (2007),

pp. 709-712.

[229] T.J. Griffin, H. Xie, S. Bandhakavi, J. Popko, A. Mohan, J.V. Carlis *et al.*, iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer, *J Proteome Res* **6** (11) (2007), pp. 4200-4209.

[230] J.J. Coon, D. Phanstiel, R. Unwinand G.C. McAlister, Peptide Quantification Using 8-Plex Isobaric Tags and Electron Transfer Dissociation Tandem Mass Spectrometry, *Analytical Chemistry* **81** (4) (2009), pp. 1693-1698.

[231] J.J. Coon, D. Phanstiel, Y. Zhangand J.A. Marto, Peptide and protein quantification using iTRAQ with electron transfer dissociation, *Journal of the American Society for Mass Spectrometry* **19** (9) (2008), pp. 1255-1262.

[232] S.A. McLuckey, H.L. Han, D.J. Pappinand P.L. Ross, Electron transfer dissociation of iTRAQ labeled peptide ions, *Journal of Proteome Research* **7** (9) (2008), pp. 3643-3648.

[233] N. Mischerikow, P. van Nierop, K.W. Li, H.G. Bernstein, A.B. Smit, A.J. Heck *et al.*, Gaining efficiency by parallel quantification and identification of iTRAQ-labeled peptides using HCD and decision tree guided CID/ETD on an LTQ Orbitrap, *Analyst* **135** (10) pp. 2643-2652.

[234] F. Yang, S. Wu, D.L. Stenoien, R. Zhao, M.E. Monroe, M.A. Gritsenko *et al.*, Combined pulsed-Q dissociation and electron transfer dissociation for identification and quantification of iTRAQ-labeled phosphopeptides, *Anal Chem* **81** (10) (2009), pp. 4137-4143.

[235] B.L. Williamson, P.L. Ross, S. Pillai, B. Purkayastha, S. Danielsand D. Pappin, Protein quantitation using a novel 8-plex set of isobaric peptide labels, *Molecular &*

Cellular Proteomics **5** (10) (2006), pp. S55-S55.

[236] P.C. Wright, S.Y. Ow, M. Salim, J. Noirel, C. Evansand I. Rehman, iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly", *Journal of Proteome Research* **8** (11) (2009), pp. 5347-5355.

[237] A. Thompson, J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt *et al.*, Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, *Anal Chem* **75** (8) (2003), pp. 1895-1904.

[238] C.L. Nilsson, R. Dillon, A. Devakumar, S.D. Shi, M. Greig, J.C. Rogers *et al.*, Quantitative phosphoproteomic analysis of the STAT3/IL-6/HIF1alpha signaling network: an initial study in GSC11 glioblastoma stem cells, *J Proteome Res* **9** (1) pp. 430-443.

[239] O.A. Mirgorodskaya, Y.P. Kozmin, M.I. Titov, R. Korner, C.P. Sonksenand P. Roepstorff, Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using O-18-labeled internal standards, *Rapid Communications in Mass Spectrometry* **14** (14) (2000), pp. 1226-1232.

[240] M. Smith Jr Fau - Olivier, A.S. Olivier M Fau - Greeneand A.S. Greene, Relative quantification of peptide phosphorylation in a complex mixture using 18O labeling, (1531-2267 (Electronic)) pp.

[241] Y. Wang, S.J. Ding, W. Wang, J.M. Jacobs, W.J. Qian, R.J. Moore *et al.*, Profiling signaling polarity in chemotactic cells, *Proc Natl Acad Sci U S A* **104** (20) (2007), pp. 8328-8333.

[242] F. Yang, N. Jaitly, H. Jayachandran, Q. Luo, M.E. Monroe, X. Du *et al.*, Applying a targeted label-free approach using LC-MS AMT tags to evaluate changes in protein phosphorylation following phosphatase inhibition, *J Proteome Res* **6** (11) (2007), pp.

4489-4497.

[243] L. Cao, K. Yu, C. Banh, V. Nguyen, A. Ritz, B.J. Raphael *et al.*, Quantitative time-resolved phosphoproteomic analysis of mast cell signaling, *J Immunol* **179** (9) (2007), pp. 5864-5876.

[244] X. Xie, S. Feng, H. Vuong, Y. Liu, S. Goodisonand D.M. Lubman, A comparative phosphoproteomic analysis of a human tumor metastasis model using a label-free quantitative approach, *Electrophoresis* **31** (11) pp. 1842-1852.

[245] W.M. Old, K. Meyer-Arendt, L. Aveline-Wolf, K.G. Pierce, A. Mendoza, J.R. Sevinsky *et al.*, Comparison of label-free methods for quantifying human proteins by shotgun proteomics, *Mol Cell Proteomics* **4** (10) (2005), pp. 1487-1502.

[246] B. Zybailov, M.K. Coleman, L. Florensand M.P. Washburn, Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling, *Anal Chem* **77** (19) (2005), pp. 6218-6224.

[247] P.R. Cutillas, B. Geeringand M.D. Waterfield, Quantification of gel-separated proteins and their phosphorylation sites by LC-MS using unlabeled internal standards - Analysis of phosphoprotein dynamics in a B cell lymphoma cell line, *Molecular & Cellular Proteomics* **4** (8) (2005), pp. 1038-1051.

[248] H. Steen, J.A. Jebanathirajah, M. Springerand M.W. Kirschner, Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS, *Proceedings of the National Academy of Sciences of the United States of America* **102** (11) (2005), pp. 3948-3953.

[249] J.D. Hoffert, T. Pisitkun, G.H. Wang, R.F. Shenand M.A. Knepper, Quantitative

phosphoproteomics of vasopressin-sensitive renal cells: Regulation of aquaporin-2 phosphorylation at two sites, *Proceedings of the National Academy of Sciences of the United States of America* **103** (18) (2006), pp. 7159-7164.

[250] O.F. Bueno, L.J. De Windt, K.M. Tymitz, S.A. Witt, T.R. Kimball, R. Klevitsky *et al.*, The MEK1-ERK1/2 signaling pathway promotes compensated cardiac hypertrophy in transgenic mice, *EMBO J* **19** (23) (2000), pp. 6341-6350.

[251] S.O. Marx, S. Reiken, Y. Hisamatsu, T. Jayaraman, D. Burkhoff, N. Rosemblit *et al.*, PKA phosphorylation dissociates FKBP12.6 from the calcium release channel (ryanodine receptor): defective regulation in failing hearts, *Cell* **101** (4) (2000), pp. 365-376.

[252] C.L. Antos, N. Frey, S.O. Marx, S. Reiken, M. Gaburjakova, J.A. Richardson *et al.*, Dilated cardiomyopathy and sudden death resulting from constitutive activation of protein kinase a, *Circ Res* **89** (11) (2001), pp. 997-1004.

[253] B. Hoch, R. Meyer, R. Hetzer, E.G. Krauseand P. Karczewski, Identification and expression of delta-isoforms of the multifunctional Ca2+/calmodulin-dependent protein kinase in failing and nonfailing human myocardium, *Circ Res* **84** (6) (1999), pp. 713-721.

[254] Y. Takeishi, J. Abe, J.D. Lee, H. Kawakatsu, R.A. Walshand B.C. Berk, Differential regulation of p90 ribosomal S6 kinase and big mitogen-activated protein kinase 1 by ischemia/reperfusion and oxidative stress in perfused guinea pig hearts, *Circ Res* **85** (12) (1999), pp. 1164-1172.

[255] Y. Devary, R.A. Gottlieb, T. Smealand M. Karin, The mammalian ultraviolet response is triggered by activation of Src tyrosine kinases, *Cell* **71** (7) (1992), pp. 1081-1091.

[256] J.M. Kyriakisand J. Avruch, Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation, *Physiol Rev* **81** (2) (2001), pp. 807-869.

[257] C.M. Pombo, J.M. Kyriakis, J.V. Bonventreand T.L. Force, The P54 Stress-Activated Protein-Kinases (Sapks) Are Major C-Jun Amino-Terminal Kinases Activated by Ischemia and Reperfusion, *Circulation* **90** (4) (1994), pp. 303-303.

[258] J.C. Lee, S. Kumar, D.E. Griswold, D.C. Underwood, B.J. Vottaand J.L. Adams, Inhibition of p38 MAP kinase as a therapeutic strategy, *Immunopharmacology* **47** (2-3) (2000), pp. 185-201.

[259] F. Gao, T.L. Yue, D.W. Shi, T.A. Christopher, B.L. Lopez, E.H. Ohlstein *et al.*, P38 MAPK inhibition reduces myocardial reperfusion injury via inhibition of endothelial adhesion molecule expression and blockade of PMN accumulation, *Cardiovascular Research* **53** (2) (2002), pp. 414-422.

[260] A.S. Clanachan, J.S. Jaswal, M. Gandhi, D.A. Bottorff, J. Coughlin, B.A. Finegan *et al.*, Effects of inhibition of myocardial extracellular-responsive kinase and P38 mitogen-activated protein kinase on mechanical function of rat hearts after prolonged hypothermic ischemia, *Transplantation* **75** (2) (2003), pp. 173-180.

[261] S.E. Harding, L.A. Brown, D.G. Wynne, C.H. Daviesand P.A. Poole-Wilson, Mechanisms of beta adrenoceptor desensitisation in the failing human heart, *Cardiovasc Res* **28** (10) (1994), pp. 1451-1460.

[262] B. Pieske, B. Beyermann, V. Breu, B.M. Loffler, K. Schlotthauer, L.S. Maier *et al.*, Functional effects of endothelin and regulation of endothelin receptors in isolated human nonfailing and failing myocardium, *Circulation* **99** (14) (1999), pp. 1802-1809.

[263] K. Asano, D.L. Dutcher, J.D. Port, W.A. Minobe, K.D. Tremmel, R.L. Roden *et al.*, Selective downregulation of the angiotensin II AT1-receptor subtype in failing human ventricular myocardium, *Circulation* **95** (5) (1997), pp. 1193-1200.

[264] Y. Takeishi, A. Bhagwat, N.A. Ball, D.L. Kirkpatrick, M. Periasamyand R.A. Walsh, Effect of angiotensin-converting enzyme inhibition on protein kinase C and SR proteins in heart failure, *Am J Physiol* **276** (1 Pt 2) (1999), pp. H53-62.

[265] Y.T. Xuan, Y. Guo, H. Han, Y. Zhuand R. Bolli, An essential role of the JAK-STAT pathway in ischemic preconditioning, *Proc Natl Acad Sci U S A* **98** (16) (2001), pp. 9050-9055.

[266] M. Harada, Y. Qin, H. Takano, T. Minamino, Y. Zou, H. Toko *et al.*, G-CSF prevents cardiac remodeling after myocardial infarction by activating the Jak-Stat pathway in cardiomyocytes, *Nat Med* **11** (3) (2005), pp. 305-311.

[267] K. Boengler, D. Hilfiker-Kleiner, H. Drexler, G. Heuschand R. Schulz, The myocardial JAK/STAT pathway: from protection to failure, *Pharmacol Ther* **120** (2) (2008), pp. 172-185.

[268] S. Negoro, K. Kunisada, E. Tone, M. Funamoto, H. Oh, T. Kishimoto *et al.*, Activation of JAK/STAT pathway transduces cytoprotective signal in rat acute myocardial infarction, *Cardiovasc Res* **47** (4) (2000), pp. 797-805.

[269] M.S. Bhuiyan, N. Shiodaand K. Fukunaga, Targeting protein kinase B/Akt signaling with vanadium compounds for cardioprotection, *Expert Opinion on Therapeutic Targets* **12** (10) (2008), pp. 1217-1227.

[270] Y. Yao, L. Li, C. Gaoand C. Shi, Sevoflurane postconditioning protects chronically-infarcted rat hearts against ischemia-reperfusion injury by activation of pro-survival

kinases and inhibition of mitochondrial permeability transition pore opening upon reperfusion, *Biol Pharm Bull* **32** (11) (2009), pp. 1854-1861.

[271] T. Matsui, J. Tao, F. del Monte, K.H. Lee, L. Li, M. Picard *et al.*, Akt activation preserves cardiac function and prevents injury after transient cardiac ischemia in vivo, *Circulation* **104** (3) (2001), pp. 330-335.

[272] P. Ping, C. Song, J. Zhang, Y. Guo, X. Cao, R.C.X. Li *et al.*, Formation of protein kinase Cε-Lck signaling modules confers cardioprotection, *Journal of Clinical Investigation* **109** (4) (2002), pp. 499-507.

[273] P. Ping, J. Zhang, W.M. Pierceand R. Bolli, Functional Proteomic Analysis of Protein Kinase C {{epsilon}} Signaling Complexes in the Normal Heart and During Cardioprotection, *Circ Res* **88** (1) (2001), pp. 59-62.

[274] R.C.X. Li, P. Ping, J. Zhang, W.B. Wead, X. Cao, J. Gao *et al.*, PKCε modulates NF-κB and AP-1 via mitogen-activated protein kinases in adult rabbit cardiomyocytes, *American Journal of Physiology - Heart and Circulatory Physiology* **279** (4) (2000), pp. H1679 -H1689-H1679 -H1689.

[275] R. Bolli, The Late Phase of Preconditioning, *Circ Res* **87** (11) (2000), pp. 972-983.

[276] P. Ping, J. Zhang, X. Cao, R.C. Li, D. Kong, X.L. Tang *et al.*, PKC-dependent activation of p44/p42 MAPKs during myocardial ischemia-reperfusion in conscious rabbits, *The American Journal of Physiology* **276** (5 Pt 2) (1999), pp. H1468-1481-H1468-1481.

[277] P. Ping, H. Takano, J. Zhang, X.-L. Tang, Y. Qiu, R.C.X. Li *et al.*, Isoform-Selective Activation of Protein Kinase C by Nitric Oxide in the Heart of Conscious Rabbits : A Signaling Mechanism for Both Nitric Oxide﷿Induced and Ischemia-Induced

Preconditioning, *Circ Res* **84** (5) (1999), pp. 587-604.

[278] G.S. Liu, M.V. Cohen, D. Mochly-Rosenand J.M. Downey, Protein Kinase C- [xi] is Responsible for the Protection of Preconditioning in Rabbit Cardiomyocytes, *Journal of Molecular and Cellular Cardiology* **31** (10) (1999), pp. 1937-1948.

[279] P. Ping, J. Zhang, Y. Qiu, X.-L. Tang, S. Manchikalapudi, X. Cao *et al.*, Ischemic Preconditioning Induces Selective Translocation of Protein Kinase C Isoforms {epsilon} and {eta} in the Heart of Conscious Rabbits Without Subcellular Redistribution of Total Protein Kinase C Activity, *Circ Res* **81** (3) (1997), pp. 404-414.

[280] M.O. Gray, J.S. Karlinerand D. Mochly-Rosen, A Selective ε-Protein Kinase C Antagonist Inhibits Protection of Cardiac Myocytes from Hypoxia-induced Cell Death, *Journal of Biological Chemistry* **272** (49) (1997), pp. 30945-30951.

[281] D.J. Hausenloy, A. Tsang, M.M. Mocanuand D.M. Yellon, Ischemic preconditioning protects by activating prosurvival kinases at reperfusion, *American Journal of Physiology-Heart and Circulatory Physiology* **288** (2) (2005), pp. H971-H976.

[282] D.J. Hausenloy, A. Tsangand D.M. Yellon, The reperfusion injury salvage kinase pathway: a common target for both ischemic preconditioning and postconditioning, *Trends Cardiovasc Med* **15** (2) (2005), pp. 69-75.

[283] D.J. Hausenloyand D.M. Yellon, New directions for protecting the heart against ischaemia-reperfusion injury: targeting the Reperfusion Injury Salvage Kinase (RISK)-pathway, *Cardiovasc Res* **61** (3) (2004), pp. 448-460.

[284] J. Layland, R.J. Solaroand A.M. Shah, Regulation of cardiac contractile function by troponin I phosphorylation, *Cardiovasc Res* **66** (1) (2005), pp. 12-21.

[285] R. Zhang, J. Zhao, A. Mandvenoand J.D. Potter, Cardiac troponin I

phosphorylation increases the rate of cardiac muscle relaxation, *Circ Res* **76** (6) (1995), pp. 1028-1035.

[286] J. van der Velden, Z. Papp, R. Zaremba, N.M. Boontje, J.W. de Jong, V.J. Owen *et al.*, Increased Ca2+-sensitivity of the contractile apparatus in end-stage human heart failure results from altered phosphorylation of contractile proteins, *Cardiovascular Research* **57** (1) (2003), pp. 37-47.

[287] O. Copeland, S. Sadayappan, A.E. Messer, G.J. Steinen, J. van der Veldenand S.B. Marston, Analysis of cardiac myosin binding protein-C phosphorylation in human heart muscle, *J Mol Cell Cardiol* **49** (6) pp. 1003-1011.

[288] L. Carrier, Cardiac myosin-binding protein C in the heart, *Arch Mal Coeur Vaiss* **100** (3) (2007), pp. 238-243.

[289] J. Machackova, J. Bartaand N.S. Dhalla, Myofibrillar remodelling in cardiac-hypertrophy, heart failure and cardiomyopathies, *Canadian Journal of Cardiology* **22** (11) (2006), pp. 953-968.

[290] A. Phrommintikuland N. Chattipakorn, Roles of cardiac ryanodine receptor in heart failure and sudden cardiac death, *International Journal of Cardiology* **112** (2) (2006), pp. 142-152.

[291] A. Kushnirand A.R. Marks, The ryanodine receptor in cardiac physiology and disease, *Adv Pharmacol* **59** pp. 1-30.

[292] R. Treinysand J. Jurevicius, L-type Ca2+ channels in the heart: structure and regulation, *Medicina (Kaunas)* **44** (7) (2008), pp. 491-499.

[293] S.O. Marx, J. Kurokawa, S. Reiken, H. Motoike, J. D'Armiento, A.R. Marks *et al.*, Requirement of a macromolecular signaling complex for beta adrenergic receptor

modulation of the KCNQ1-KCNE1 potassium channel, *Science* **295** (5554) (2002), pp. 496-499.

[294] J.T. Hulme, T. Scheuerand W.A. Catterall, Regulation of cardiac ion channels by signaling complexes: role of modified leucine zipper motifs, *J Mol Cell Cardiol* **37** (3) (2004), pp. 625-631.

[295] C.J. Vlahos, S.A. McDowelland A. Clerk, Kinases as therapeutic targets for heart failure, *Nature Reviews Drug Discovery* **2** (2) (2003), pp. 99-113.

[296] T. Force, K. Kuida, M. Namchuk, K. Parangand J.M. Kyriakis, Inhibitors of protein kinase signaling pathways: emerging therapies for cardiovascular disease, *Circulation* **109** (10) (2004), pp. 1196-1205.

[297] R.D. Edmondson, T.M. Vondriska, K.J. Biederman, J. Zhang, R.C. Jones, Y. Zheng *et al.*, Protein kinase C epsilon signaling complexes include metabolism- and transcription/translation-related proteins: complimentary separation techniques with LC/MS/MS, *Mol Cell Proteomics* **1** (6) (2002), pp. 421-433.

[298] H.C. Chou, Y.W. Chen, T.R. Lee, F.S. Wu, H.T. Chan, P.C. Lyu *et al.*, Proteomics study of oxidative stress and Src kinase inhibition in H9C2 cardiomyocytes: a cell model of heart ischemia-reperfusion injury and treatment, *Free Radical Biology and Medicine* **49** (1) (2010), pp. 96-108.

[299] P. Fernando, W. Deng, B. Pekalska, Y. DeRepentigny, R. Kothary, J.F. Kelly *et al.*, Active kinase proteome screening reveals novel signal complexity in cardiomyopathy, *Molecular & Cellular Proteomics* **4** (5) (2005), pp. 673-682.

[300] R. Tian, N. Musi, J. D'Agostino, M.F. Hirshmanand L.J. Goodyear, Increased adenosine monophosphate-activated protein kinase activity in rat hearts with pressure-

overload hypertrophy, *Circulation* **104** (14) (2001), pp. 1664-1669.

[301] S. Satoh, Y. Ueda, M. Koyanagi, T. Kadokami, M. Sugano, Y. Yoshikawa *et al.*, Chronic inhibition of Rho kinase blunts the process of left ventricular hypertrophy leading to cardiac contractile dysfunction in hypertension-induced heart failure, *J Mol Cell Cardiol* **35** (1) (2003), pp. 59-70.

[302] M.R. Morissette, V.P. Sah, C.C. Glembotskiand J.H. Brown, The Rho effector, PKN, regulates ANF gene transcription in cardiomyocytes through a serum response element, *American Journal of Physiology-Heart and Circulatory Physiology* **278** (6) (2000), pp. H1769-H1774.

[303] A.C. Morrison, P.A. Doris, A.R. Folsom, F.J. Nieto, E. Boerwinkleand R.A. Hegele, G-Protein {beta}3 Subunit and {{alpha}}-Adducin Polymorphisms and Risk of Subclinical and Clinical Stroke Editorial Comment : Candidate Genes for Stroke: If Elected, Will They Serve?, *Stroke* **32** (4) (2001), pp. 822-829.

[304] T. Toyofuku, H. Zhang, A. Kumanogoh, N. Takegahara, M. Yabuki, K. Harada *et al.*, Guidance of myocardial patterning in cardiac development by Sema6D reverse signalling, *Nat Cell Biol* **6** (12) (2004), pp. 1204-1211.

[305] B.J. Wilkinsand J.D. Molkentin, Calcium-calcineurin signaling in the regulation of cardiac hypertrophy, *Biochem Biophys Res Commun* **322** (4) (2004), pp. 1178-1191.

[306] A.K. Paulsson, S. Franklin, S.A. Mitchell-Jordan, S. Ren, Y. Wangand T.M. Vondriska, Post-translational regulation of calsarcin-1 during pressure overload-induced cardiac hypertrophy, *J Mol Cell Cardiol* **48** (6) pp. 1206-1214.

[307] M. Aplin, G.L. Christensenand J.L. Hansen, Pharmacologic perspectives of functional selectivity by the angiotensin II type 1 receptor, *Trends Cardiovasc Med* **18** (8)

(2008), pp. 305-312.

[308] B. Schulenberg, R. Aggeler, J.M. Beechem, R.A. Capaldiand W.F. Patton, Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye, *Journal of Biological Chemistry* **278** (29) (2003), pp. 27251-27255.

[309] N. Deng, J. Zhang, C. Zong, Y. Wang, H. Lu, P. Yang *et al.*, Phosphoproteome analysis reveals regulatory sites in major pathways of cardiac mitochondria, *Mol Cell Proteomics* **10** (2) pp. M110 000117.

[310] X. Yin, F. Cuello, U. Mayr, Z. Hao, M. Hornshaw, E. Ehler *et al.*, Proteomics analysis of the cardiac myofilament subproteome reveals dynamic alterations in phosphatase subunit distribution, *Mol Cell Proteomics* **9** (3) pp. 497-509.

[311] H. Lu, C. Zong, Y. Wang, G.W. Young, N. Deng, P. Souda *et al.*, Revealing the dynamics of the 20 S proteasome phosphoproteome: a combined CID and electron transfer dissociation approach, *Mol Cell Proteomics* **7** (11) (2008), pp. 2073-2089.

[312] E. Dubois, V. Richard, P. Mulder, N. Lamblin, H. Drobecq, J.P. Henry *et al.*, Decreased Serine(207) phosphorylation of troponin T as a biomarker for left ventricular remodelling after myocardial infarction, *European Heart Journal* **32** (1) (2011), pp. 115-123.

CHAPTER 3

CONCENTRATION-DEPENDENT EFFECTS OF THE SOY PHYTOESTROGEN GENISTEIN ON THE PROTEOME OF CULTURED CARDIOMYOCYTES

1. Introduction

Epidemiological evidence suggests that the intake of dietary soy-derived phytoestrogens may lower the incidence of cardiovascular disease (CVD) [1]. Recently, several studies have suggested that isoflavones, particularly genistein (GEN), an abundant phytoestrogen in soy, may provide protection to cardiac myocytes against ischemic stress [2-7]. However, the underlying mechanisms are still elusive and may involve diverse cellular processes in a concentration-dependent manner as previously reported in various cell types [8-10].

GEN is known for its estrogenic [11] and anti-estrogenic properties [12]. Specifically, GEN at low concentration (1-10 μM) has been shown to be an effective agonist for both estrogen receptors (ER) 1 and 2 [13, 14]. Experimentally, estrogen has been shown to reduce the infarct size in various *in vivo* models against ischemia insult [15-21]. Natural estrogen analogs like GEN have been proposed as a candidate for hormone replacement therapies and have been shown to have cardioprotective effects via the modulation of PI3K/Akt and mitochondrial permeability [3]. Like estrogen, GEN was also reported to modulate cAMP-PKA signaling in a non-genomic fashion

[22, 23].

However, in addition to estrogenic bioactivity, GEN also acts as a non-specific protein tyrosine kinase (TK) inhibitor [24], typically at concentrations >10 μM. Tyrosine kinase signaling has been shown to play critical roles in preconditioning or pharmacological cardioprotection and this signaling can be blocked by GEN [25-37].

Given the complexity of known cellular consequences and other possible biological impacts of GEN treatment, a proteomic approach was chosen to provide a comprehensive view of the molecular impact of GEN on cardiomyocytes. HL-1 cells were used as a model because they maintain a similar gene expression profile and a differentiated cardiac phenotype to that of adult mouse cardiomyocytes [38, 39]. To elucidate the concentration-dependent alteration in protein expression profile modulated by GEN exposure, we analyzed the proteome of HL-1 cells treated with physiologically relevant (1 μM) and pharmacological (50 μM) concentrations of GEN compared to that of a control group via two-dimensional electrophoresis (2DE). In this protocol, proteins were first pre-fractionated by two-stage sequential protein extraction technique followed by 2DE-silver stain procedure for proteomic analysis. We identified 25 and 62 differentially expressed proteins via tandem mass spectrometry and database matching in the low- and high-concentration treatments, respectively. Subsequent Gene Ontology (GO) and pathway analyses were launched to discover the links between differentially expressed protein responses to GEN treatment and possible cardioprotective mechanisms. To our knowledge, this is the

first systems biology investigation on the cellular effects of GEN in cardiomyocytes.

2. Material and Methods

2.1. Cell culture and treatments

HL-1 cardiomyocyte cells [39] were cultured in T75 flasks pre-coated with gelatin and fibronectin. Claycomb medium [39] supplemented with 10% fetal bovine serum, 100 U/mL penicillin, 100 μg/mL streptomycin, 0.1 mM norepinephrine, and 2 mM L-glutamine was changed daily. Upon confluency, cells were treated with either 1 or 50 μM of GEN for 24 h before harvesting. Control cells were treated with the same volume (<3 μL/mL of medium) of DMSO. Each treatment was performed in triplicate. For each condition, approximately 10^7 cells were collected by trypsinizing.

2.2. Two-stage hydrophilic and hydrophobic protein extraction

The cell suspension was washed twice with pH 7.4 PBS and centrifuged at 300 xg for 3 min before protein extraction. Protein extraction was achieved with a two-stage hydrophilic/hydrophobic solubilization protocol. Water-soluble proteins were first extracted by ultrasonication for 5 min in 1mL hydrophilic extraction solution containing 10 mM Tris-HCL (pH 8.0), 0.5 mM Pefabloc, 5 mM magnesium acetate, and 0.01% Triton X-100. Samples were centrifuged at 17 000 x g for 20 min which the supernatant was designated as the hydrophilic protein fraction. The resulting pellet was further washed twice with 0.5 mL of hydrophilic extraction solution then dissolved with 0.5 mL hydrophobic extraction solution containing 10 mM Tris-HCL

(pH 8.0), 7 M urea, 2 M thiourea, 0.5 mM Pefabloc, 5 mM magnesium acetate and 4% CHAPS. After centrifugation at 17 000 x g, the supernatant was collected as the hydrophobic protein fraction. Total protein concentration of both hydrophilic and hydrophobic protein fractions was determined using the RC-DC protein assay kit (Bio-Rad, Hercules, CA, USA) according to the instructions from the manufacturer. Bovine serum albumin (Sigma-Aldrich) was used as protein assay standard.

2.3. Two-dimensional electrophoresis

For the hydrophilic protein fraction, 300 μg of protein was brought up to a volume of 400 μL (1:3 or greater dilution) with rehydration buffer (8 M urea, 4% CHAPS, 10% glycerol, and 0.3% dithiothreitol). Eight microliters of IPG buffer (Bio-Rad) were added before loading the sample into a 18-cm pH 4-7 IPG ReadyStrip (Bio-Rad). For the hydrophobic protein fraction, 200 μg of protein were loaded using the same procedure. IPG strip was rehydrated in an Immobiline DryStrip tray (GE Healthcare, Piscataway, NJ, USA) for 16 h. IEF was performed using a Multiphor II electrophoresis system (GE Healthcare) at 20 ^{o}C. The IEF voltage was programmed to increase to 500 V linearly within 1 min, followed by a linear increase to 3500 V over 5 h, and then maintained constant at 3500 V for 17.5 h. Proteins were reduced by submerging the IPG strips in 3 mL equilibration buffer (6 M urea, 30% v/v glycerol, 2% w/v SDS, and 24 mM Tris-HCl pH 6.8) supplemented with 2% w/v dithiothreitol for 15 min. Alkylation was performed by submerging the strips in 3 mL equilibration buffer supplemented with 2.5% w/v iodoacetamide and a trace of bromophenol blue for 5 min. The strips were then loaded onto a 13.5% polyacrylamide SDS-PAGE gel

(18 cm x 20 cm x 1 mm). The second dimension of electrophoresis was performed at a constant current of 40 mA per gel for 4.5 h in a Protean II XL 2-D Multi-Cell system (Bio-Rad).

A MS-compatible silver stain protocol was used to visualize gel spots. Briefly, gels were first fixed in 40% v/v ethanol and 10% v/v acetic acid for 30 min, followed by sensitization in 30% v/v ethanol, 0.2% w/v sodium thiosulphate, and 7% w/v sodium acetate for 30 min. After gels were washed three times with Nanopure water (Barnstead, Garner, NC) for 5 min, a 20-min staining was carried out in a 0.25% w/v silver nitrate solution. Gels were washed twice with Nanopure water for 1 min each and visualized with 0.074‰ formaldehyde in a 2.5% w/v sodium carbonate solution. After visualization, gels were preserved in 1.5% w/v EDTA. In all steps, a total volume of 250 mL of solution was used per gel. Digital gel images were created using the UVP Bioimaging System and further processed by LabWorks 4.6 (UVP LLC, Upland, CA). Gel images were analyzed with Delta 2D v3.4 image analysis software (Decodon GmbH, Greifswald, Germany). Briefly, image wrapping and spot matching was performed automatically and manually validated afterward. After background subtraction, a fused image was created by adding all spots from all images as a collective universal proteome map. A unique-spots ID list was created based on the proteome map and reassigned back to each individual gel to assure complete spot matching. To prevent identification of false spots, spots were removed from consideration if the spot quality was < 0.25 and if the spot percent volume in the densitometry intensity analysis was < 0.002. Spot quantification was based on fold-

changes based on percent volume (% $vol_i = vol_i/\sum vol_n$) of all spots resolved in the gel, where vol_i is the volume of spot *i* in a gel containing *n* spots. Statistical analysis was carried out with the *Student's t* test included in the Delta 2D package, using a cut-off value of $p < 0.05$. Three biological replicates each with two technical replicate gels were performed for each sample. Spots with significantly densitometric value changes were further checked visually to exclude false spots, mismatches, and spots showing evidence of co-migration. All spots from both hydrophilic and hydrophobic fractions with significantly altered intensity were used to visualize the change in global expression pattern upon GEN treatments. Hierarchical cluster analysis employing Euclidean correlation and average linkaging was performed with Gene Cluster v3.0 [40]. The resulting heat map was displayed using TreeView v1.6 [41].

2.4. Protein identification

Gel spots were cut with a spot picker with 1.5 mm or 3.0 mm diameter PDM tips (Gel Company, San Francisco, CA) depending on spot size. Excised spots were destained with 80 μL 1:1 solution of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate for 15 min, and washed twice for 15 min in 100 μL 100 mM ammonium bicarbonate [42], and dehydrated with 100 μL 100% ACN. The supernatant was removed and gel plugs were dried in a SPD SpeedVac (Thermo Electron, Waltham, MA). Tryptic digestion was accomplished by incubation of the gel spots with 0.6 μg sequencing-grade trypsin (Promega, Madison, WI, USA) in 40 μL of 100 mM ABC for 12 h at 37 ^{o}C. After the incubation, the supernatant was collected. Peptides were further extracted by sequentially incubating gel plugs with 80 μL of 0.1 % formic acid

(FA) in 50% ACN, 0.1 % FA in 70% ACN, 40 mM ABC, and 100% ACN each for 15 min. The pooled supernatants from these steps were dried by SPD SpeedVac to a final volume of 10 μL. Peptides were desalted using ZipTips C18 pipette tips (Millipore) and reconstituted in 10 μL of 0.1% FA with 3% ACN prior to analysis on a ESI-Q-TOF (Agilent 6150, Santa Clara CA). Desalted protein digests were injected onto a G4240-62001 C-18 HPLC-Chip (40 nL enrichment column, 75 μm x 43 mm analytical column, 5 μm C-18SB-ZX, 300 Å, Agilent) interfaced with the ESI. Peptides were separated during a 30-min gradient of 16–90% ACN in 0.1% FA at 4 μL/min. Peptides were analyzed using the data-dependent MS/MS mode over the m/z range of 59–3000. Three precursor ions were selected for MS/MS following each survey scan and only ions exhibiting a detection intensity exceeding 1000 counts were selected for MS/MS fragmentation by collision induced dissociation (CID). MS/MS data were acquired in centroid mode. MS/MS spectra were converted into Mascot generic format (.mgf) and uploaded to Mascot v2.3.01 (Matrix Science, London, UK) for database searching against a target-reverse concatenated International Protein Index (IPI, European Bioinformatics Institute, http://www.ebi.ac.uk/IPI) mouse database (v3.74, 113868 sequences). The search parameters were set to allow for up to two missed cleavages, carboxyamidomethylation on cysteine and variable modification of oxidation on methionine, a peptide tolerance of 0.2 Da, and an MS/MS tolerance of 0.1 Da. All identifications from Mascot were transferred into Scaffold v3.00.03 (Proteome Software, Inc., Portland, OR) for validation, where peptides that met the following two criteria, whichever was stricter, were used for protein identifications: (1) ion

score must be equal to or higher than the identification score from each individual Mascot search result, and (2) the score of a +2, +3, or +4 ion must be equal or higher than 34, 37 or 40, respectively. For protein identification, the match with the highest protein score (must be $p < 0.05$) and with at least with two unique peptide matches screened by the aforementioned peptide criteria was considered as the protein identification. Common contaminants such as keratins and trypsin were excluded from the results.

2.5. Bioinformatic analysis

To better enhance the biological significance of the quantitative proteomics results, lists of differentially expressed proteins from either the 1 μM or 50 μM GEN-treated samples vs. the control samples were analyzed using MetaCore software (GeneGO). The MetaCore database was constructed by abstracting interconnecting information from a large fraction of biomedical literature on the functions of genes, protein, and compounds, interactions of biological molecules, and disease inter-relationships, and is mainly focused on mice, rats, and humans. Once the list of altered proteins from either comparison was uploaded to MetaCore, the most relevant biological process was enriched based on public GO databases. Moreover, the size of the intersection between the subset of uploaded proteins and the proteins on all pre-built pathway maps in the MetaCore database was computed. In order to know whether the experimentally identified proteins had an association to a particular biological process GO or a pre-built canonical pathway, a *p*-value was computed based on the hypergeometric distribution probability test. The *p*-value represents the odds of

having a given biological process GO or a pre-built network to be associated with the experimentally-identified proteins simply by chance. The *p*-value was calculated using the formula:

$$p\text{-}Value = \frac{R!\,n!\,(N-R)!\,(N-n)!}{N!} \sum_{i=\max(r,R+n-N)}^{\min(n,R)} \frac{1}{i!\,(R-i)!\,(n-i)!\,(N-R-n+i)!}$$

where N is the total number of objects (i.e. all mouse genes), in the MetaCore database, R is the number of network objects corresponding to the list of differentially expressed proteins identified from the experiment, n is the total number of nodes in each small network, and r is the number of proteins from the experimental data that associate with each small network. The most relevant GO biological processes and pathways were then prioritized based on their statistical relevance with respect to the uploaded data sets. For GO enrichment, a false discovery rate (FDR) threshold of 0.01% was used to select significant processes considering the *p*-value of each associated process with respect to the total number of process in the public database. In order to maintain specificity in the GO analysis, GO terms within the top three ancestor terms in the GO Hierarchical rank and GO terms specifically designated to biological process in other cell types were dropped. Additionally, network construction based on Dijkstra's Shortest Path Algorithm was used to portray the potential interaction between GEN-responsive proteins. A maximum of two steps in each path were pre-set so that if there were no direct interaction between two proteins, an intermediate component connecting to both proteins was be added to the map to fill the gap.

3. Results and Discussion

3.1. Effect of genistein treatment on the proteome of HL-1 cardiomyocytes

For both soluble and insoluble protein fractions, a representative 2-D proteome map was generated by Delta 2D using the image fusion function in the union mode (Fig. 3.1). Overall, approximately 1,500 spots and 800 spots were visualized in gels from water-soluble and insoluble protein fractions, respectively. Among these 2,300 spots, 39 and 99 spots had different volumes (t-test, $p<0.05$) in samples from the low and high concentration GEN treatments compared to the untreated cells. Fig. 3.2A/B shows volcano plots of the entire data, set highlighting proteins expressed at significantly different (*t*-test $p < 0.05$) levels in the two GEN-treated HL-1 samples vs. controls. The global proteomic change patterns in GEN-treated groups vs. control populations were visualized by cluster analysis (Fig.3.2C). As expected, the six replicates in each experimental group cluster together, and the expression alteration pattern differs between the 1 μM GEN-treated and the 50 μM GEN-treated groups.

In total, 25 and 62 proteins were successfully identified by ESI-MS/MS from the low and high concentration GEN groups, respectively. Some of the spots were very faint and close to the limit of 2DE detection. Consequently, after gel excision and digestion, these spots were below the sensitivity of the MS instrument. Identified spots are summarized in Appendix I, along with protein name, IPI ID, number of unique peptide IDs, molecular weight, the fold-changes of both low and high concentration treatment versus control, and associated biological process GO terms are listed. Protein identification details for each spot can be found in Appendix II.

To assess the global trends of changes in cellular functions, lists of differentially expressed proteins in both low and high concentration GEN treatments were analyzed by MetaCore (GeneGO). The top 20 GO processes affected by 1 μM and 50 μM GEN treatment are summarized in Figure 3.3, and the complete list of significant relevant GO processes can be found in the Appendix III. Interestingly, both levels of GEN exposure had significant impacts on the energy metabolism GO processes, including glucose catabolic process and regulation of ATPase activity. According to the GO analysis, the low concentration GEN treatment also significantly impacted the anti-apoptosis process.

GeneGO pathways strongly associated with GEN treatments are summarized in Table 3.1. Several pathways, including cytoskeleton remodeling and hypoxia induced Akt-HIF1 activation, were highly enriched in both low and high concentration GEN-treated samples. The glycolysis and gluconeogenesis pathways were also enriched in the high-level GEN treatment. In Figure 3.4, interactive networks of the identified proteins that are modulated by GEN treatments are shown. Through analysis of these two networks, core transcriptional factors that hypothetically could mediate the GEN-induced protein expression alteration can be selected for future investigation. Transcription factors including estrogen receptor 1 (ER1), androgen receptors, jun proto-oncogene (AP-1/c-Jun), SP1 transcription factor, heat shock transcription factor 1 (HSF1), tumor protein p53, and vitamin D receptor (VDR) were found to be interactive hubs in the two networks. In addition, we found that v-myc

myelocytomatosis viral oncogene homolog (c-Myc), E2F transcription factor 1, early growth response 1 (EGR1), forkhead box O3 (FOXO3A), hypoxia inducible factor 1 (HIF1), signal transducer and activator of transcription 3 (STAT3), and YY1 transcription factor may also be involved with the expression regulation of genes induced by the high concentration of GEN, while cAMP-responsive element binding protein 1 (CREB1) may be involved with expression regulation of genes response to low concentration GEN treatment.

In the following sections, proteins and pathways involved in six functional areas are described and their potential linkages to the observed cardioprotective role of GEN discussed.

3.2. Stress response and protein folding machinery

The proteins identified in this study suggest that GEN treatment can affect the expression level of chaperones or heat shock proteins. Among the differentially expressed chaperones is heat shock 70 kDa protein 5 (GRP78 or Bip), the expression of which was enhanced by the high-concentration GEN treatment (+1.53 fold). This is consistent with previous findings that exposure to 100 μM GEN can significantly induce GRP78 expression in an ER-independent mode in both estrogen receptor-positive (MCF-7) and ER-negative (MDA-MB-231) cells [43]. GRP78 is known to work with protein disulfide-isomerase to fold misfolded proteins. GRP78 is also considered to be an anti-apoptotic factor since it can deactivate several pro-apoptotic factors such as BIK and BAX [44]. In addition, GRP78 has been reported to form

complexes with pro-caspases such as caspase-7 and caspase-12, preventing the activation of the pro-apoptotic caspase cascade [45]. Also, a study employing HeLa cells confirmed that suppression of GRP78 via siRNA silencing can cause apoptosis [46].

Another chaperone, the heat shock protein beta-1 (or HSP27), was also induced by high concentration GEN (+1.36 to +3.98 fold). Interestingly, in cells exposed to the low concentration of GEN, the expression pattern of HSP27 suggests potential changes in the HSP27 distribution (+1.33 fold in soluble fraction, -1.36 in insoluble fraction). Another mitochondrial chaperone involved in protein folding, the stress-70 protein (mitochondrial HSP70, 75 kDa glucose regulated protein, or GRP 75), was found to be up-regulated by the low-concentration GEN treatment (+1.57 to +2.06-fold). Jayakumar et al. indicated that up-regulation of HSP70 protects mitochondrial function in rats during cardiac ischemia-reperfusion injury [47]. Mitochondrial chaperone 60 kDa heat shock protein (HSP60) was up-regulated in the high-concentration GEN treatment (+1.63 to +4.42-fold). HSP60 is known for its important role in mitochondrial protein transport and assembly, and it also prevents polypeptide misfolding under stress conditions in the mitochondrial matrix. HSP60 is also considered to be an apoptosis regulator via its ability to form complexes with pro-apoptotic factors such as BAX, BAK, and Bcl-XL [48]. Interestingly, estradiol treatment was shown to activate heat shock factor-1 (HSF-1) [49]. HSF-1 is known to be the transcription regulator of HSP60 [50], and our results suggest that GEN may up-regulate HSP60 via the activation of HSF-1. The induction of N-myc downstream-

regulated protein 1 (NDRG1) was observed in cells exposed to 1 µM GEN (+1.38-fold). NGRG1 is inducible under hypoxic conditions and has been shown to play a protective role against hypoxia [51]. Interestingly, AP-1, which is positively regulated by ERs, was documented as a transcription enhancer of NDRG1 [52].

3.3. Cellular redox balance

Free radicals and oxidative stress are common mediators of cardiac ischemic injury [53]. Genistein, like many other plant-derived isoflavones, is considered to be an antioxidant [54]. Moreover, isoflavones have been shown to stimulate the antioxidant enzyme system (AOEs) in the vasculature [55-57]. In particular, GEN has been shown to have the ability to up-regulate endothelial nitric oxide synthase in rat vascular endothelial cells [58]. However, our observations suggest that GEN may have an overall negative impact on cardiac antioxidant enzyme system.

Peroxiredoxin-4 was slightly down-regulated in the low-GEN treatment (-1.31-fold). The peroxiredoxin family works with thioredoxin to remove hydroperoxides and to reduce oxidized proteins. The presence of GEN, an antioxidant [59], may have led to the feedback inhibition of the endogenous AOE system. Both the protein disulfide-isomerase (PDI) precursor, prolyl 4-hydroxylase subunit, and the protein disulfide isomerase A3 precursor were found to be significantly inhibited in cells dosed with the high level of GEN (-3.39 and -6.01-fold). Since they are key players in maintaining cellular redox homeostasis, and also members of the unfolded protein response network, PDIs are thought to have a vital role in decreasing oxidative stress-

induced apoptosis in cardiomyocytes in a mouse *in vivo* model, and in protecting against cardiac ischemia [60]. The significant decrease observed for PDIs may contribute to the attenuation of cardioprotection by GEN at high concentrations. Interestingly, glutathione synthetase was found to be overexpressed (+3.65-fold) in cells from the high-concentration GEN treatment. High levels of cellular antioxidant GSH can be induced by hypoxic stress and have been shown to provide cardiac protection against apoptosis-induced oxidative stress [61].

3.4. Apoptosis

Our 2DE data suggest that GEN treatment influences the apoptosis pathway in cardiomyocytes. Mitochondrial voltage-dependent anion-selective channel protein 2 (VDAC-2) was up-regulated by GEN at both concentrations (+1.52-fold at 1 μM, and +1.36-fold at 50 μM). VDACs help small hydrophilic molecules such as ATP move through the mitochondrial membrane and help to maintain the membrane integrity. In recent years, it was also found that VDACs participate in the regulation of apoptosis. Other than the fact that VDACs can deactivate pro-apoptosis factors such as BAK [62], VDACs *per se* may have a more direct impact on apoptosis via their open/closed status [63]. However, evidence for this hypothesis is inconsistent between reports [64] and goes beyond the scale of this proteomics study.

Another mitochondrial membrane protein up-regulated in the low GEN treatment group (+1.38-fold) is the import inner membrane translocase subunit TIM50. TIM50 regulates the translocation of transit peptide-containing proteins across the

mitochondrial inner membrane. Loss of TIM50 causes permeabilization of the mitochondrial membrane and release of cytochrome c, which initiates apoptosis [65].

Bcl-2-associated athanogene-2 (BAG2) expression was lower in cardiomyocytes treated with 50 μM GEN (-1.72-fold). BAG inhibits HSC70 by binding to HSC70 as a competitive antagonist of the co-chaperone Hip [66]. Thus, it has been proposed that BAG promotes cell survival by coordinating the function of these chaperones with the proteasome [67]. One study on rat primary cardiac myocytes showed that the oxidative stress-induced over-expression of BAGs can provide cardioprotection against apoptosis triggered by hypoxia [68]. Over-expression of the acidic protease cathepsin D was found in the 50 μM GEN-treated group (+1.61-fold). Cathepsin D is emerging as an apoptotic contributor of cardiomyocyte death under oxidative stress, which may cause lysosomal destabilization and acidification [69]. Interestingly, in other cell model, GEN was shown to trigger apoptosis via the ER-dependent up-regulation of cathepsin D [70]. Our analysis suggests that high levels of GEN may induce apoptosis in cardiomyocytes accordance with previous reports on other cell types treated by GEN [70-73].

3.5. Energy metabolism

Four critical enzymes that convert glucose to pyruvate were down-regulated in cardiomyocytes exposed to 50 μM GEN (but not at the lower level): triosephosphate isomerase (-5.01-fold), phosphoglycerate kinase 1 (-1.47 to -2.00-fold), alpha-enolase (-1.72-fold) and pyruvate kinase isozyme M2 (-1.63 to -3.06-fold). This suggests that

glycolysis was inhibited by the 50 μM GEN treatment but not by the 1 μM GEN treatment. Interestingly, the alpha subunit of pyruvate dehydrogenase E1, which converts pyruvate to acetyl-CoA, was up-regulated in cells exposed to both GEN levels (+1.59, +2.34-fold), suggesting that the cells might shift to other catabolic pathways to generate pyruvate upon GEN treatment.

A significant decrease in the level of NADH dehydrogenase [ubiquinone] flavoprotein 2 was found in the 50 μM GEN treatment (-1.23 to -3.59-fold), while another important component of electron transfer chain, electron transfer flavoprotein subunit beta (ETF), was up-regulated in the same cells (+1.65-fold). ETF carries electrons generated from fatty acid and amino acid catabolism to the electron transfer chain in mitochondria. Interestingly, two enzymes involved with fatty acid and amino acid catabolism were up-regulated in 50 μM GEN-treated cells. Aspartate transaminase (AST) was induced by exposure to 50 μM GEN (+2.33-fold). One key enzyme in the beta-oxidation of unsaturated fatty acid, the delta(3,5)-delta(2,4)-dienoyl-CoA isomerase, which also has enoyl-CoA hydratase activity, was up-regulated by the 50 μM GEN treatment (+1.52-fold). These changes suggest that the bioenergetics of cardiomyocytes shift from glycolysis to fatty acid and amino acid catabolism upon treatment with 50 μM GEN. Such a shift in energy utilization may render the cardiomyocytes even more unprepared for anaerobic metabolism upon ischemic insult.

3.6. DNA integrity and RNA transcription/processing

The levels of two proteins involved in RNA transport and processing, heterogeneous nuclear ribonucleoproteins (hnRNPs) A/B and H, were impacted by GEN. The hnRNP A/B was under-expressed in the 1-μM GEN-treated cells (-1.97-fold). These hnRNP proteins form the core of the ribonucleoprotein complex that associates with pre-mRNA splicing, and nucleo-cytoplasmic export of nascent mRNAs. It is also involved with DNA replication and repair [74]. In addition, hnRNP H was up-regulated in cells exposed to both low and high concentrations of GEN (+1.48 and 1.58-fold). The hnRNP H has been shown to modulate the expression of several genes in a tissue-specific fashion [75], Interestingly, hnRNP H was shown to be up-regulated by GEN in leukemia HL-60 cells in a proteomics study [76]. Another proteomics study [77] showed that hnRNP H was also up-regulated when human U937 cells were exposed to hydrogen peroxide. However, little is known about the functional significance of hnRNP isoforms in cardiomyocytes.

A decrease in prohibitin expression was found when the cells were treated with 50 μM GEN (-1.51 to -2.19-fold). Prohibitin is known to inhibit DNA synthesis and thus inhibit cell proliferation. Little is known about its role in differentiated cells; however, two recent studies revealed that it can suppress estrogen signaling [78] and androgen signaling [79]. Another protein involved with DNA synthesis/repair, the proliferating cell nuclear antigen (PCNA), was altered in cells treated with GEN. In contrast to previous reports of enhanced expression of PCNA by estrogen treatment [80], PCNA was suppressed following exposure to the low concentration GEN (-2.19-fold).

3.7. Cytoskeletal remodeling mobility/contractility

The actin-related protein 2/3 (Arp2/3) complex is known to mediate the formation of branched actin networks. Arp2/3 subunit 5 was found to be up-regulated in cardiomyocytes exposed to 1 μM and 50 μM GEN (+1.51-fold and +2.27-fold). However, little is known about the exact role of this p16 subunit. Interestingly, tubulin beta-5, one subunit of cytoplasmic microtubules, was down-regulated by exposure to 1 μM GEN. The actin thin filament binding protein, tropomyosin, regulates actin mechanics and muscle contraction. The low-level GEN treatment slightly increased the expression (+1.32-fold) of tropomyosin alpha chain, while a decrease in the amounts of both alpha and beta chain of tropomyosin was observed in the high-concentration GEN treatment (-1.61 and -1.48-fold). Several isoforms of contractile protein myosins such as myosin light chain 2, 6, and heavy chain 6 were all up-regulated (+1.79, +1.39, and +6.47-fold, respectively) in the 50-μM GEN treatment, although myosin light chain 4 was down-regulated (-2.18-fold). Contractility modulator calponin-3 was also down-regulated by 50 μM GEN (-1.47-fold). Calponin has been shown to inhibit Mg^{2+} ATPase activity in a reconstituted actomyosin system, thereby inhibiting the unloaded shortening velocity [81]. Based on these findings, exposure to GEN at the 50 μM level may have an important effect on the cell morphology and contractility of cardiac myocytes.

3.8. Protein phosphorylation and cellular signaling

While GEN is a well-documented tyrosine kinase inhibitor [24], our results suggest

that GEN may affect the phosphoproteome via another mechanism. A significant change in the level of low molecular weight phosphotyrosine protein phosphatase was observed in cells exposed to GEN at both concentrations (+2.14 and +2.34-fold). This phosphatase can non-specifically hydrolyze the phosphor group from phosphor-tyrosine residues. This evidence further supports the conclusion that GEN can suppress the cellular signaling via tyrosine phosphorylation.

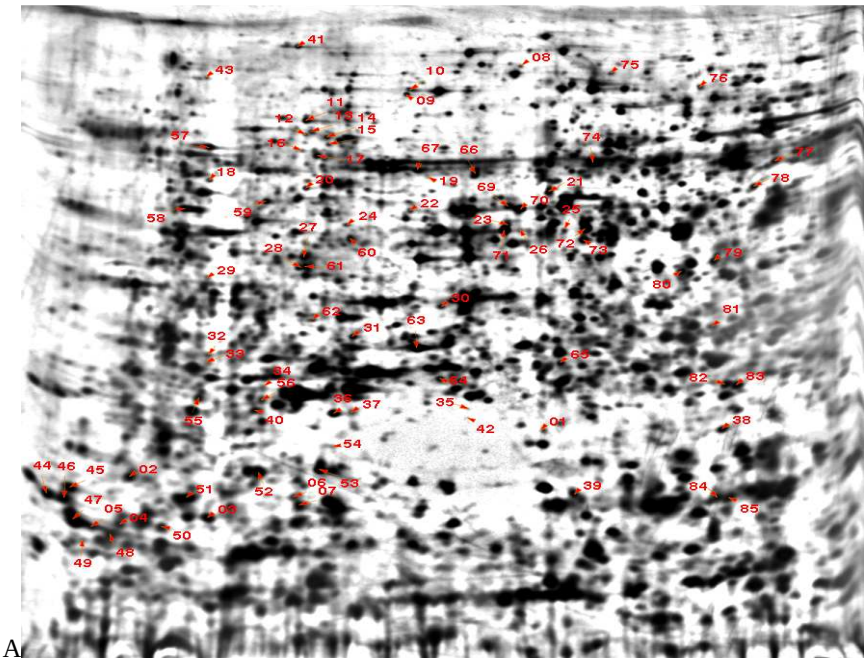
Another signaling protein, the growth factor receptor-bound protein 2 (Grb2) was down-regulated (-2.59-fold) in the 50-μM GEN treatment. Grb2 contains one SH2 domain, which recognizes and binds to the phosphorylated tyrosine motif of multiple signaling partners [82]. This suggests that the high level of GEN treatment may block the Grb2-related signaling.

4. Concluding Remarks

Genistein has been suggested to protect cardiomyocytes against adverse stress in recent years. Although GEN is a well-known estrogenic component and a tyrosine kinase inhibitor, our understanding of the complex cellular and molecular impact of GEN on cardiomyocytes is still incomplete. In this study, a global expression analysis using 2DE-based proteomics revealed the molecular impact of GEN in HL-1 cardiomyocytes on a concentrations dependent fashion. The use of two-stage hydrophilic and hydrophobic protein extraction significantly improves the coverage of cardiac proteome. Out of 2300 spots resolved, 25 and 62 significantly changed

proteins in response to low and high concentrations of GEN were identified, respectively. The protein expression profile suggests that exposure to 1 μM GEN enhances expression of heat shock proteins and anti-apoptotic proteins. At 50 μM, GEN down-regulates glycolytic proteins and antioxidant enzymes and potentially makes cardiomyocytes more susceptible to energy depletion and apoptosis. Significant expression changes in cytoskeletal protein machinery were also observed with the 50 μM-GEN treated cardiomyocytes. Moreover, MetaCore GO analysis suggests that GEN at low concentration significantly influences the anti-apoptosis process and both low- and high- concentration of GEN treatment have significant impact on glucose catabolic process and regulation of ATPase activity. Pathways analysis shown cytoskeleton remodeling and hypoxia induced Akt-HIF1 activation was represented by proteins responding to both low- and high- concentration GEN treatments. Enzymes of the glycolysis and gluconeogenesis pathway were also enriched following the high concentration GEN treatment. Although proteomics discoveries require further validation works, we hope this study serves as a valuable step towards our long-term goal of clarifying the complexity of genistein's impact on cardiomyocytes.

A



B

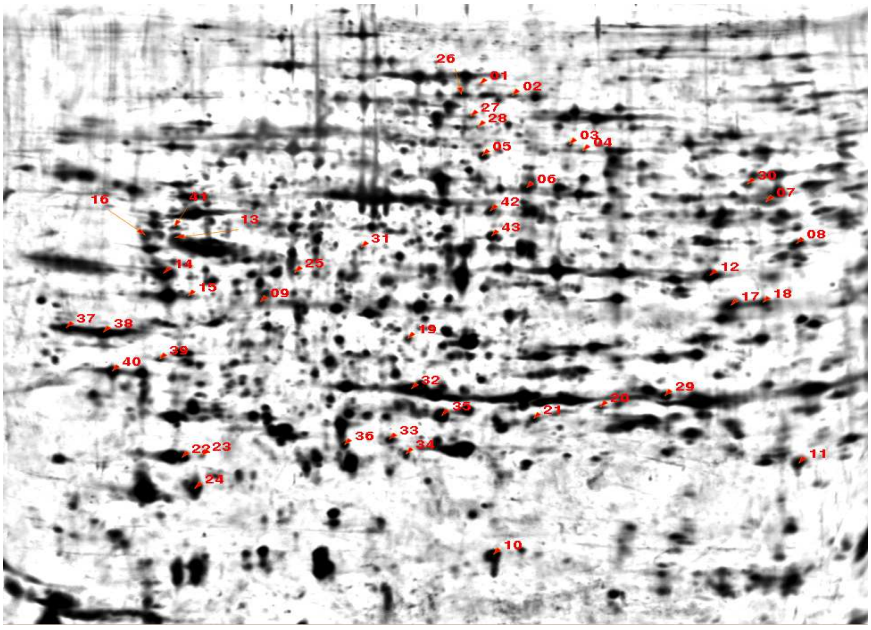


Figure 3.1. Representative 2DE map of hydrophilic (A) and hydrophobic (B) protein extracts from HL-1 cardiomyocytes. Three-hundred μg (hydrophilic fraction) or 200 μg (hydrophobic fraction) were separated in a pH 4-7 IPG strip as the first dimension and SDS-PAGE gels (13.5% acrylamide) in the second dimension. Gels were visualized by silver nitrate staining. From the analysis with the Delta 2D software, approximately 1500 spots and 800 spots were detected from the gels with the hydrophilic and hydrophobic fractions, respectively. From these 2,300 spots, 39 and 99 spots (labeled) were differentially expressed ($p<0.05$) in cells treated with low (1 μM) and high (50 μM) concentrations of GEN compared to untreated cells, respectively.

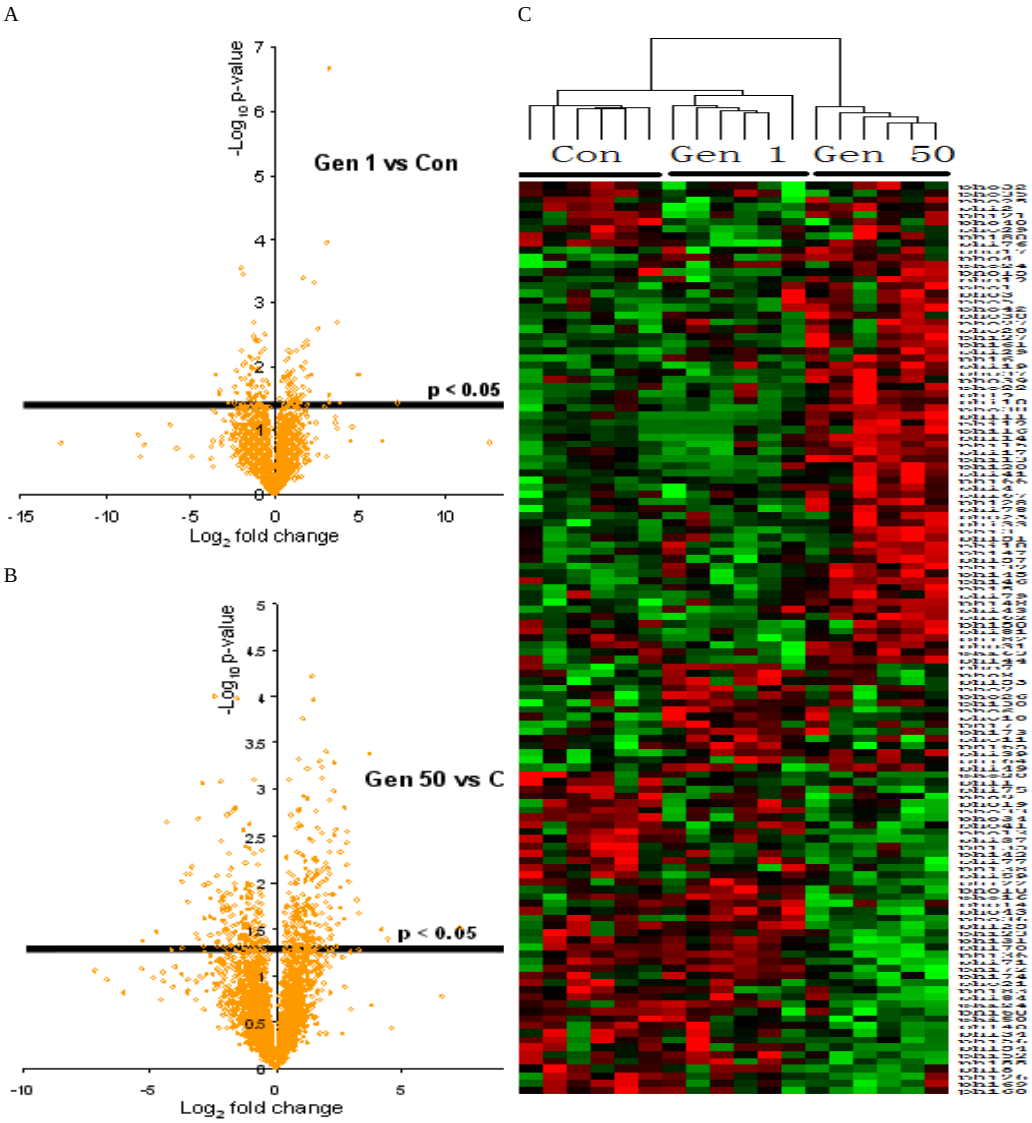
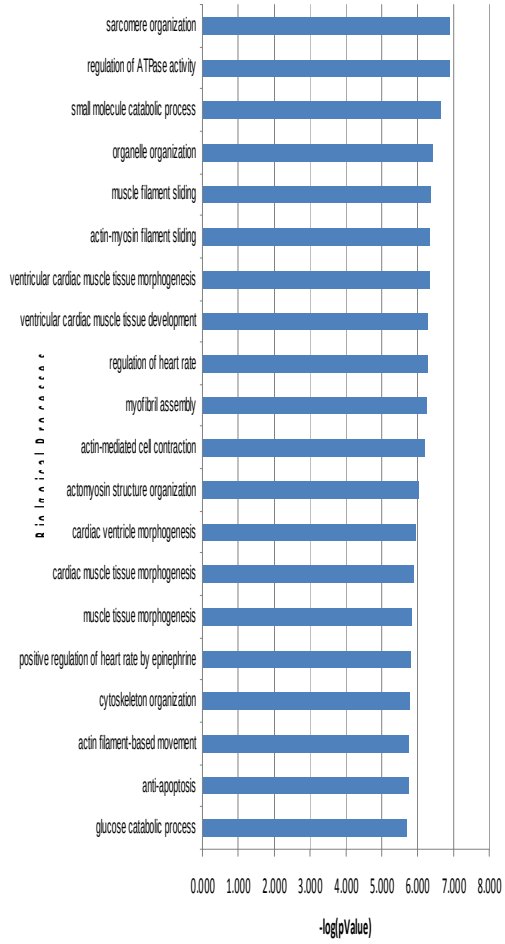


Figure 3.2. Proteomic patterns with spots from hydrophobic and hydrophilic fractions merged together. A/B. Volcano plots of the entire set of spots quantified during 2DE image analysis comparing Gen 1 vs Con and Gen 50 vs Con. Each point represents the fold-change in log2 scale between Gen treated group vs control group plotted against the level of statistical significance in log10 scale. Solid lines represent the significance threshold of $p < 0.05$ (Student's *t*-test). Spots above this line were considered as differentially expressed. C. Heat map representation of clustering analysis. Experimental samples are clustered on the horizontal axis and protein spots on the vertical axis with a relative color scale ranging from -0.75 (green) to +0.75 (red). Compared to control group, Gen 1 and Gen 50 shown distinctive global proteomic features respectively.

A



B

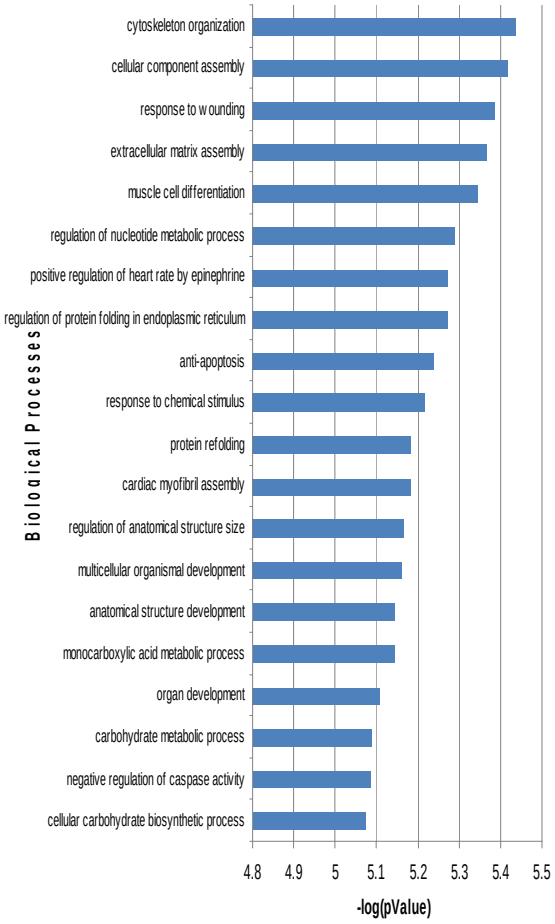


Figure 3.3. High-scored biological processes in 1 μM GEN treated (A) and 50 μM GEN treated (B) cardiomyocytes. –log(*p* value) > 2 (*p* value<0.01) was considered significant.

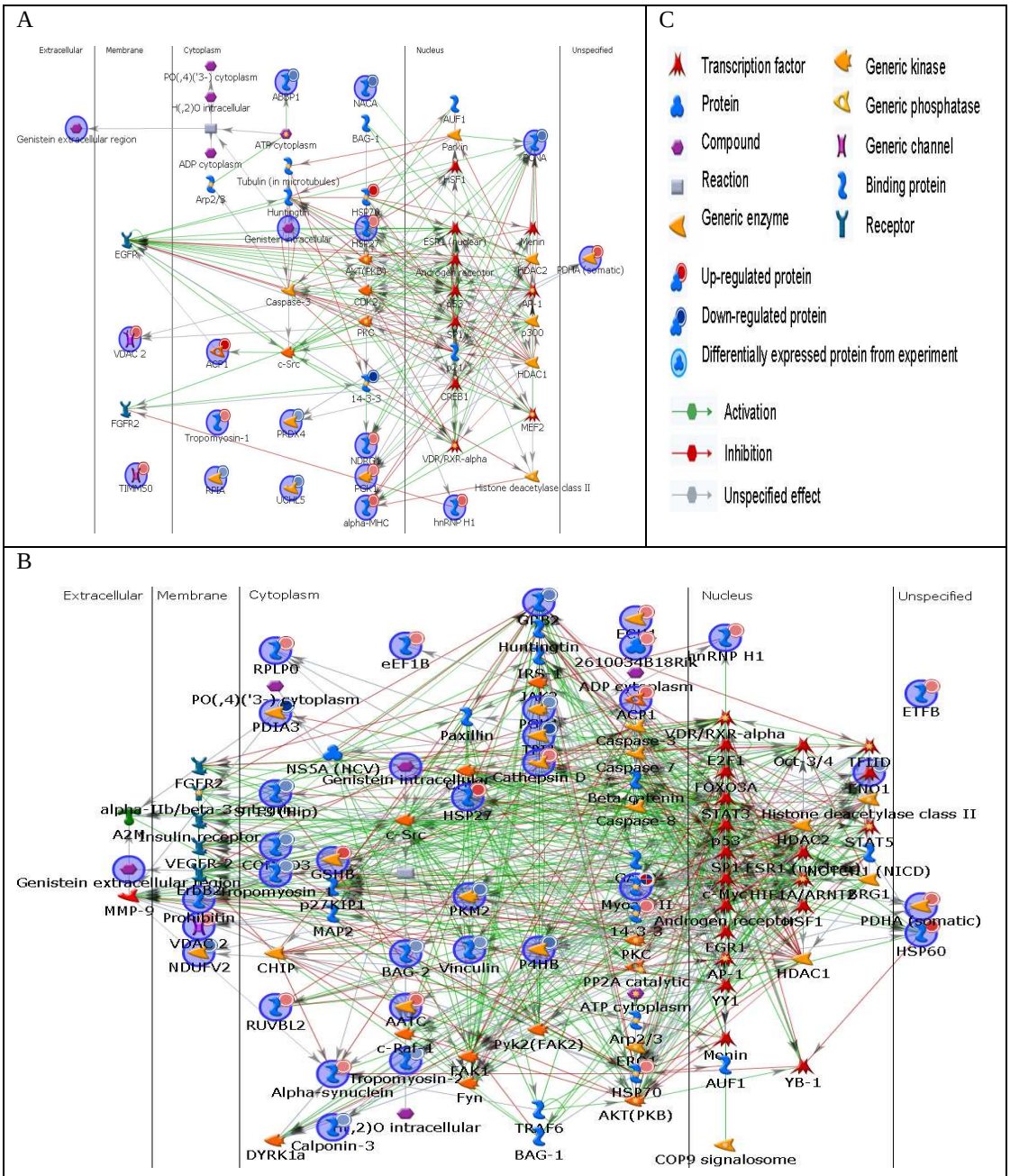


Figure 3.4. Biological network analysis of proteins differentially expressed response to 1μM (A) and 50 μM (B) GEN treatment using the Build Network tool provided by MetaCore. The network was generated using Dijkstra's shortest path algorithm to find the shortest directed interactions between proteins, with maximum of 2 intermediate steps. All connections were curated by MetaCore database. Network legend (C): Nodes represent proteins or compounds, lines between nodes indicate molecular interaction. For each interaction, the biological consequences, either activation or inhibition, were indicated by red or green color. Gray line means interaction with unspecified effect. Experimentally identified proteins that are modulated by GEN are denoted by a circle around the nodes. Both extracellular and cytoplasmic GEN were added to the map to show all known connections to those proteins.

Table 3.1. Differentially altered proteins identified in 1μM and 50 μM GEN treated cardiomyocytes were grouped for GeneGO MetaCore pathway analysis. A hypergeometric distribution probability test was carried out to decide if the experimentally identified proteins (objects) were involved with a particular biological pathway in the GeneGO MetaCore database. The p-value represents the odds of having a given pre-built network be associated with the list of experimentally- identified proteins simply by chance. NS means no significant relevance.

pathway name	total objects in MetaCore Pathway	objects in our study	p-value
		1 Gen/Control	1 Gen/Control
		50 Gen/Control	50 Gen/Control
Cytoskeleton remodeling_Regulation of actin cytoskeleton by Rho GTPases	23	1	4.12E-02
		4	1.40E-06
Glycolysis and gluconeogenesis	66	1	NS
		5	4.16E-06
Cytoskeleton remodeling_Cytoskeleton remodeling	102	1	NS
		5	3.52E-05
Transcription_Role of Akt in hypoxia induced HIF1 activation	27	2	1.06E-03
		3	1.34E-04
Muscle contraction_ GPCRs in the regulation of smooth muscle tone	83	1	NS
		4	2.48E-04
Proteolysis_Role of Parkin in the Ubiquitin-Proteasomal Pathway	24	2	8.40E-04
		2	3.61E-03
Mechanisms of CFTR activation by S-nitrosoglutathione	46	2	3.08E-03
		1	NS
Leucune, isoleucine and valine metabolism	54	2	4.22E-03
		2	1.75E-02
Regulation of lipid metabolism_Insulin regulation of fatty acid metabolism	89	2	1.11E-02
		3	4.44E-03
Apoptosis and survival_BAD phosphorylation	42	1	NS
		2	1.08E-02
Development_PIP3 signaling in cardiac myocytes	47	1	NS
		2	1.34E-02
Regulation of lipid metabolism_Insulin signaling: generic cascades	47	1	NS
		2	1.34E-02
Pyruvate metabolism	49	1	NS
		2	1.45E-02

REFERENCES

[1] A. Menotti, A. Keys, H. Blackburn, D. Kromhout, M. Karvonen, A. Nissinen *et al.*, Comparison of multivariate predictive power of major risk factors for coronary heart diseases in different countries: results from eight nations of the Seven Countries Study, 25-year follow-up, *J Cardiovasc Risk* **3** (1) (1996), pp. 69-75.

[2] L. Al-Nakkash, B. Markus, K. Bowden, L.M. Batia, W.C. Prozialeckand T.L. Broderick, Effects of acute and 2-day genistein treatment on cardiac function and ischemic tolerance in ovariectomized rats, *Gend Med* **6** (3) (2009), pp. 488-497.

[3] R. Tissier, X. Waintraub, N. Couvreur, M. Gervais, P. Bruneval, C. Mandet *et al.*, Pharmacological postconditioning with the phytoestrogen genistein, *J Mol Cell Cardiol* **42** (1) (2007), pp. 79-87.

[4] B. Deodato, D. Altavilla, G. Squadrito, G.M. Campo, M. Arlotta, L. Minutoli *et al.*, Cardioprotection by the phytoestrogen genistein in experimental myocardial ischaemia-reperfusion injury, *Br J Pharmacol* **128** (8) (1999), pp. 1683-1690.

[5] N. Couvreur, R. Tissier, S. Pons, M. Chenoune, X. Waintraub, A. Berdeaux *et al.*, The ceiling effect of pharmacological postconditioning with the phytoestrogen genistein is reversed by the GSK3beta inhibitor SB 216763 [3-(2,4-dichlorophenyl)-4(1-methyl-1H-indol-3-yl)-1H-pyrrole-2,5-dione] through mitochondrial ATP-dependent potassium channel opening, *J Pharmacol Exp Ther* **329** (3) (2009), pp. 1134-1141.

[6] E. Souzeau, S. Belanger, S. Picardand C.F. Deschepper, Dietary isoflavones during pregnancy and lactation provide cardioprotection to offspring rats in adulthood, *Am J Physiol Heart Circ Physiol* **289** (2) (2005), pp. H715-721.

[7] P. Zhai, T.E. Eurell, R.P. Cotthaus, E.H. Jeffery, J.M. Bahrand D.R. Gross, Effects of

dietary phytoestrogen on global myocardial ischemia-reperfusion injury in isolated female rat hearts, *Am J Physiol Heart Circ Physiol* **281** (3) (2001), pp. H1223-1232.

[8] Z.C. Dang, Dose-dependent effects of soy phyto-oestrogen genistein on adipocytes: mechanisms of action, *Obes Rev* **10** (3) (2009), pp. 342-349.

[9] A. Stadnicka, W.M. Kwok, D.C. Warltierand Z.J. Bosnjak, Protein tyrosine kinase-dependent modulation of isoflurane effects on cardiac sarcolemmal K(ATP) channel, *Anesthesiology* **97** (5) (2002), pp. 1198-1208.

[10] Z.C. Dangand C. Lowik, Dose-dependent effects of phytoestrogens on bone, *Trends Endocrinol Metab* **16** (5) (2005), pp. 207-213.

[11] M.W. Carter, W.W. Smart, Jr.and G. Matrone, Estimation of estrogenic activity of genistein obtained from soybean meal, *Proc Soc Exp Biol Med* **84** (2) (1953), pp. 506-508.

[12] K. Miyazaki, Novel approach for evaluation of estrogenic and anti-estrogenic activities of genistein and daidzein using B16 melanoma cells and dendricity assay, *Pigment Cell Res* **17** (4) (2004), pp. 407-412.

[13] M. Maggiolini, D. Bonofiglio, S. Marsico, M.L. Panno, B. Cenni, D. Picard *et al.*, Estrogen receptor alpha mediates the proliferative but not the cytotoxic dose-dependent effects of two major phytoestrogens on human breast cancer cells, *Mol Pharmacol* **60** (3) (2001), pp. 595-602.

[14] G.G. Kuiper, J.G. Lemmen, B. Carlsson, J.C. Corton, S.H. Safe, P.T. van der Saag *et al.*, Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta, *Endocrinology* **139** (10) (1998), pp. 4252-4263.

[15] M. van Eickels, R.D. Patten, M.J. Aronovitz, A. Alsheikh-Ali, K. Gostyla, F. Celestin *et al.*, 17-beta-estradiol increases cardiac remodeling and mortality in mice with

myocardial infarction, *J Am Coll Cardiol* **41** (11) (2003), pp. 2084-2092.

[16] C.H. Tsai, S.F. Su, T.F. Chouand T.M. Lee, Differential effects of sarcolemmal and mitochondrial K(ATP) channels activated by 17 beta-estradiol on reperfusion arrhythmias and infarct sizes in canine hearts, *J Pharmacol Exp Ther* **301** (1) (2002), pp. 234-240.

[17] T.M. Lee, S.F. Su, C.C. Tsai, Y.T. Leeand C.H. Tsai, Cardioprotective effects of 17 beta-estradiol produced by activation ofmitochondrial ATP-sensitive K(+)Channels in canine hearts, *J Mol Cell Cardiol* **32** (7) (2000), pp. 1147-1158.

[18] T.M. Lee, M.S. Lin, T.F. Chou, C.H. Tsaiand N.C. Chang, Adjunctive 17beta-estradiol administration reduces infarct size by altered expression of canine myocardial connexin43 protein, *Cardiovasc Res* **63** (1) (2004), pp. 109-117.

[19] S.L. Hale, Y. Birnbaumand R.A. Kloner, Estradiol, Administered Acutely, Protects Ischemic Myocardium in Both Female and Male Rabbits, *J Cardiovasc Pharmacol Ther* **2** (1) (1997), pp. 47-52.

[20] E.A. Booth, N.R. Obeidand B.R. Lucchesi, Activation of estrogen receptor-alpha protects the in vivo rabbit heart from ischemia-reperfusion injury, *Am J Physiol Heart Circ Physiol* **289** (5) (2005), pp. H2039-2047.

[21] E.A. Booth, M. Marchesi, E.J. Kilbourneand B.R. Lucchesi, 17 beta-estradiol as a receptor-mediated cardioprotective agent, *Journal of Pharmacology and Experimental Therapeutics* **307** (1) (2003), pp. 395-401.

[22] W.W. Ng, W. Keung, Y.C. Xu, K.F. Ng, G.P. Leung, P.M. Vanhoutte *et al.*, Genistein potentiates protein kinase A activity in porcine coronary artery, *Mol Cell Biochem* **311** (1-2) (2008), pp. 37-44.

[23] H. Teohand R.Y. Man, Enhanced relaxation of porcine coronary arteries after acute

exposure to a physiological level of 17beta-estradiol involves non-genomic mechanisms and the cyclic AMP cascade, *British Journal of Pharmacology* **129** (8) (2000), pp. 1739-1747.

[24] T. Akiyama, J. Ishida, S. Nakagawa, H. Ogawara, S. Watanabe, N. Itoh *et al.*, Genistein, a specific inhibitor of tyrosine-specific protein kinases, *J Biol Chem* **262** (12) (1987), pp. 5592-5595.

[25] C.P. Baines, L. Wang, M.V. Cohenand J.M. Downey, Protein tyrosine kinase is downstream of protein kinase C for ischemic preconditioning's anti-infarct effect in the rabbit heart, *J Mol Cell Cardiol* **30** (2) (1998), pp. 383-392.

[26] I.F. Benter, J.S. Juggi, I. Khan, M.H. Yousif, H. Canatanand S. Akhtar, Signal transduction mechanisms involved in cardiac preconditioning: role of Ras-GTPase, Ca2+/calmodulin-dependent protein kinase II and epidermal growth factor receptor, *Mol Cell Biochem* **268** (1-2) (2005), pp. 175-183.

[27] R.M. Fryer, J.E. Schultz, A.K. Hsuand G.J. Gross, Pretreatment with tyrosine kinase inhibitors partially attenuates ischemic preconditioning in rat hearts, *Am J Physiol* **275** (6 Pt 2) (1998), pp. H2009-2015.

[28] R.M. Fryer, J.E. Schultz, A.K. Hsuand G.J. Gross, Importance of PKC and tyrosine kinase in single or multiple cycles of preconditioning in rat hearts, *Am J Physiol* **276** (4 Pt 2) (1999), pp. H1229-1235.

[29] R.M. Fryer, Y. Wang, A.K. Hsu, H. Nagaseand G.J. Gross, Dependence of delta1-opioid receptor-induced cardioprotection on a tyrosine kinase-dependent but not a Src-dependent pathway, *J Pharmacol Exp Ther* **299** (2) (2001), pp. 477-482.

[30] R. Germack, M. Griffinand J.M. Dickenson, Activation of protein kinase B by

adenosine A1 and A3 receptors in newborn rat cardiomyocytes, *J Mol Cell Cardiol* **37** (5) (2004), pp. 989-999.

[31] N. Hedayati, S.J. Schomisch, J.L. Carino, J. Timothy Sherwood, E.J. Lesnefskyand B.L. Cmolik, Cardioprotection by St Thomas' solution is mediated by protein kinase C and tyrosine kinase, *J Surg Res* **113** (1) (2003), pp. 121-127.

[32] P. Htun, W.D. Ito, I.E. Hoefer, J. Schaperand W. Schaper, Intramyocardial infusion of FGF-1 mimics ischemic preconditioning in pig myocardium, *J Mol Cell Cardiol* **30** (4) (1998), pp. 867-877.

[33] M. Krenz, C.P. Baines, G. Heusch, J.M. Downeyand M.V. Cohen, Acute alcohol-induced protection against infarction in rabbit hearts: differences from and similarities to ischemic preconditioning, *J Mol Cell Cardiol* **33** (11) (2001), pp. 2015-2022.

[34] S. Lee, G. Chanoit, R. McIntosh, D.A. Zvaraand Z. Xu, Molecular mechanism underlying Akt activation in zinc-induced cardioprotection, *Am J Physiol Heart Circ Physiol* **297** (2) (2009), pp. H569-575.

[35] D.A. Liem, C.C. Gho, B.C. Gho, S. Kazim, O.C. Manintveld, P.D. Verdouw *et al.*, The tyrosine phosphatase inhibitor bis(maltolato)oxovanadium attenuates myocardial reperfusion injury by opening ATP-sensitive potassium channels, *J Pharmacol Exp Ther* **309** (3) (2004), pp. 1256-1262.

[36] S. Okubo, Y. Tanabe, K. Takeda, M. Kitayama, S. Kanemitsu, R.C. Kukreja *et al.*, Pretreatment with tyrosine kinase inhibitor attenuates the reduction of apoptosis 24 h after ischemic preconditioning, *Jpn J Physiol* **54** (2) (2004), pp. 143-151.

[37] T.L. Wang, Y.H. Huangand H. Chang, Somatostatin analogue mimics acute ischemic preconditioning in a rat model of myocardial infarction, *J Cardiovasc Pharmacol* **45** (4)

(2005), pp. 327-332.

[38] S.M. White, P.E. Constantinand W.C. Claycomb, Cardiac physiology at the cellular level: use of cultured HL-1 cardiomyocytes for studies of cardiac muscle cell structure and function, *American Journal of Physiology-Heart and Circulatory Physiology* **286** (3) (2004), pp. H823-H829.

[39] W.C. Claycomb, N.A. Lanson, Jr., B.S. Stallworth, D.B. Egeland, J.B. Delcarpio, A. Bahinski *et al.*, HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte, *Proc Natl Acad Sci U S A* **95** (6) (1998), pp. 2979-2984.

[40] M.B. Eisen, P.T. Spellman, P.O. Brownand D. Botstein, Cluster analysis and display of genome-wide expression patterns, *Proc Natl Acad Sci U S A* **95** (25) (1998), pp. 14863-14868.

[41] R.D. Page, TreeView: an application to display phylogenetic trees on personal computers, *Comput Appl Biosci* **12** (4) (1996), pp. 357-358.

[42] V.A. Zhukov, L.N. Shishkina, A.A. Sergeev, E.M. Malkova, E.I. Riabchikova, V.A. Petrishchenko *et al.*, [Comparative analysis of the susceptibility and productivity of respiratory tract target cells of mice and rats exposed to inflienza virus in vitro], *Vestn Ross Akad Med Nauk* (2) (2008), pp. 12-16.

[43] H.A. Lim, J.H. Kim, M.K. Sung, M.K. Kim, J.H. Parkand J.S. Kim, Genistein induces glucose-regulated protein 78 in mammary tumor cells, *J Med Food* **9** (1) (2006), pp. 28-32.

[44] Y. Fu, J. Liand A.S. Lee, GRP78/BiP inhibits endoplasmic reticulum BIK and protects human breast cancer cells against estrogen starvation-induced apoptosis, *Cancer*

Res **67** (8) (2007), pp. 3734-3740.

[45] R.K. Reddy, C. Mao, P. Baumeister, R.C. Austin, R.J. Kaufmanand A.S. Lee, Endoplasmic reticulum chaperone protein GRP78 protects cells from apoptosis induced by topoisomerase inhibitors: role of ATP binding site in suppression of caspase-7 activation, *J Biol Chem* **278** (23) (2003), pp. 20915-20924.

[46] T. Suzuki, J. Lu, M. Zahed, K. Kitaand N. Suzuki, Reduction of GRP78 expression with siRNA activates unfolded protein response leading to apoptosis in HeLa cells, *Arch Biochem Biophys* **468** (1) (2007), pp. 1-14.

[47] J. Jayakumar, K. Suzuki, I.A. Sammut, R.T. Smolenski, M. Khan, N. Latif *et al.*, Heat shock protein 70 gene transfection protects mitochondrial and ventricular function against ischemia-reperfusion injury, *Circulation* **104** (12 Suppl 1) (2001), pp. I303-307.

[48] S. Guptaand A.A. Knowlton, HSP60, Bax, apoptosis and the heart, *J Cell Mol Med* **9** (1) (2005), pp. 51-58.

[49] K.L. Hamilton, S. Guptaand A.A. Knowlton, Estrogen and regulation of heat shock protein expression in female cardiomyocytes: cross-talk with NF kappa B signaling, *J Mol Cell Cardiol* **36** (4) (2004), pp. 577-584.

[50] N.D. Trinklein, W.C. Chen, R.E. Kingstonand R.M. Myers, Transcriptional regulation and binding of heat shock factor 1 and heat shock factor 2 to 32 human heat shock genes during thermal stress and differentiation, *Cell Stress Chaperones* **9** (1) (2004), pp. 21-28.

[51] B. Chen, D.M. Nelsonand Y. Sadovsky, N-myc down-regulated gene 1 modulates the response of term human trophoblasts to hypoxic injury, *J Biol Chem* **281** (5) (2006), pp. 2764-2772.

[52] K. Salnikow, T. Kluz, M. Costa, D. Piquemal, Z.N. Demidenko, K. Xie *et al.*, The regulation of hypoxic genes by calcium involves c-Jun/AP-1, which cooperates with hypoxia-inducible factor 1 in response to hypoxia, *Mol Cell Biol* **22** (6) (2002), pp. 1734-1741.

[53] S. Sekili, P.B. McCay, X.Y. Li, M. Zughaib, J.Z. Sun, L. Tang *et al.*, Direct evidence that the hydroxyl radical plays a pathogenetic role in myocardial "stunning" in the conscious dog and demonstration that stunning can be markedly attenuated without subsequent adverse effects, *Circ Res* **73** (4) (1993), pp. 705-723.

[54] A. Arora, M.G. Nairand G.M. Strasburg, Antioxidant activities of isoflavones and their biological metabolites in a liposomal system, *Arch Biochem Biophys* **356** (2) (1998), pp. 133-141.

[55] R.C. Siow, F.Y. Li, D.J. Rowlands, P. de Winterand G.E. Mann, Cardiovascular targets for estrogens and phytoestrogens: transcriptional regulation of nitric oxide synthase and antioxidant defense genes, *Free Radic Biol Med* **42** (7) (2007), pp. 909-925.

[56] G.E. Mann, D.J. Rowlands, F.Y. Li, P. de Winterand R.C. Siow, Activation of endothelial nitric oxide synthase by dietary isoflavones: role of NO in Nrf2-mediated antioxidant gene expression, *Cardiovasc Res* **75** (2) (2007), pp. 261-274.

[57] G.E. Mann, B. Bonacasa, T. Ishiiand R.C. Siow, Targeting the redox sensitive Nrf2-Keap1 defense pathway in cardiovascular disease: protection afforded by dietary isoflavones, *Curr Opin Pharmacol* **9** (2) (2009), pp. 139-145.

[58] H. Siand D. Liu, Genistein, a soy phytoestrogen, upregulates the expression of human endothelial nitric oxide synthase and lowers blood pressure in spontaneously hypertensive rats, *J Nutr* **138** (2) (2008), pp. 297-304.

[59] R.M. Han, Y.X. Tian, Y. Liu, C.H. Chen, X.C. Ai, J.P. Zhang *et al.*, Comparison of flavonoids and isoflavonoids as antioxidants, *J Agric Food Chem* **57** (9) (2009), pp. 3780-3785.

[60] A. Severino, M. Campioni, S. Straino, F.N. Salloum, N. Schmidt, U. Herbrand *et al.*, Identification of protein disulfide isomerase as a cardiomyocyte survival factor in ischemic cardiomyopathy, *J Am Coll Cardiol* **50** (11) (2007), pp. 1029-1037.

[61] L. Lin, M.W. Andersand S.S. Shen, Mitochondrial glutathione protects cardiac myocytes from H2O2-induced apoptosis, *Biophysical Journal* **84** (2) (2003), pp. 97a-97a.

[62] E.H. Cheng, T.V. Sheiko, J.K. Fisher, W.J. Craigenand S.J. Korsmeyer, VDAC2 inhibits BAK activation and mitochondrial apoptosis, *Science* **301** (5632) (2003), pp. 513-517.

[63] A.P. Halestrap, S.J. Clarkeand S.A. Javadov, Mitochondrial permeability transition pore opening during myocardial reperfusion--a target for cardioprotection, *Cardiovasc Res* **61** (3) (2004), pp. 372-385.

[64] T.K. Rostovtseva, W. Tanand M. Colombini, On the role of VDAC in apoptosis: fact and fiction, *J Bioenerg Biomembr* **37** (3) (2005), pp. 129-142.

[65] Y. Guo, N. Cheong, Z. Zhang, R. De Rose, Y. Deng, S.A. Farber *et al.*, Tim50, a component of the mitochondrial translocator, regulates mitochondrial integrity and cell death, *J Biol Chem* **279** (23) (2004), pp. 24813-24825.

[66] S. Takayama, Z. Xieand J.C. Reed, An evolutionarily conserved family of Hsp70/Hsc70 molecular chaperone regulators, *J Biol Chem* **274** (2) (1999), pp. 781-786.

[67] J. Luders, J. Demandand J. Hohfeld, The ubiquitin-related BAG-1 provides a link between the molecular chaperones Hsc70/Hsp70 and the proteasome, *J Biol Chem* **275** (7)

(2000), pp. 4613-4617.

[68] P.A. Townsend, R.I. Cutress, C.J. Carroll, K.M. Lawrence, T.M. Scarabelli, G. Packham *et al.*, BAG-1 proteins protect cardiac myocytes from simulated ischemia/reperfusion-induced apoptosis via an alternate mechanism of cell survival independent of the proteasome, *J Biol Chem* **279** (20) (2004), pp. 20723-20728.

[69] K. Robergand K. Ollinger, Oxidative stress causes relocation of the lysosomal enzyme cathepsin D with ensuing apoptosis in neonatal rat cardiomyocytes, *Am J Pathol* **152** (5) (1998), pp. 1151-1156.

[70] L. Fioravanti, V. Cappelletti, P. Miodini, E. Ronchi, M. Brivioand G. Di Fronzo, Genistein in the control of breast cancer cell growth: insights into the mechanism of action in vitro, *Cancer Lett* **130** (1-2) (1998), pp. 143-152.

[71] F. Spinozzi, M.C. Pagliacci, G. Migliorati, R. Moraca, F. Grignani, C. Riccardi *et al.*, The natural tyrosine kinase inhibitor genistein produces cell cycle arrest and apoptosis in Jurkat T-leukemia cells, *Leuk Res* **18** (6) (1994), pp. 431-439.

[72] J.-X. Xiao, G.-Q. Huang, X. Gengand H.-W. Qiu, Soy-derived Isoflavones Inhibit HeLa Cell Growth by Inducing Apoptosis, *Plant Foods for Human Nutrition (Dordrecht, Netherlands)* pp.

[73] T.A. Mansoor, R.M. Ramalho, X. Luo, C. Ramalhete, C.M.P. Rodriguesand M.-J.U. Ferreira, Isoflavones as Apoptosis Inducers in Human Hepatoma HuH-7 Cells, *Phytotherapy Research: PTR* pp.

[74] Y. Heand R. Smith, Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B, *Cell Mol Life Sci* **66** (7) (2009), pp. 1239-1256.

[75] B. Honore, U. Baandrupand H. Vorum, Heterogeneous nuclear ribonucleoproteins F

and H/H' show differential expression in normal and selected cancer tissues, *Exp Cell Res* **294** (1) (2004), pp. 199-209.

[76] D. Zhang, Y.C. Tai, C.H. Wong, L.K. Tai, E.S. Koayand C.S. Chen, Molecular response of leukemia HL-60 cells to genistein treatment, a proteomics study, *Leuk Res* **31** (1) (2007), pp. 75-82.

[77] J.K. Seong, D.K. Kim, K.H. Choi, S.H. Oh, K.S. Kim, S.S. Lee *et al.*, Proteomic analysis of the cellular proteins induced by adaptive concentrations of hydrogen peroxide in human U937 cells, *Exp Mol Med* **34** (5) (2002), pp. 374-378.

[78] B. He, Q. Feng, A. Mukherjee, D.M. Lonard, F.J. DeMayo, B.S. Katzenellenbogen *et al.*, A repressive role for prohibitin in estrogen signaling, *Mol Endocrinol* **22** (2) (2008), pp. 344-360.

[79] S.C. Gamble, D. Chotai, M. Odontiadis, D.A. Dart, G.N. Brooke, S.M. Powell *et al.*, Prohibitin, a protein downregulated by androgens, represses androgen receptor activity, *Oncogene* **26** (12) (2007), pp. 1757-1768.

[80] J.A. Vendrell, F. Magnino, E. Danis, M.J. Duchesne, S. Pinloche, M. Pons *et al.*, Estrogen regulation in human breast cancer cells of new downstream gene targets involved in estrogen metabolism, cell proliferation and cell transformation, *J Mol Endocrinol* **32** (2) (2004), pp. 397-414.

[81] P.T. Szymanski, Calponin (CaP) as a latch-bridge protein--a new concept in regulation of contractility in smooth muscles, *J Muscle Res Cell Motil* **25** (1) (2004), pp. 7-19.

[82] E.J. Lowenstein, R.J. Daly, A.G. Batzer, W. Li, B. Margolis, R. Lammers *et al.*, The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras

signaling, *Cell* **70** (3) (1992), pp. 431-442.

CHAPTER 4

QUANTITATIVE PROTEOMIC PROFILING OF ESTROGEN RECEPTOR-DEPENDENT/INDEPENDENT TARGETS OF GENISTEIN USING ISOBARIC TAGS

1. Introduction

Hormone replacement therapy or HRT has been adopted to lower the incidence of cardiovascular disease (CVD) among other benefits for postmenopausal women [1-4]. However, HRT is also controversial due to potential adverse effects such as increasing risk in breast cancer [5, 6]. Therefore, phytoestrogen-based approaches have been advocated as a potential alternative to HRT [7-9]. Such speculation was further augmented by epidemiological evidence that the intake of phytoestrogens, in particular the soy-derived isoflavone genistein (GEN), decreased CVD risk [10-14]. The phytoestrogens or isoflavones is a group of compounds bearing weak estrogenic effects as well as anti-cancer properties. Among the bioactive compounds present in soy products, GEN is the most abundant phytoestrogen. Several laboratory studies have suggested that GEN can provide direct protection to against damaging stresses in cardiac myocytes (CM) [15-20].

The estrogenic attributes of GEN have been emphasized as the main explanation for its cardioprotective role. Genistein at physiological relevant concentrations (1-10 μM) has been shown to bind to both estrogen receptor (ER) subtypes [21-23]. However

little is known on which group of proteins or pathways are the downstream effectors of GEN via ER-dependent mechanism and whether those proteins are related to the cardioprotective role of GEN.

It is also noteworthy that GEN at concentrations greater than 10 μM can also trigger ER-independent effects, in particular non-specific inhibition of protein tyrosine kinases (PTK). The PTK inhibitory effects of GEN have primarily been explored as experimental methodology to block key PTK activation that mediate the cardioprotection signaling triggered by ischemic preconditioning procedure [24-28]. Therefore, it is possible that ER-independent PTK inhibition contributes to the biological effects of GEN provide cardioprotection in a dose-dependent biphasic fashion as high-dose GEN is detrimental to CM survival.

Here we introduce a high throughput quantitative shotgun proteomic approach based on isobaric tag multiplexing technology (iTRAQ) to identify molecular end effectors of GEN co-incubated with or without ER antagonist ICI 182,780 on primary cardiac myocytes isolated from adult male SD rat. Using this approach, 509 unique proteins were identified using a consensus/iterative database searching strategy. By contrasting the proteome alteration in these two conditions relative to the control samples, we identify cellular function changes by ER-dependent or ER-independent action of GEN which can be further decipher the underlying molecular basis of GEN cardioprotection.

2. Materials and Methods

2.1. Isolation of rat primary cardiomyocytes

All laboratory animal protocols were approved by the CSU Institutional Animal Care and Use Committee in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals*. Primary cardiomyocytes were isolated from male Sprague-Dawley rats weighing 300-350 g (n=4). Hearts were surgically removed according to the protocol described previously (Hamilton, Gupta, Knowlton, JMCC, 2004). Excised hearts were subject to retrograde-perfusion of Joklik solution (Sigma, MO) containing 0.8 mg/mL type II collagenase (Worthington Biochemical, NJ). Following tissue digestion, the cell suspension was filtered and pelleted in a solution containing 10% bovine serum albumin (BSA, Sigma, MO), 111 mM NaCl, 5mM KCl, 10 mM NaH_2PO_4, 1 mM $MgSO_4$, 50 uM $CaCl_2$, 5mM HEPES, 60 mM taurine and 20 mM creatine. The pellet was then resuspended in Joklik solution with 6% BSA and $CaCl_2$ was gradually reintroduced to final concentration of 1 mM. After incubation with $CaCl_2$ for 20 mins, cells were collected by centrifugation (~300 rpm for 3 min) and resuspended in M119 medium supplemented with 10% fetal bovine serum (Sigma, MO), 10 mM butanedione monoxime (Sigma, MO), 2 mM glutamine (Invitrogen, CA), 100 U penicillin and 100 mg/ml streptomycin (Invitrogen, CA). Cells were then plated on 0.01% laminin-coated dishes followed by incubation at 37°C, 5%CO_2 for 4 hrs.

2.2. Treatment protocols

Myocytes from each animal were divided into 3 separate groups for treatments. Treatment experiments were performed by replenishing cells with fresh supplemented M119 medium containing 10 μM genistein without (GEN) or with 10 μM ICI 182,780, a full estrogen receptor antagonist (GEN+ERA). Control (CON) cells were treated with DMSO vehicle only. Myocytes were then incubated in at 37°C with 5%CO_2 for 12 hrs.

2.3. Protein extraction

Following treatments, cells were gently washed by 5 mL PBS twice each for 10 sec. Lysis solution (1 ml) containing 10 mM Tris-HCL (pH 8.0), 7 M urea, 2 M thiourea, 1X protease inhibitor cocktail (Sigma, MO), 5 mM magnesium acetate and 1% CHAPS were then directly applied to the cells for quick protein extraction. Protein solutions were then collected in falcon tubes subjected to ultrasonication for 5 min in ice bath. Protein samples were centrifuged at 17 000 x g for 15 min and the supernatant was collected. Total protein concentration for each sample was determined using the RC-DC protein assay kit (Bio-Rad, Hercules, CA, USA) according to manufacturer's instruction. Bovine serum albumin (Sigma) was used as protein assay standard.

2.4. iTRAQ multiplexing and HILIC separation

One-hundred microgram of protein from each sample was directly reduced by 10 mM dithiothreitol for 1 h in lysis buffer and alkylated by 40 mM iodoacetamide for 30

min in dark. Alkylation was quenched by adding dithiothreitol to final concentration of 20 mM. Protein was then precipitated by 2DE Ready Prep clean up kits (Bio-Rad) according to manufacturer's protocol. The resulting protein pellet was reconstituted by 20 μL 50 mM triethyl ammonium bicarbonate supplemented by 4 μL 1% ProteaseMAX. Once the pellets was dissolved completely by votexing, 110 μl 50 mM triethyl ammonium bicarbonate and 5 μg trypsin (Promega) was added to the sample. Digestion was carried out in 37 ºC for 5h and stopped by adding 10 μL 10% trifluoroacetic acid. Tryptic peptides were cleaned up by 200 mg C18 Sep-Pak SPE column (Waters) and then dried by Speed Vac (Thermo Electron). Peptides were then resuspended by 20 μL iTRAQ dissolution buffer (0.5 M triethyl ammonium bicarbonate) and labeled according to manufacturer's protocol (Applied Biosystems). Two separate 8plex iTRAQ sets were used to accommodate samples from 4 biological replicates each with 3 treatments using the design summarized in Fig 4.1. Briefly, all treatments from rat 1 and 3 were arranged in the run A and all samples from rat 2 and 4 were arranged in the run B with randomized order within each run. Two common channels (113 and 119) were used in each run as run-to-run referent channel. Referent sample was constructed by pooling CON and GEN samples from rat 1 and 4. Following 2 h incubation with iTRAQ tags, all samples within each run were then pooled and concentrated by Speed Vac to approximately 30 μL. Samples were then diluted by 30 μL 100 mM ammonium formate at pH3 and then 340 μL of acetonitrile (ACN) added incrementally with votexing. For peptide fractionation, 4 injections each with 100 μL of 8-plexed sample was injected into a PolyHYDROXYETHYL A^{TM} column (200 x 4.6 mm, 5 μm, 200 Å; PolyLC, MD) for

separation under hydrophilic interaction chromatography (HILIC) mode with flow rate at 0.5 mL/min. The peptides were resolved by a gradient first maintained at 100% buffer B (85% ACN, 15 mM ammonium formate, pH 3) for 5 min and then to 10% buffer A (15 mM ammonium formate, pH 3) in 2 min, then 40% A in 48 min and to 100 % A in 3 min and holding at 100% A for 1 min, and finishing up by ramp back to 100% B in 2 min. Fractions were collected every 2 min for the middle 48 min separation period and the starting and stripping eluents were pooled to the final fraction. All fractions were dried and re-dissolved in 10 μL of 3% ACN, 0.1% formic acid before LC-MS/MS analysis.

2.5. Protein identification by ESI-Q-TOF MS/MS

Each HILIC fraction was injected in triplicates onto the G4240-62010 LC-chip (Agilent) hyphenated with a 6150 hybrid ESI-Q-TOF (Agilent). Nano-LC was performed with a 43 min gradient flowing at 500 nL/min starting with 97% solvent A (0.1% formic acid) and 3% solvent B (0.1% formic acid in 90% ACN) to 25% B over 2 min, from 25 to 30% B over the next 13 min, and to 40% B over 21 min, and to 100% B over the next 3 min and maintained for 1 min, and finishing up by drop back to 0%. Mass spectrometric analysis was performed on under the positive mode with nanoelectrospray generated at 2.1 kV. The m/z response of the instrument was calibrated regularly with standards from manufacturer. Precursor scan and product ion spectra were acquired in centroid mode using data dependent data acquisition in MassHunter (Agilent) with the following parameters: mass ranges for MS and MS/MS were m/z 250–2400 and 59–3000, respectively. Every second, a TOF MS

spectrum was scanned, followed by maximum five product ion spectra. The switching from TOF-MS to MS/MS is triggered by precursors with ion intensity >1000 counts with dynamic exclusion for 30 sec. The collision energy was set with 5 V/100 Da slope offset with 2.5 V.

2.6. Spectrum analysis and iTRAQ quantification

The data analysis workflow combined consensus-iterative search strategy with iTRAQ quantitation is summarized in Fig 4.2. Briefly, peak lists in mgf format were generated from raw .d files by Mascot Distiller 2.3.2 (Matrix Science) for peak deisotope and charge state determination. Peak lists were then submit to Mascot v2.3 (Matrix Science) at local server in CSU and X!Tandem Cyclone v 2010.06.01.6 for consensus search against a target-reverse concatenated International Protein Index (IPI, European Bioinformatics Institute, http://www.ebi.ac.uk/IPI) rat protein database (v3.70, 79158 sequences). The search parameters were set to allow for up to two missed cleavages, carboxyamidomethylation on cysteine, iTRAQ 8plex tag on N-terminus and lysine as fixed modifications, oxidation on methionine, iTRAQ 8plex tag on tyrosine as variable modifications, a MS mass tolerance of 100 ppm and MS/MS mass tolerance of 0.05 Da on monoisotopic mode. No iTRAQ quantitation was carried out in this searching step. All identifications from Mascot and X!Tandem were compiled by Scaffold v3.00.03 (Proteome Software, OR) for peptide and protein identification probability calculation [29]. All spectra with good quality but fail to pass a preliminary probability filter with 90% peptide probability and 20% protein probability were exported for 2nd round of iterative search. The iterative search was

conducted by Mascot and X!Tandem adding more dynamic parameters: three missed cleavages, deamidation on asparagine and glutamine, methylation on aspartic acid and glutamic acid as variable modifications, a wider MS mass tolerance of 300 ppm. In X!Tandem iterative search, additional variable modifications such as acetylation on lysine, oxidation on tryptophan, dioxidation on methionine and tryptophan were also included. Iterative search results were then merged with previous search results using Scaffold and peptide and protein probability was recalculated. Finally consensus-iterative search results from both run A and B were merged in Scaffold. Using the same filter, an excel spectrum report consist of good peptide identification results were exported. Scan with ambiguous/multiple protein hits was excluded. Proteins with only one peptide hit across both run A and B were discarded from the sequential quantitative analysis. Following the scan filtering, false discover rate (FDR) was calculated dividing the number of false hits by the number of all hits as summarized by table 4.1:

All target hits	= all hits above the filtering criteria
False positives (FP)	= Decoy hits
True positives (TP)	= All target hits - FP
FDR	= FP/(FP+TP)

An in-house Fortran 90 script was written to extract raw iTRAQ intensity from the original mgf peak list. Briefly, the script searches for peak iTRAQ 8plex tag peaks (113.1, 114.1, 115.1, 116.1, 117.1, 118.1, 119.1, 121.1) within each MS/MS scan with the following two criteria whichever is more stringent: peak m/z falls into the tag

mass ± 0.02 Da mass window or m/z shows the minimum mass shift from the tags. Once the tag peaks are detected and the tag intensities are extracted from a MS/MS scan, the script will add 8plex intensities to corresponding scan in the spectrum report by matching the unique scan ID from the mgf. Quantitative analysis was done in Excel according the following steps: relative tag intensity was generated by normalizing the tag raw intensity to the reference channel in each spectrum; relative intensities in the form of 11x/reference ratio in log2 scale were normalized again by force the medium value to 0; log2(11x/reference) from all scans that belongs to the same protein ID were averaged to generate relative protein intensity in log2 scale within each run; log2(protein intensity) for all 3 treatments from both run A and B was then used in a paired *t*-test to select protein ID with significant ($p < 0.05$) changes.

2.7. Pathway analysis

To identify the potential biological significance of the quantitative proteomics results, lists of differentially expressed proteins as the results of ER-dependent and –independent action of GEN were analyzed by Ingenuity Pathway Analysis (IPA, Ingenuity Systems, Inc.). The IPA database was constructed by abstracting, interconnecting information on functions of gene/protein/compound, interaction of biological molecules, and disease inter-relationships from a large fraction of biomedical literature covering various species including rodents. Once the list of GEN altered proteins via ER-dependent and –independent action were uploaded to IPA, the most relevant pre-built canonical biological pathways was enriched based on

its proprietary database. A *p*-value was computed based on the hypergeometric distribution probability test. The *p*-value represents the odds of having a given pre-built pathways be associated with the list of experimentally- identified proteins simply by chance. The *p*-value is calculated using the formula:

$$p\text{-}Value = \frac{R!\,n!\,(N-R)!\,(N-n)!}{N!} \sum_{i=\max(r,R+n-N)}^{\min(n,R)} \frac{1}{i!\,(R-i)!\,(n-i)!\,(N-R-n+i)!}$$

where N is the total number of nodes, i.e. all rat genes in this case, in IPA database, R is the number of network objects corresponding to the list of differentially expressed proteins identified from the experiment, n is the total number of nodes in each small network that associate with the list, r is the number of proteins from the experimental data that associate with each small network generated from the list.

3. Results

3.1. Protein identification

Before combining the results, Mascot and X!Tandem analyses identified 365 and 521 peptides, respectively, with 321 common protein IDs (data not shown). When compiling the peptide IDs from both search results using Scaffold, single peptide hits from either search engine get chance to pair up with single peptide hits from the other search engine which add 16 marginal gains on protein IDs. The iterative search by both engines with loose search parameter gave 110 additional protein identifications. To this point, the consensus-iterative search strategy has shown superior sensitivity to widen the proteome coverage by 691 protein identifications from target database,

which 89.3% and 32.6% more protein IDs than using Mascot and X!Tandem alone (Fig 4.3A). It is a known issue that consensus search may introduce high error rate as the total decoy hits are very likely the sum of those from each individual search algorithms [30]. In order to effectively estimate the combined FDR for this consensus-iterative search protocol, we constructed a target-reverse concatenated protein database so that each search engine can search the same decoy database to pool all false hits by the end. Under the current identification criteria, this searching protocol provided overall excellent accuracy that the final FDR of run A and B were 1.43% and 2.26%, respectively, as summarized by table 4.2. At this point, 568 and 632 unique protein identification were generated from run A and B, from which 509 common IDs were selected for iTRAQ quantitation (Fig 4.3B).

3.2. iTRAQ quantitation

Raw iTRAQ label intensities were extracted directly from the spectra for all 509 proteins. An example of MS/MS spectrum with iTRAQ report ion regions was given in Fig 4.4. All spectrum been used in the quantitation of those proteins were summarized in Appendix IV. Overall, label intensities shown consistency between all channels across two iTRAQ runs as shown by Fig 4.5A. Instrumental variance between run A and B was corrected by normalizing the intensity of each channel (11X) against the referent 119 channel in run A and 113 channel in run B. The relative intensity of all channels has shown no data skewing towards over-expression or under-expression in meeting the assumption that the majority of proteins remain unchanged. Small labeling variance was then corrected by forcing the medium value

of log2 relative intensity to be 0 as shown in Fig 4.5B. An average log2 relative intensity of all spectra was used for protein quantitation. Poor dynamic range of iTRAQ quantitation was reported, and concern has been raised that iTRAQ quantitation may underestimate the true biological variance between samples [31, 32]. In this experiment, the abundance differences of majority of 509 proteins between any groups fell within ±2 fold-change window (Fig 4.6) including those been deemed as significantly changed proteins. Using pair wised *t*-test, 14 and 15 proteins was found differentially expressed in GEN and GEN+ERA treated groups, respectively, compare to the vehicle group. For those differentially expressed proteins, we also performed the pair wised *t*-test between GEN and GEN+ERA treated group to further evaluate if the protein expression alteration is solely triggered by ER -dependent or – independent mechanism. Additionally, expression pattern was analyzed to further divide GEN responsive gene into 5 groups:

- ***ER-dependent only***

  Protein expression changes in the GEN treated group but such change was completely abolished in the ERA+GEN treated group; comparing to the ERA+GEN group, the GEN treated group shown change with the same magnitude as comparing to the control group.

- ***ER-dependent, possibly with ER-independent in synergism***

  Protein expression changes in the GEN treated group but such change was abolished in the ERA+GEN treated group; however, statistical analysis shown no significant difference between ERA+GEN and GEN treated groups.

- ***ER-independent only***

GEN and GEN+ERA cause changes with same direction and same magnitude.

- ***ER-independent and ER-dependent in antagonism***

  GEN alone result in no chances, but the addition of ERA cause changes; comparing to the GEN group, the GEN+ERA treated group shown change with the same magnitude as comparing to control group.

- ***ER-independent, possibly with ER-dependent in antagonism***

  GEN alone result in no chances, but the addition of ERA cause changes; however, statistical analysis shown no significant difference between ERA+GEN and GEN treated groups.

For each protein identification, protein name, IPI ID, molecular weight, number of unique peptide and number of spectrum used for quantitation, associated biological function GO terms and the fold-changes of GEN, GEN+ERA versus control, together with the fold-changes of GEN versus GEN treated group were summarized in Appendix V. MS/MS spectra and identification details for each spot can be found in supporting information.

The IPA pathway analysis reports 5 top canonical pathways that were significantly associated with the gene differentially changes by GEN via ER-dependent and –independent mechanisms as summarized by table 4.3. Interesting, both groups of genes hit glucose and fatty acid metabolism as the most significantly relevant IPA pathways as will be discussed later in the paper.

4. Discussion

4.1. A novel proteomic workflow using consensus-iterative searching strategy with iTRAQ quantitation

In this study, we developed a shotgun proteomics based on HILIC-RP 2DLC separation coupled with high resolution Q-tof MS/MS and iTRAQ multiplexing technology to quantify protein expression changes in primary rat CM. By constructing a reference sample to be used in the common normalization channel in each iTRAQ experiment, this protocol can be extended to accommodate more samples into multiple parallel iTRAQ experiments for meta-analysis. The use of HILIC separation for peptide fractionation has been advocated for its high orthogonality to reverse-phase LC and higher separation resolution compare to ion-exchange based LC [33]. However, very few studies have tested the applicability of HILIC separation for iTRAQ-labeled peptides. Peptides with iTRAQ tags are generally more hydrophobic than their underivatized peptide counterparts especially with 8plex iTRAQ tags. Therefore, the HILIC gradient start with 85% ACN and the majority of iTRAQ-tagged peptides were separated in a long 85%-50% ACN gradient over 55min to ensure thorough fractionation. Ammonium formate instead of trifluoroacetic acid was used to prevent making peptides more hydrophobic and enhance peptide retention in the HILIC column. Therefore, peptides were identified from all middle 23 fractions (data not shown) suggesting a thorough separation using the current HILIC protocol.

The use of the consensus-iterative database searching strategy further helped to

maximize the identification sensitivity for better proteome coverage [34, 35]. The number of peptide-spectrum matches (PSM) using Mascot alone was 37274 and 38760 run A and B, respectively, while using or X!tandem alone gave 57112 and 58343 PSMs in run A and B. When using consensus-iterative search strategy, total number of PSMs was 60741 and 63443 in run A and B respectively. The additional PSMs enhance the sensitivity by adding more low abundant protein IDs as most of them are based on 2 unique peptides with low number of PSM. However, there has no direct application of consensus search strategy with protein quantitation such as iTRAQ. This is largely due to the fact that different search engines may or may not support iTRAQ data, even they do, they used different data process routine for iTRAQ quantitation. Thus, in our consensus searching procedure, a stand-alone Fortran 90 script was introduced to circumvent the problem by directly extract the iTRAQ report ion intensities from the raw MS/MS spectra and combine them to the final identification report. Moreover, this procedure generates no negative value as in the case of iTRAQ quantitation in Mascot and also gives the user full control of the downstream data analysis.

Despite the improved sensitivity and overall large number of PSMs assigned, the number of unique peptide and protein IDs from both run A and B are lower than we expected. This is probably due to the presence of high abundance proteins from the myocytes. Peptides from cytoskeleton and myofibril proteins occupied a large portion of the assigned PSMs. For example myosin-6 alone generates 9578 and 8589 PSMs in run A and B, respectively. The PSMs from the top 10 most abundant proteins in run A

and B were assigned by 18.58% and 15.17% of total spectra while PSMs from all protein IDs from run A and B occupied 44.99% and 43.33% of total spectra, respectively.

4.2. ER-dependent proteome alteration

The estrogenic properties of GEN have been postulated as a contributing factor for its cardioprotective role. However, which end effectors are targeted by GEN in an ER-dependent fashion is still poorly understood. Here we report two glycolysis proteins the fructose-biphosphate aldolase A and α-enolase were up-regulated in GEN treated group but not in the GEN+ERA treated group, possibly resulted as the ER-dependent effects of GEN. It is long being known that heart depends on glycolysis for energy production to survival under ischemic conditions, thus enhanced glycolysis provides critical cytoprotection for CM against ischemic-reperfusion insults [36, 37]. Therefore, the argument of glucose metabolism by estrogen treatment has been contribute to the direct cardioprotective efficacy of estrogen [38]. This evidence strongly support that the cardioprotective effects of GEN, particularly in the ischemic heart, may also be partly attributed to the ER-dependent up-regulation of glucose metabolism.

We also found that two mitochondrial membrane proteins involved in fatty acid metabolism were affected by GEN. The translocation of fatty acyl-CoA across the mitochondrial membrane for β-oxidation requires its conjugation with carnitine. Carnitine palmitoyltransferase 2 (Carnitine O-palmitoyltransferase2, mitochondrial, Cpt2), a critical enzyme for conversion of acyl-carnitine to acyl-CoA, was up-

regulated by 1.08 fold. In contrast, mitochondrial carnitine/acylarnitine carrier protein (carnitine/acylcarnitine translocase, solute carrier family 25 member 20, Slc25a20) which transport acylcarnitine across the membrane into mitochondria and shuttle carnitine back to cytosol was down-regulated 1.21 fold by GEN treatment. Notably, it is well documented that transcription of fatty acid metabolism related genes like Cpt2 and Slc25a20 can be activated by peroxisome proliferator-activated receptors α (PPARα) [39-42]. Further, genistein and other isoflavones have been shown to activate PPARs in bone cells [43] and macrophages [44]. It has been suggested that the activation of PPARs contributes to the antiestrogenic effects of GEN [45]. Given that fact that estrogen generally promotes glycolysis while activation of PPARs generally promote fatty acid metabolism and inhibit the glucose metabolism, the non-uniform expression patterns of energy metabolism genes after GEN treatment shown in our data lead us to speculate that GEN triggers the concurrent modulation of energy utilization in rat CM via both PPARα and ER-dependent pathways.

In addition to proteins involved in energy metabolism, we also found two proteins that may relate to cardioprotection were modulated by GEN in an ER-dependent fashion. Soluble epoxide hydrolase (sEH) was down-regulated by GEN without ERA co-incubation. Soluble epoxide hydrolase metabolizes the cardioprotective epoxyeicosatrienoic acids (EET) which derived from endogenous arachidonic acid by cytochrome P450 [46, 47]. Deficiency of sEH and pharmacological inhibition of sEH elicit cardiac resistance to ischemia via EET-mediated STAT3 signaling *in vitro* and *in vivo* models [48, 49]. The down-regulation of sEH by GEN potentially identifies a

novel target of GEN for cardioprotection. However, another protein T-complex protein 1 (TCP-1) subunit delta was significantly down-regulation in the GEN treated group. The TCP-1 is one of the eight subunits forming cytosolic chaperone heterooligomeric complex and specifically involves with actin and tubulin polymerization, which suggest its importance for the maintenance of cytoskeleton integrity.

4.3. ER-independent genes

Our analysis of proteomic expression pattern also reveals a list of proteins that are regulated by GEN via ER-independent routes. However, some proteins were also regulated by GEN treatment alone indicating the coexistence of both ER-dependent and ER-independent effects of GEN.

Genistein has been reported to modulate steroidogenesis in both an inhibitory [50, 51] and stimulatory [52] fashion. Genistein is known to inhibit the activity of aromatase [53], which converts androgens to estrogens [54, 55]. Additionally, GEN was reported to inhibit the expression of hydroxysteroid dehydrogenases (HSDs) which convert estradiol to the less active estrone. Impairment of HSD gene expression has been reported as the result GEN treatment in porcine granulosa cells [56, 57]. Our data show a 1.05-fold decrease in 3-hydroxyacyl-CoA dehydrogenase type-2 (or 17-beta-hydroxysteroid dehydrogenase 10, HSD10) by GEN in the presence of ERA, suggesting an ER-independent mechanism. Interestingly, PTK inhibition has been shown to increase estradiol production in granulosa cells [52]. Therefore, we

hypothesize that the inhibition of HSD10 by GEN may result from the PTK inhibitory property of GEN. Moreover, our data suggest there are also ER-dependent effects that abolished the down-regulation of HSD 10 when CM are treated with GEN alone, suggesting GEN may trigger an ER-dependent action that antagonizes TK inhibitory action. Although the implication of these changes in steroidogenesis in CM is still unknown, but these data shed light on the complexity GEN induced changes in steroid metabolism.

Consistent with the notion that GEN may impact cell signaling in ways similar to estrogen but independent of ER binding, we found GEN down-regulated the estrogen signaling protein, MACRO domain-containing protein 1 (or leukemia related protein 16, LRP16). LRP16 has been shown to positively regulate estrogen signaling in cancer cells promoting cell growth [58]. Here we report an approximate 1.5 fold decrease in LRP16 with GEN regardless of whether an ER antagonist was present suggesting that the down-regulation of LRP16 in CM by GEN is via an ER independent mechanism. Interestingly, up-regulation of LRP16 by 17beta-estradiol through activation of estrogen receptors was reported in human cancer cell lines [58-60] and LRP16 has been shown take a role in a positive-feedback regulatory loop for estrogen signaling resulting in invasive growth of cancer cell [58]. Therefore, our results suggest LRP16 is a potential novel target for GEN to exert biological activity with importance in both CM and cancer cells.

Small molecular chaperones were identified as important cytoprotective proteins that

regulate protein folding and apoptosis pathways [61]. However, we report that heat shock protein 75 kDa (HSP75, HSP90L, Tumor necrosis factor type 1 receptor-associated protein, TNFR-associated protein 1, TRAP1) was down-regulated by GEN+ERA but not by GEN alone. HSP90 class has been shown to be inducible by heat and I/R insults [62]. This group of chaperones specifically bind to steroid receptor complexes such as estrogen, androgen and progesterone receptors and function as negative regulators of their activity [63]. HSP90 overexpression was shown to afford cytoprotection against heat shock in H9c2 cultured CM [64] and rat primary CM [65] and against ischemic shock in mice heart [66]. HSP75 or TRAP1 is a HSP90 isoform that only found in mitochondria. Xiang et al reported that TRAP1 plays a role in maintaining mitochondrial function and preventing CM apoptosis by reducing the mitochondrial permeability during hypoxia [67]. Masuda et al also suggested that ROS can down-regulated of TRAP1 leading to apoptosis by releasing cytochrome c from mitochondria [68]. Interestingly, HSP90 was documented as an estrogen-responsive protein that can be up-regulated by estrogen in ovarian cancer cell line [69]. Considering our finding that GEN down-regulated the HSP90 ER-independently while GEN alone did not affect the expression level, we suspect that there might be a counteracting ER-dependent action of GEN to up-regulate the HSP90 in CM. However, the biological significance of TRAP1 down-regulation by GEN in CM still needed to be investigated.

5. Concluding Remarks

Genistein and soy isoflavones have been suggested to provide direct protection for CM and have been advocated as an alternative HRT agent in replacement of conventional estrogen. However, the molecular impact of GEN is multifaceted and may be mediated by activation of ERs and/or other molecular targets concurrently. In this study, a high throughput shotgun proteomic investigation coupled with iTRAQ quantitation and advanced bioinformatic tools were applied to dissect the ER-dependent and -independent effects of GEN treatment on CM. The use of a consensus-iterative searching strategy increased the sensitivity for protein identification with a total 691 protein IDs from two parallel iTRAQ experiments. A house-written Fortran 90 script was employed to facilitate the iTRAQ quantitation in this consensus-iterative searching protocol. As the result, 14 and 15 proteins were found differentially expressed from a total of 509 proteins quantified comparing GEN, and GEN+ERA treated CM to untreated controls. In those genes regulated by GEN in an ER-dependent manner, we found glycolysis proteins were up-regulated by GEN. Together with the modulation of mitochondrial fatty acid transportation related proteins, our proteomics data provide critical evidence shown that the GEN may have complex influence on the myocardial energy metabolism. Additionally, we found the soluble epoxide hydrolase (sEH) was down-regulated indicating a novel target of GEN for the cardioprotection. From those gene regulated by GEN potential independent from ER binding, we discovered novel molecular targets in steroidogenesis and estrogen signaling pathways that were down-regulated by GEN. The use of bioinformatics tools like Ingenuity pathway analysis aided in the

interpretation of the proteomics data collected and helped develop an approach to recognize or predict the major pathways affected by GEN treatment. According to IPA analysis, GEN appears to exert a crucial role on glucose and fatty acid metabolism in both ER-dependent and –independent manner. Although the proteomics discovery and resulting hypothesis will require validation and test by further studies, but we hope they serve as an important step stone to the understanding of complex cellular impact induced by GEN in CM.

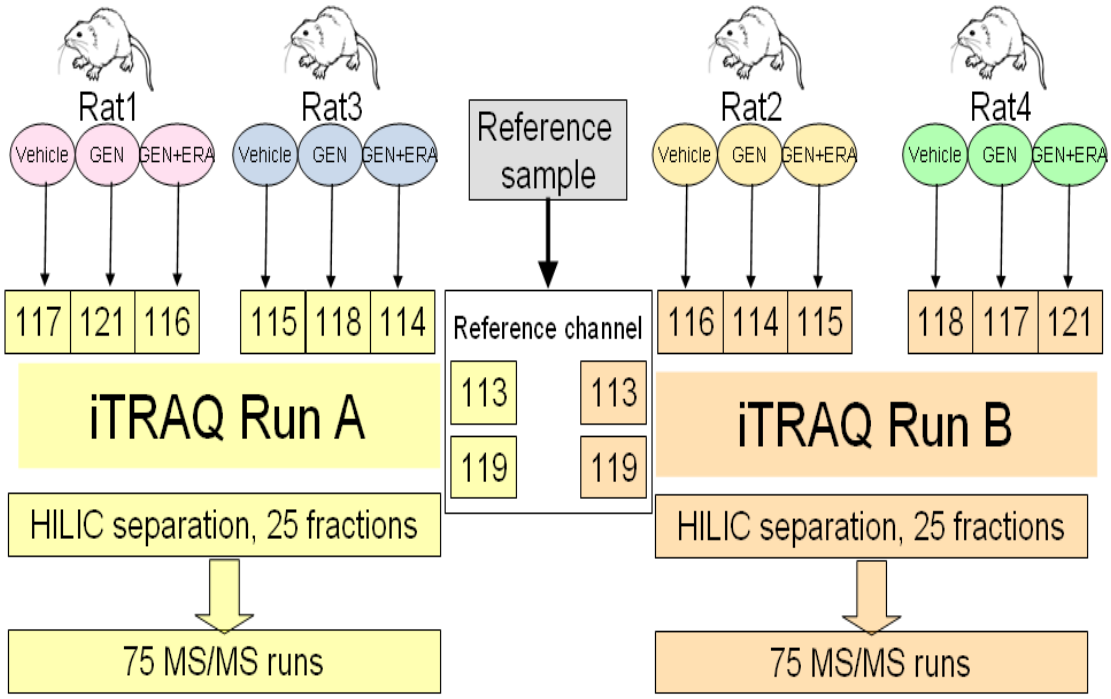


Fig 4.1, Overall iTRAQ LC-MS/MS experiment design. Two separate 8plex iTRAQ sets were used to accommodate samples from 4 biological replicates each with 3 treatments. In each run, pooled iTRAQ-tagged peptides will be fractionated by HILIC into 25 fractions, which will be further analyzed by Q-tof MS/MS in triplicates.

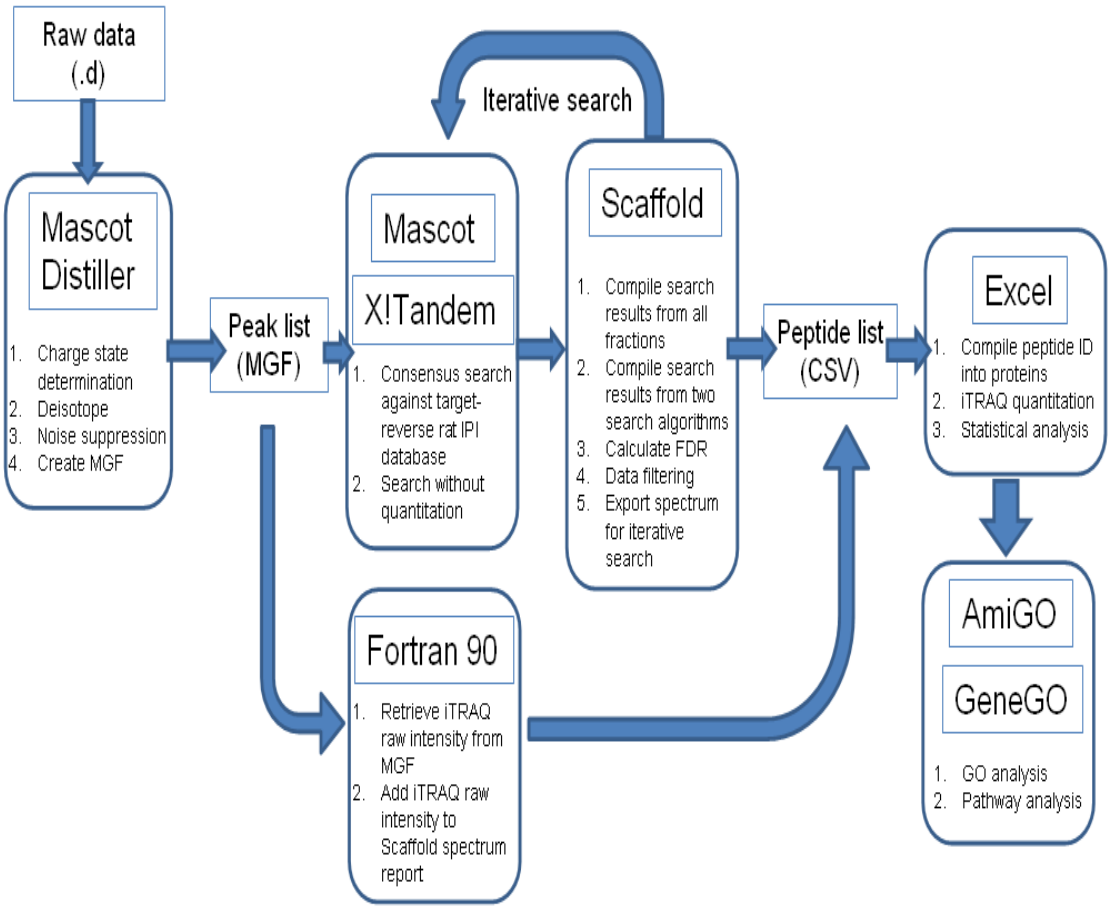
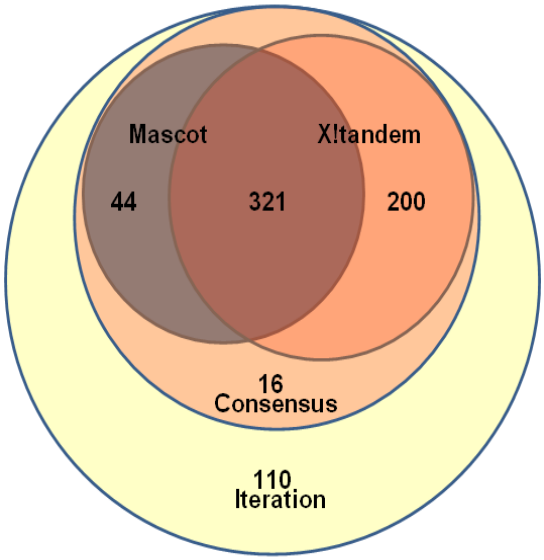


Fig 4.2 Overview of consensus-iterative search strategy. The Scaffold was used for compiling search results from Mascot and X!tandem against a target-reverse concatenated IPI rat database. Scaffold was also used to export good quality spectrum with no matches for iterative search which is later combined with previous search results. A Fortran 90 script was written to extract iTRAQ raw intensity directly from mgf and combined them with the final spectrum report from the consensus-iterative database searching.

A



B

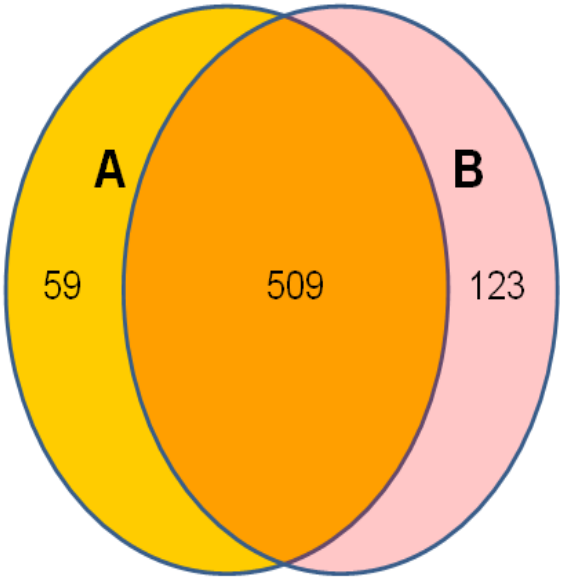
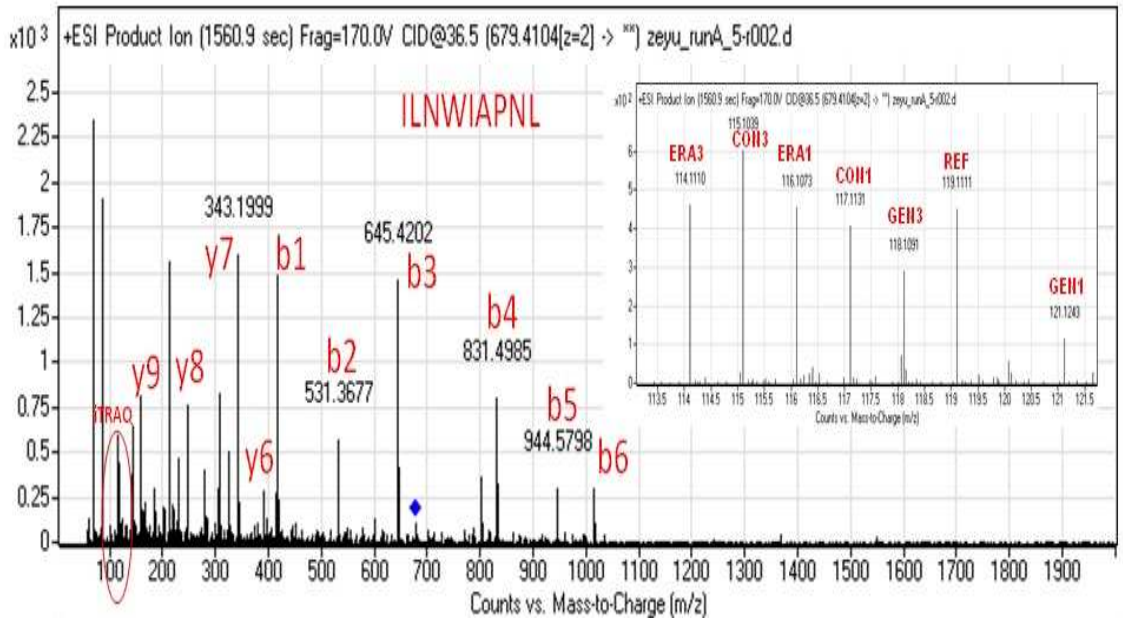


Fig 4.3, Summary of protein identification via the consensus-iterative search strategy. Using 90% peptide probability and 20% protein probability and at least 2 peptides IDs across 2 iTRAQ runs as the criteria for identification, Fig 4.3A shown that totally 691 protein IDs was detected after consensus-iterative search strategy from the targeted database, while Mascot and X!Tandem alone gave 365 and 521 IDs, respectively. Fig 4.3B shown that iTRAQ run A and B results in 568 and 632 unique protein IDs respectively. We choose to do iTRAQ quantitation on the 509 common IDs from both runs for reliable statistics.

A



B

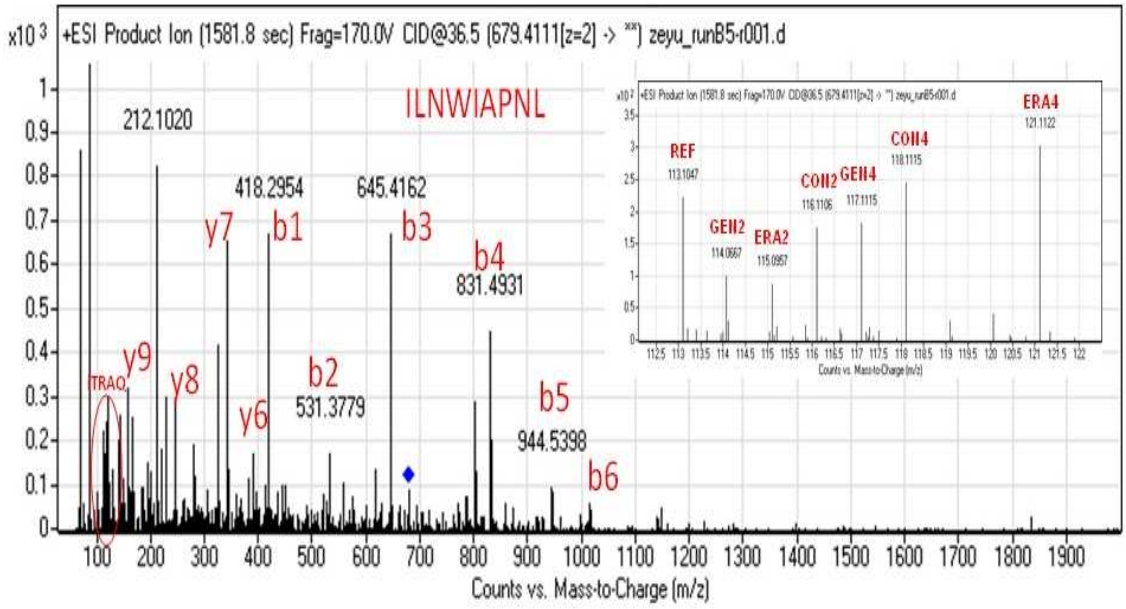
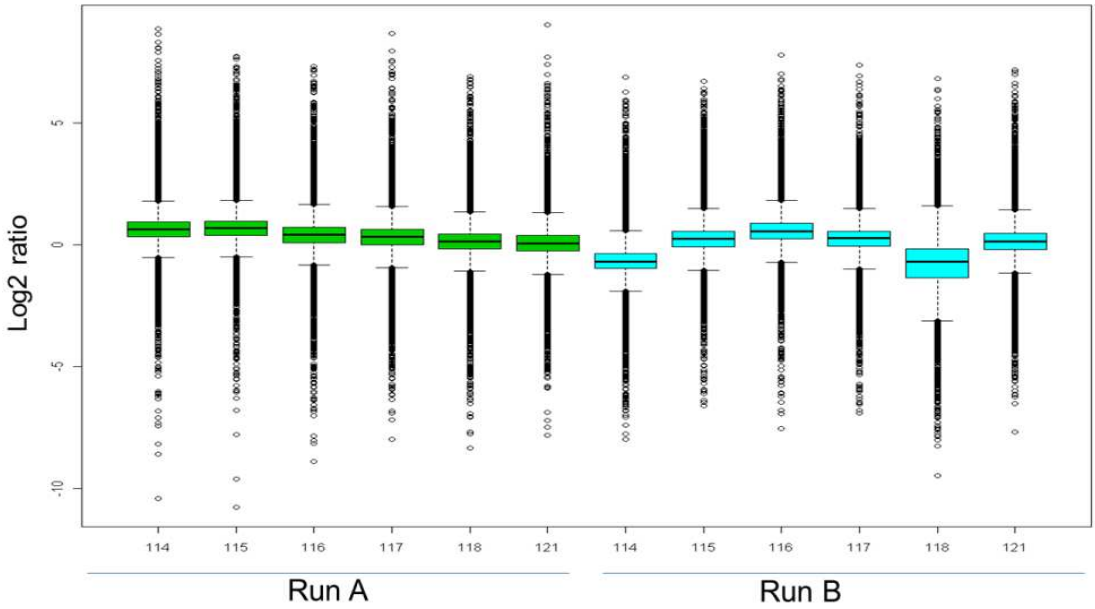


Fig 4.4 Example MS/MS spectra from a mitochondrial carnitine/acylcarnitine carrier protein that down-regulated by GEN treatment. Both spectra from run A and B shown consistent fragmentation pattern which gives the same peptides ID with same series of y and b ions. Close-in view of the iTRAQ report ion region from both MS/MS shown peptide ILNWIAPNL from GEN treated groups are overall down-regulated.

A



B

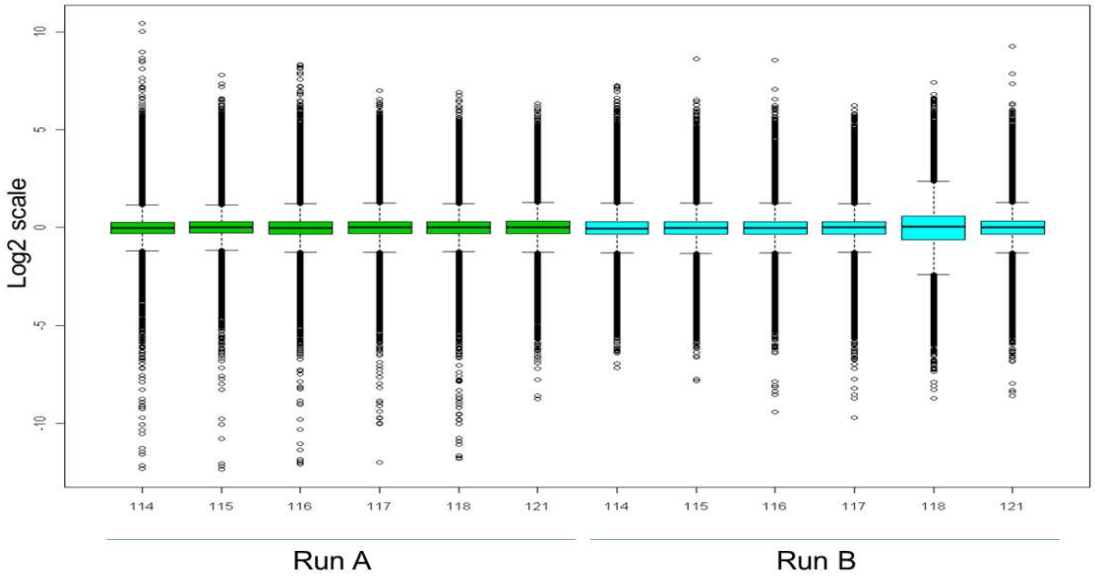


Fig 4.5, Summary of iTRAQ raw data directly extracted from raw MS/MS spectrum. Instrumental variance between run A and B was corrected by normalizing the intensity of each channel against the reference channel. Fig 4.5A shown the relative intensity of 11x/reference in both runs. The log2 relative intensity of all channels shown symmetric distribution around 0. Labeling variance was then corrected by forcing the medium value of log2 relative intensity to be 0 as shown in Fig 4.5B in assumption that the expression of majority of proteins remain unchanged.

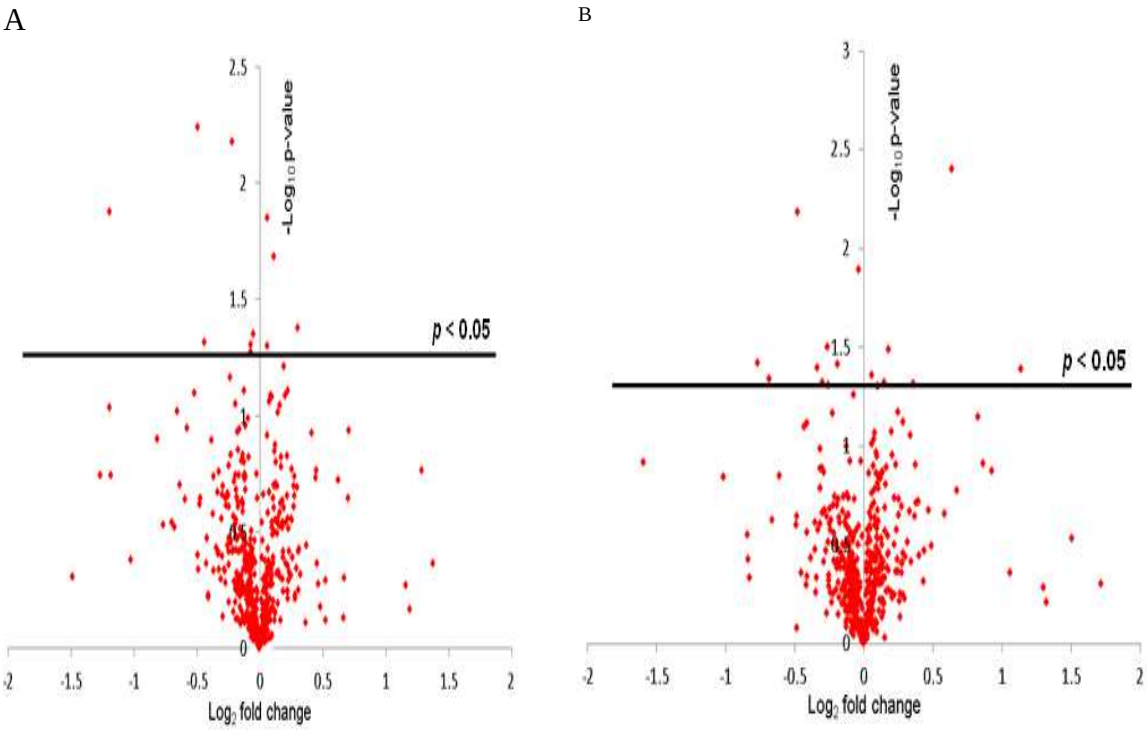


Fig. 4.6 Volcano plot of differential expressed proteins comparing GEN treated group (A) and GEN+ERA treated group (B) verses the control group. The horizontal axis represents the both up and down fold changes in log2 scale while the vertical axis measures the change significance according the p-value from pair wise t-test in –log10 scale. Significant change were qualified by $p < 0.05$.

Table 4.2 Identification Statistics for consensus-iterative database searching strategy. The identification acceptance criteria were set to 90% peptide probability, 20% protein probability and at least 2 independent observations of the same protein from either run A and B.

	Run A	**Run B**
total protein ID	586	646
Target protein ID	568	632
Decoy protein ID	8	14
FDR	1.43%	2.26%
Unique peptide IDs	2478	2583
Peptide-spectrum matches	60741	63443
Total number of search queries	134508	146552
spectrum identification rate	45%	43%
Overall statistics for quantitation		
Protein detected by both runs	511	
Target ID in both runs	509	
Decoy ID in both runs	2	
Overall FDR	0.4%	

Table 4.3. Differentially altered proteins by GEN via ER-dependent and –independent mechanisms in primary rat CM were grouped for GeneGO MetaCore pathway analysis. A hypergeometric distribution probability test was carried out to decide if the experimentally identified proteins (objects) were involved with a particular biological pathway in the GeneGO MetaCore database. The p-value represents the odds of having a given pre-built network be associated with the list of experimentally-identified proteins simply by chance.

Canonical pathway name	**Total objects in pathway**	**Objects differentially expressed**	***p*-value**
ER-dependent alteration			
Glycolysis/gluconeogenesis	134	3	1.97E-05
Fatty Acid Metabolism	184	2	2.46E-03
Inositol Metabolism	18	1	3.72E-03
Phenylalanine, Tyrosine and Tryptophan Biosynthesis	67	1	1.05E-02
Purine Metabolism	391	2	1.11E-02
ER-independent alteration			

Glycolysis/gluconeogenesis	134	2	1.64E-03
Fatty Acid Elongation in Mitochondria	47	1	1.29E-02
Citrate Cycle	57	1	1.96E-02
Bile Acid Biosynthesis	106	1	3.63E-02
Butanoate Metabolism	127	1	3.89E-02

REFERENCES

[1] F. Grodstein, M.J. Stampfer, J.E. Manson, G.A. Colditz, W.C. Willett, B. Rosner *et al.*, Postmenopausal estrogen and progestin use and the risk of cardiovascular disease, *New England Journal of Medicine* **335** (7) (1996), pp. 453-461.

[2] G.M. Rosano, P.M. Sarrel, P.A. Poole-Wilsonand P. Collins, Beneficial effect of oestrogen on exercise-induced myocardial ischaemia in women with coronary artery disease, *Lancet* **342** (8864) (1993), pp. 133-136.

[3] M.J. Stampfer, G.A. Colditz, W.C. Willett, J.E. Manson, B. Rosner, F.E. Speizer *et al.*, Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from the nurses' health study, *N Engl J Med* **325** (11) (1991), pp. 756-762.

[4] M.J. Stampferand G.A. Colditz, Estrogen Replacement Therapy and Coronary Heart-Disease - a Quantitative Assessment of the Epidemiologic Evidence, *Preventive Medicine* **20** (1) (1991), pp. 47-63.

[5] J.E. Rossouw, G.L. Anderson, R.L. Prentice, A.Z. LaCroix, C. Kooperberg, M.L. Stefanick *et al.*, Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial, *JAMA* **288** (3) (2002), pp. 321-333.

[6] G.L. Anderson, M. Limacher, A.R. Assaf, T. Bassford, S.A. Beresford, H. Black *et al.*, Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial, *JAMA* **291** (14) (2004), pp. 1701-1712.

[7] D. Altavilla, A. Crisafulli, H. Marini, M. Esposito, R. D'Anna, F. Corrado *et al.*, Cardiovascular effects of the phytoestrogen genistein, *Current Medicinal Chemistry*

Cardiovascular and Hematological Agents **2** (2) (2004), pp. 179-186.

[8] M. Atteritano, H. Marini, L. Minutoli, F. Polito, A. Bitto, D. Altavilla *et al.*, Effects of the phytoestrogen genistein on some predictors of cardiovascular risk in osteopenic, postmenopausal women: a two-year randomized, double-blind, placebo-controlled study, *J Clin Endocrinol Metab* **92** (8) (2007), pp. 3068-3075.

[9] A. Crisafulli, D. Altavilla, H. Marini, A. Bitto, D. Cucinotta, N. Frisina *et al.*, Effects of the phytoestrogen genistein on cardiovascular risk factors in postmenopausal women, *Menopause* **12** (2) (2005), pp. 186-192.

[10] A. Menotti, A. Keys, H. Blackburn, D. Kromhout, M. Karvonen, A. Nissinen *et al.*, Comparison of multivariate predictive power of major risk factors for coronary heart diseases in different countries: results from eight nations of the Seven Countries Study, 25-year follow-up, *J Cardiovasc Risk* **3** (1) (1996), pp. 69-75.

[11] X. Zhang, X.O. Shu, Y.-T. Gao, G. Yang, Q. Li, H. Li *et al.*, Soy food consumption is associated with lower risk of coronary heart disease in Chinese women, *The Journal of Nutrition* **133** (9) (2003), pp. 2874-2878.

[12] K.D. Setchelland A. Cassidy, Dietary isoflavones: biological effects and relevance to human health, *J Nutr* **129** (3) (1999), pp. 758S-767S.

[13] M.G. Hertog, D. Kromhout, C. Aravanis, H. Blackburn, R. Buzina, F. Fidanza *et al.*, Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study, *Arch Intern Med* **155** (4) (1995), pp. 381-386.

[14] A. Keys, A. Menotti, C. Aravanis, H. Blackburn, B.S. Djordevic, R. Buzina *et al.*, The seven countries study: 2,289 deaths in 15 years, *Prev Med* **13** (2) (1984), pp. 141-154.

[15] R. Tissier, X. Waintraub, N. Couvreur, M. Gervais, P. Bruneval, C. Mandet *et al.*,

Pharmacological postconditioning with the phytoestrogen genistein, *J Mol Cell Cardiol* **42** (1) (2007), pp. 79-87.

[16] P. Zhai, T.E. Eurell, R.P. Cotthaus, E.H. Jeffery, J.M. Bahrand D.R. Gross, Effects of dietary phytoestrogen on global myocardial ischemia-reperfusion injury in isolated female rat hearts, *Am J Physiol Heart Circ Physiol* **281** (3) (2001), pp. H1223-1232.

[17] N. Couvreur, R. Tissier, S. Pons, M. Chenoune, X. Waintraub, A. Berdeaux *et al.*, The ceiling effect of pharmacological postconditioning with the phytoestrogen genistein is reversed by the GSK3beta inhibitor SB 216763 [3-(2,4-dichlorophenyl)-4(1-methyl-1H-indol-3-yl)-1H-pyrrole-2,5-dione] through mitochondrial ATP-dependent potassium channel opening, *J Pharmacol Exp Ther* **329** (3) (2009), pp. 1134-1141.

[18] L. Al-Nakkash, B. Markus, K. Bowden, L.M. Batia, W.C. Prozialeckand T.L. Broderick, Effects of acute and 2-day genistein treatment on cardiac function and ischemic tolerance in ovariectomized rats, *Gend Med* **6** (3) (2009), pp. 488-497.

[19] E. Souzeau, S. Belanger, S. Picardand C.F. Deschepper, Dietary isoflavones during pregnancy and lactation provide cardioprotection to offspring rats in adulthood, *Am J Physiol Heart Circ Physiol* **289** (2) (2005), pp. H715-721.

[20] B. Deodato, D. Altavilla, G. Squadrito, G.M. Campo, M. Arlotta, L. Minutoli *et al.*, Cardioprotection by the phytoestrogen genistein in experimental myocardial ischaemia-reperfusion injury, *Br J Pharmacol* **128** (8) (1999), pp. 1683-1690.

[21] M. Maggiolini, D. Bonofiglio, S. Marsico, M.L. Panno, B. Cenni, D. Picard *et al.*, Estrogen receptor alpha mediates the proliferative but not the cytotoxic dose-dependent effects of two major phytoestrogens on human breast cancer cells, *Mol Pharmacol* **60** (3) (2001), pp. 595-602.

[22] G.G. Kuiper, J.G. Lemmen, B. Carlsson, J.C. Corton, S.H. Safe, P.T. van der Saag *et al.*, Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta, *Endocrinology* **139** (10) (1998), pp. 4252-4263.

[23] M.W. Carter, W.W. Smart, Jr.and G. Matrone, Estimation of estrogenic activity of genistein obtained from soybean meal, *Proc Soc Exp Biol Med* **84** (2) (1953), pp. 506-508.

[24] C.P. Baines, L. Wang, M.V. Cohenand J.M. Downey, Protein tyrosine kinase is downstream of protein kinase C for ischemic preconditioning's anti-infarct effect in the rabbit heart, *J Mol Cell Cardiol* **30** (2) (1998), pp. 383-392.

[25] I.F. Benter, J.S. Juggi, I. Khan, M.H. Yousif, H. Canatanand S. Akhtar, Signal transduction mechanisms involved in cardiac preconditioning: role of Ras-GTPase, Ca2+/calmodulin-dependent protein kinase II and epidermal growth factor receptor, *Mol Cell Biochem* **268** (1-2) (2005), pp. 175-183.

[26] R.M. Fryer, J.E. Schultz, A.K. Hsuand G.J. Gross, Pretreatment with tyrosine kinase inhibitors partially attenuates ischemic preconditioning in rat hearts, *Am J Physiol* **275** (6 Pt 2) (1998), pp. H2009-2015.

[27] R.M. Fryer, J.E. Schultz, A.K. Hsuand G.J. Gross, Importance of PKC and tyrosine kinase in single or multiple cycles of preconditioning in rat hearts, *Am J Physiol* **276** (4 Pt 2) (1999), pp. H1229-1235.

[28] S. Okubo, Y. Tanabe, K. Takeda, M. Kitayama, S. Kanemitsu, R.C. Kukreja *et al.*, Pretreatment with tyrosine kinase inhibitor attenuates the reduction of apoptosis 24 h after ischemic preconditioning, *Jpn J Physiol* **54** (2) (2004), pp. 143-151.

[29] A. Keller, A.I. Nesvizhskii, E. Kolkerand R. Aebersold, Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search,

Anal Chem **74** (20) (2002), pp. 5383-5392.

[30] W. Yu, J.A. Taylor, M.T. Davis, L.E. Bonilla, K.A. Lee, P.L. Auger *et al.*, Maximizing the sensitivity and reliability of peptide identification in large-scale proteomic experiments by harnessing multiple search engines, *Proteomics* **10** (6) pp. 1172-1189.

[31] P.C. Wright, S.Y. Ow, M. Salim, J. Noirel, C. Evansand I. Rehman, iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly", *Journal of Proteome Research* **8** (11) (2009), pp. 5347-5355.

[32] N.A. Karp, W. Huber, P.G. Sadowski, P.D. Charles, S.V. Hesterand K.S. Lilley, Addressing accuracy and precision issues in iTRAQ quantitation, *Mol Cell Proteomics* **9** (9) pp. 1885-1897.

[33] M. Gilar, P. Olivova, A.E. Dalyand J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, *Anal Chem* **77** (19) (2005), pp. 6426-6434.

[34] R.K. Dagda, T. Sultanaand J. Lyons-Weiler, Evaluation of the Consensus of Four Peptide Identification Algorithms for Tandem Mass Spectrometry Based Proteomics, *Journal of proteomics & bioinformatics* **3** pp. 39-47.

[35] T. Sultana, R. Jordanand J. Lyons-Weiler, Optimization of the Use of Consensus Methods for the Detection and Putative Identification of Peptides via Mass Spectrometry Using Protein Standard Mixtures, *Journal of proteomics & bioinformatics* **2** (6) (2009), pp. 262-273.

[36] W.C. Stanley, G.D. Lopaschuk, J.L. Halland J.G. McCormack, Regulation of myocardial carbohydrate metabolism under normal and ischaemic conditions. Potential for pharmacological interventions, *Cardiovasc Res* **33** (2) (1997), pp. 243-257.

[37] G.D. Lopaschukand W.C. Stanley, Glucose metabolism in the ischemic heart, *Circulation* **95** (2) (1997), pp. 313-315.

[38] H. Fraser, S.T. Davidgeand A.S. Clanachan, Enhancement of post-ischemic myocardial function by chronic 17 beta -estradiol treatment: role of alterations in glucose metabolism, *J Mol Cell Cardiol* **31** (8) (1999), pp. 1539-1549.

[39] M. Rakhshandehroo, B. Knoch, M. Mullerand S. Kersten, Peroxisome proliferator-activated receptor alpha target genes, *PPAR Res* **2010** pp.

[40] S. Mandard, M. Mullerand S. Kersten, Peroxisome proliferator-activated receptor alpha target genes, *Cell Mol Life Sci* **61** (4) (2004), pp. 393-416.

[41] M.J. Barrero, N. Camarero, P.F. Marreroand D. Haro, Control of human carnitine palmitoyltransferase II gene transcription by peroxisome proliferator-activated receptor through a partially conserved peroxisome proliferator-responsive element, *Biochem J* **369** (Pt 3) (2003), pp. 721-729.

[42] K. Tachibana, K. Takeuchi, H. Inada, D. Yamasaki, K. Ishimoto, T. Tanaka *et al.*, Regulation of the human SLC25A20 expression by peroxisome proliferator-activated receptor alpha in human hepatoblastoma cells, *Biochem Biophys Res Commun* **389** (3) (2009), pp. 501-505.

[43] Z.C. Dangand C. Lowik, Dose-dependent effects of phytoestrogens on bone, *Trends Endocrinol Metab* **16** (5) (2005), pp. 207-213.

[44] O. Mezei, W.J. Banz, R.W. Steger, M.R. Peluso, T.A. Wintersand N. Shay, Soy isoflavones exert antidiabetic and hypolipidemic effects through the PPAR pathways in obese Zucker rats and murine RAW 264.7 cells, *J Nutr* **133** (5) (2003), pp. 1238-1243.

[45] Z. Dangand C.W. Lowik, The balance between concurrent activation of ERs and

PPARs determines daidzein-induced osteogenesis and adipogenesis, *J Bone Miner Res* **19** (5) (2004), pp. 853-861.

[46] K. Nithipatikom, J.M. Moore, M.A. Isbell, J.R. Falckand G.J. Gross, Epoxyeicosatrienoic acids in cardioprotection: ischemic versus reperfusion injury, *Am J Physiol Heart Circ Physiol* **291** (2) (2006), pp. H537-542.

[47] S.N. Batchu, S.B. Lee, R.S. Qadhi, K.R. Chaudhary, H. El-Sikhry, R. Kodela *et al.*, Cardioprotective effect of a dual acting epoxyeicosatrienoic acid analogue towards ischaemia reperfusion injury, *Br J Pharmacol* **162** (4) pp. 897-907.

[48] A. Motoki, M.J. Merkel, W.H. Packwood, Z. Cao, L. Liu, J. Iliff *et al.*, Soluble epoxide hydrolase inhibition and gene deletion are protective against myocardial ischemia-reperfusion injury in vivo, *Am J Physiol Heart Circ Physiol* **295** (5) (2008), pp. H2128-2134.

[49] M.J. Merkel, L. Liu, Z. Cao, W. Packwood, J. Young, N.J. Alkayed *et al.*, Inhibition of soluble epoxide hydrolase preserves cardiomyocytes: role of STAT3 signaling, *Am J Physiol Heart Circ Physiol* **298** (2) pp. H679-687.

[50] J. Bektic, R. Guggenberger, I.E. Eder, A.E. Pelzer, A.P. Berger, G. Bartsch *et al.*, Molecular effects of the isoflavonoid genistein in prostate cancer, *Clin Prostate Cancer* **4** (2) (2005), pp. 124-129.

[51] D. Haynes-Johnson, M.T. Lai, C. Campenand S. Palmer, Diverse effects of tyrosine kinase inhibitors on follicle-stimulating hormone-stimulated estradiol and progesterone production from rat granulosa cells in serum-containing medium and serum-free medium containing epidermal growth factor, *Biol Reprod* **61** (1) (1999), pp. 147-153.

[52] A. Makarevich, A. Sirotkin, T. Taradajnikand P. Chrenek, Effects of genistein and

lavendustin on reproductive processes in domestic animals in vitro, *J Steroid Biochem Mol Biol* **63** (4-6) (1997), pp. 329-337.

[53] Y.J. Moon, X. Wangand M.E. Morris, Dietary flavonoids: effects on xenobiotic and carcinogen metabolism, *Toxicol In Vitro* **20** (2) (2006), pp. 187-210.

[54] D. Ghosh, J. Griswold, M. Ermanand W. Pangborn, Structural basis for androgen specificity and oestrogen synthesis in human aromatase, *Nature* **457** (7226) (2009), pp. 219-223.

[55] E.R. Simpson, M.S. Mahendroo, G.D. Means, M.W. Kilgore, M.M. Hinshelwood, S. Graham-Lorence *et al.*, Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis, *Endocr Rev* **15** (3) (1994), pp. 342-355.

[56] U. Tiemann, F. Schneider, J. Vanselowand W. Tomek, In vitro exposure of porcine granulosa cells to the phytoestrogens genistein and daidzein: effects on the biosynthesis of reproductive steroid hormones, *Reprod Toxicol* **24** (3-4) (2007), pp. 317-325.

[57] M. Lacey, J. Bohday, S.M. Fonseka, A.I. Ullahand S.A. Whitehead, Dose-response effects of phytoestrogens on the activity and expression of 3beta-hydroxysteroid dehydrogenase and aromatase in human granulosa-luteal cells, *J Steroid Biochem Mol Biol* **96** (3-4) (2005), pp. 279-286.

[58] W.D. Han, Y.L. Zhao, Y.G. Meng, L. Zang, Z.Q. Wu, Q. Li *et al.*, Estrogenically regulated LRP16 interacts with estrogen receptor alpha and enhances the receptor's transcriptional activity, *Endocr Relat Cancer* **14** (3) (2007), pp. 741-753.

[59] Y.G. Meng, W.D. Han, Y.L. Zhao, K. Huang, Y.L. Si, Z.Q. Wu *et al.*, Induction of the LRP16 gene by estrogen promotes the invasive growth of Ishikawa human endometrial cancer cells through the downregulation of E-cadherin, *Cell Res* **17** (10) (2007), pp. 869-

880.

[60] Y.L. Zhao, W.D. Han, Q. Li, Y.M. Mu, X.C. Lu, L. Yu *et al.*, Mechanism of transcriptional regulation of LRP16 gene expression by 17-beta estradiol in MCF-7 human breast cancer cells, *J Mol Endocrinol* **34** (1) (2005), pp. 77-89.

[61] I.J. Benjaminand D.R. McMillan, Stress (heat shock) proteins: molecular chaperones in cardiovascular biology and disease, *Circ Res* **83** (2) (1998), pp. 117-132.

[62] J. Nishizawa, A. Nakai, T. Higashi, M. Tanabe, S. Nomoto, K. Matsuda *et al.*, Reperfusion causes significant activation of heat shock transcription factor 1 in ischemic rat heart, *Circulation* **94** (9) (1996), pp. 2185-2192.

[63] E.E. Baulieu, N. Binart, F. Cadepond, M.G. Catelli, B. Chambraud, J. Garnier *et al.*, Receptor-associated nuclear proteins and steroid/antisteroid action, *Ann N Y Acad Sci* **595** (1990), pp. 300-315.

[64] R.J. Heads, D.M. Yellonand D.S. Latchman, Differential cytoprotection against heat stress or hypoxia following expression of specific stress protein genes in myogenic cells, *J Mol Cell Cardiol* **27** (8) (1995), pp. 1669-1678.

[65] D.V. Cumming, R.J. Heads, A. Watson, D.S. Latchmanand D.M. Yellon, Differential protection of primary rat cardiocytes by transfection of specific heat stress proteins, *J Mol Cell Cardiol* **28** (12) (1996), pp. 2343-2349.

[66] J.S. Burchfield, J.W. Dong, Y. Sakata, F. Gao, H.P. Tzeng, V.K. Topkara *et al.*, The cytoprotective effects of tumor necrosis factor are conveyed through tumor necrosis factor receptor-associated factor 2 in the heart, *Circ Heart Fail* **3** (1) pp. 157-164.

[67] F. Xiang, Y.S. Huang, X.H. Shiand Q. Zhang, Mitochondrial chaperone tumour necrosis factor receptor-associated protein 1 protects cardiomyocytes from hypoxic injury

by regulating mitochondrial permeability transition pore opening, *FEBS J* **277** (8) pp. 1929-1938.

[68] Y. Masuda, G. Shima, T. Aiuchi, M. Horie, K. Hori, S. Nakajo *et al.*, Involvement of tumor necrosis factor receptor-associated protein 1 (TRAP1) in apoptosis induced by beta-hydroxyisovalerylshikonin, *J Biol Chem* **279** (41) (2004), pp. 42503-42515.

[69] G. Walker, K. MacLeod, A.R. Williams, D.A. Cameron, J.F. Smythand S.P. Langdon, Estrogen-regulated gene expression predicts response to endocrine therapy in patients with ovarian cancer, *Gynecol Oncol* **106** (3) (2007), pp. 461-468.

CHAPTER 5

QUANTITATIVE PERFORMANCE OF SEQUENTIAL IMMOBILIZED METAL AFFINITY CHROMATOGRAPHIC ENRICHMENT FOR PHOSPHOPEPTIDES

1. Introduction

Reversible phosphorylation of serine (Ser), threonine (Thr) and tyrosine (Tyr) residue is involved with the regulation of many critical biological processes and is a highly dynamic aspect of the proteome. In recent years, mass spectrometry based phosphoproteomics has emerged as a useful tool to survey the phosphorylation state of a complex protein mixture in a large-scale and high-throughput fashion. However, given the fact that most phosphoproteins are in low abundance with phosphorylation in low stoichiometry, enrichment technique(s) before MS analysis become a necessary step to separate phosphopeptides from a complex background such as a total cell lysate. Immobilized metal affinity chromatography (IMAC) based on ferric ions has long been used to capture phosphopeptides non-specifically [1]. Over the years, new IMAC chemistry based on various multivalent metal cations, such as gallium [2], zirconium [3], and titanium [4] has been introduced with varying selectivity and efficiency. Much attention has also been drawn to the use of metal oxide affinity chromatography (MOAC) for phosphopeptide enrichment due to its alleged higher recovery rate and selectivity compared to IMAC [5-14]. Numerous MOAC protocols based on different multivalent metal oxides such as titanium dioxide (TiO_2) [15], zirconium dioxide (ZrO_2) [16],

aluminum oxide (Al_2O_3) [17] have been widely adopted with promising enrichment efficiency. Interesting, it was reported recently that IMAC is less efficient for enrichment of mono-phosphorylated peptides than for multiple-phosphorylated species [18-20]. In contrast, MOAC was shown to be more efficient for capturing mono-phosphorylated peptides [21]. This is probably due to the fact that IMAC provides weaker affinity that mono-phosphorylated peptides may have poor retention on IMAC material while MOAC provides strong interaction that it become difficult to elute multiple-phosphorylated peptides. Therefore, Thingholm et al. introduced a novel sequential elution protocol from IMAC (SIMAC) that uses MOAC as the secondary enrichment step to capture mono-phosphorylated peptides left out by IMAC enrichment [21]. The application of this SIMAC protocol on whole cell lysate from human mesenchymal stem cells provided more phosphopeptide identifications than using MOAC or IMAC alone [21].

Despite the fact that IMAC- or MOAC-based protocols have been used in large-scale phosphoproteomics studies in recent years, questions about the reliability of these methods remain. In particular, there have been very few investigations of whether metal-based affinity enrichment techniques can be used in quantitative phosphoproteomics scenarios. Attention has usually been given to test the selectivity and sensitivity of the enrichment methods but not of quantitative performances such as the repeatability, dynamic range and linearity. In a typical large-scale phosphoproteomics study, a liquid chromatographic separation step such as SCX, HILIC or ERLIC is usually performed as a peptide fractionation procedure to reduce the sample complexity followed by metal-based affinity chromatography to isolate phosphopeptides from each fraction. However, each fraction usually contains peptide subsets with different total peptide amounts and

complexity. Such dynamic sample characteristics have made the estimation of the quantitative performance of metal-based affinity chromatography even more difficult in a real large-scale phosphoproteomics application.

In this study, we evaluated the repeatability, dynamic range and linearity of metal-based affinity chromatography for quantitative phosphoproteomics applications. The testing protocol was modified from the SIMAC procedure in which both IMAC and MOAC were performed sequentially as described by Thingholm et al. [21]. For IMAC, a gallium-based IMAC method was selected as gallium has been shown to have higher selectivity and sensitivity than other metal-based IMAC methods in a recent study [22]. The MOAC procedure was based on the most widely used TiO_2-MOAC protocol as described in [18]. Glycolic acid was used to prevent non-specific binding of non-phosphorylated peptides with acidic amino acid residues. Glycolic acid was introduced as an effective alternative to 2,5-dihydroxybenzoic acid (DHB) and phthalic acid as an acidic quenching agent [18].

Two experiments were performed. First, to test the enrichment repeatability of SIMAC with varying background, we constructed a series of peptide mixtures with a variety of loading conditions and complexity to mimic sample characteristic of peptide mixture as the result of LC pre-fractionation. In the second experiment, a series of phosphopeptide standard mixtures with different concentrations were spiked to estimate the linearity and dynamic range of the SIMAC method. Our data suggest

overall SIMAC shows limited performance on repeatability. The linearity test shown SIMAC can only be considered as a semi-quantitative method.

2. Experimental

2.1. Materials

HPLC-grade acetonitrile (ACN), water and acetic acid were obtained from Thermo Fisher (Waltham, MA, USA). Urea, dithiothreitol (DTT), iodoacetamide (IAA), sodium dodecyl sulfate (SDS) and 2DE Ready Prep clean up kits were purchased from Bio-Rad (Hercules, CA, USA). Bovine serum albumin (BSA), α-casein and β-casein, ammonium bicarbonate (ABC), formic acid (FA), trifluoroacetic acid (TFA), glycolic acid and ammonium hydroxide were purchased from Sigma (St. Louis, MO, USA). Sep-Pak SPE column with 200 mg C18 resin were obtained from Waters (Milford, MA, USA). Spin columns with filter (Cat# M105010S) were purchased from Boca Scientific (Boca Raton, FL, USA). Ga(II)-IMAC Nutip (part No.TT2GAA) from Glygen (Columbia, MD, USA) and Titansphere TiO_2 beads from GL Science (Tokyo, Japan) were used as enrichment media in all SIMAC experiments. Trypsin was purchased from Promega (Fitchburg, WI, USA). Phosphopeptide standard mixture (P33357) was purchased from Invitrogen (Carlsbad, CA, USA), and another phosphopeptide standards mixture (PHOSPHOSTD01) was from Glygen. Table 5.1 provided a detailed description of these peptide standards.

2.2. Protein preparation and digestion

Protein mixture in each experiment were dissolved in 25 mM ABC with 0.1% SDS,

then directly reduced by 10 mM DTT for 1 h and alkylated by 40 mM IAA for 30 min in dark. Alkylation was quenched by adding DTT to the final concentration of 20 mM. Protein was then precipitated using 2DE Ready Prep clean up kits according to the manufacturer's protocol. The resulting protein pellet was reconstituted in 25 mM ABC and digested with trypsin at a 30:1 protein:protease ratio. Digestion was carried out in 37 ºC for 5 h and stopped by acidification using TFA. Tryptic peptides were purified using a 200 mg C18 Sep-Pak SPE column and dried with a Speed Vac (Thermo Electron).

2.3. Experiment 1

BSA was chosen to create a non-phosphopeptide background as it is rich in acidic amino acids such as Asp and Glu, which compete with phosphopeptides during SIMAC enrichment. For practical purpose, three different loading amounts (100, 200, and 500 μg) were tested. These values were selected because most large-scale phosphoproteomics studies start with 1-5 mg of total lysate and most LC-prefractionation procedures generate 10-20 fractions. In each loading amount test, two different levels of sample complexity were created by mixing tryptic peptides from caseins:BSA at 1:49 or 1:99 w/w ratio to mimic the low abundance of phosphopeptides in real samples from cell lysate. In all experiments, caseins consist of equal amount of α and β isoforms. All six tests were repeated in triplicate to estimate the repeatability of SIMAC procedure with different loading background.

2.4. Experiment 2

To test the linearity and dynamic range of phosphopeptides enrichment using SIMAC, a series phosphopeptide mixtures were spiked into four samples each containing 200

μg tryptic peptides from caseins:BSA at ratio of 1:49. The amounts of spiked phosphopeptides are summarized in Table 5.3. All phosphopeptides were spiked prior to IMAC enrichment.

2.5. Phosphopeptide enrichment by SIMAC

The sequential elution protocol employed the Ga(II)-IMAC Nutip as the first stage of enrichment and then used Titansphere TiO2 beads as a second enrichment step to further enrich phosphorylated peptides from the flow-through of IMAC as summarized in Fig 5.1. Solution components in each step during SIMAC enrichment can be found in Table 5.4. During the IMAC enrichment, the Nutips were first equilibrated twice with 150 μL IMAC Binding Solution and then loaded with peptide mixture in 150 μL IMAC Binding Solution. Eluate was collected and reloaded again for complete binding. Two step of washing each with 150 μL IMAC Washing Solution 1 and 2 were performed followed by a 100 μL water wash to remove acid. Phosphopeptides were then eluted by 100 μL Elution Solution 1 and 2 sequentially. Flow-through fractions from IMAC loading and all washing steps were combined and dried by Speed Vac for MOAC enrichment. For the MOAC enrichment, the 2 mg TiO2 beads were aliquoted into a spin column filter and equilibrated with 300 μL MOAC Binding Solution twice and then loaded with dried IMAC flow-through in 300 μL MOAC Binding Solution. Eluate was collected and reloaded again for complete binding. Washing steps with 300 μL MOAC Washing Solution 1 and 200 μL MOAC Washing Solution 2 were performed followed by 100 μL water wash to remove acid. Phosphopeptides were then eluted with 100 μL Elution Solution 1 and 2 sequentially. Enrichment fractions from both IMAC and MOAC were then combined,

dried and re-dissolved in 10 μL of 3% ACN with 0.1% formic acid for LC-MS analysis.

2.6. Casein peptide identification by LC-MS/MS

The tryptic peptides mixture from α-casein and β-casein were injected in triplicates onto a G4240-62001 C-18 HPLC-Chip (40 nL enrichment column, 75 μm x 43 mm analytical column, 5 μm C-18SB-ZX, 300 Å, Agilent) hyphenated with a 6150 hybrid ESI-Q-TOF (Agilent). Nano-LC was performed with a 30-min gradient of 16–90% ACN in 0.1% FA at 0.5 μL/min. Mass spectrometric analysis was performed under positive mode with nanoelectrospray generated at 2.1 kV. The m/z response of the instrument was calibrated regularly with standards from manufacturer. Precursor scan and product ion spectra were acquired in centroid mode using data dependent acquisition in MassHunter (Agilent) with mass ranges for MS and MS/MS at m/z 250–2400 and 59–3000, respectively. Every second, a MS spectrum was scanned, followed by maximum of five product ion spectra. The switching from TOF-MS to MS/MS is triggered by precursors with ion intensity >1000 counts with dynamic exclusion for 30 sec. The collision energy was set with 5 V/100 Da slope offset with 2.5 V.

Spectra were searched by Spectrum Mill 3.3.084 (Agilent) at local server in Colorado State University against a SwissProt protein database under the taxonomy of *Bos taurus*. The search parameters were set to allow for up to two missed cleavages, carboxyamidomethylation on Cys as fixed modification, oxidation on Met, phosphorylation on Ser/Thr/Tyr as dynamic modification and 100 ppm for both MS

and MS/MS mass tolerance with decoy search mode. Autovalidation was performed using default setting within the Spectrum Mill (see supplementary material) to qualify confident identifications.

2.7. LC-MS quantitation of phosphopeptides

For both Experiment 1 and Experiment 2, 2 μL of the tryptic peptides mixture from α/β-casein and BSA spiked with phosphopeptide standards were injected in triplicate for LC-MS analysis on a Q-TOF instrument. The LC setting and LC-gradient were the same as described in Section 2.6 except that the Q-tof was operated in the MS scan only mode. Data were collected in centroid mode with MS m/z ranges set to 250–2400. Raw LC-MS data were then convert to mzData format using MassHunter (Agilent) for further analysis.

2.8. Data analysis

All mzData files from both Experiment 1 and 2 were loaded into MZmine 2 [23] to quantify phosphopeptide abundance based on LC-MS peak area. Briefly, compounds from each LC-MS run were recognized by their unique m/z and retention time values, isotopic peaks of each compound were then grouped to give reliable quantitation. Chromatograms of each compound were then aligned across samples/injections for comparison. Phosphopeptide species were searched using unique m/z values, and their peak areas from each samples/injections were exported to Excel for further analysis. Detailed procedure can be found in Appendix VI.

For Experiment 1, the repeatability of SIMAC in each loading condition was evaluated based on each enriched casein phosphopeptide by SIMAC across the repeats with peptide mixture at different starting amount and mixing ratio. For Experiment 2, the signal linearity of each spiked phosphopeptides was analyzed individually.

3. Results

3.1. Protein identification

As expected, most qualified peptide IDs were from BSA or the caseins. Non-phosphorylated peptides from both BSA and caseins were observed in relatively high abundance in all experiments (data not shown), indicating that non-specific binding is still an issue with the SIMAC procedure. Phosphopeptides from caseins with confident identification (see Appendix VII) were used for the LC-MS quantitative analysis in Experiment 1 as summarized in Table 5.2. Some phosphopeptide isoforms such as DIGSESpTpEDQAMEDIK and DIGSpESTpEDQAMEDIK, which have identical molecular weight and also similar retention time as they has very similar sequences, were grouped into one m/z species for LC-MS quantitation. In contrast, phosphopeptides containing oxidized Met or additional missed cleavage sites were considered as separate m/z species for LC-MS quantitation. As a result, five, three and one phosphorylation sites from casein αSI, αSII and β were identified from which 13 different m/z species were generated. All casein phosphopeptides are summarized in Table 5.2.

3.2. SIMAC repeatability

In our hands, for most casein phosphopeptides been quantified, the SIMAC procedure demonstrated overall limited repeatability for quantitative purposes (Fig 5.2). The biggest variance came from the difference between individual phosphopeptides. For example, phosphopeptide ions such as CP1, CP2-1, CP5-1, CP6-1 have significantly higher MS response than other ions. Moreover, CP3-1 and CP3-2 were not detectable in most LC-MS runs and were not included in quantitation analysis. Significant signal variation for most casein phosphopeptides were observed regardless of the changes in loading amount and sample complexity. As the signal variation between replicate injections in LC-MS is relatively small (data not shown), we postulate that the major MS signal variance was introduced by SIMAC enrichment. Thus, our data suggesting this SIMAC protocol may generate unreliable phosphopeptides quantitation. However, phosphopeptide CP1, CP4, CP5-5 and CP6-2 in 1:49 caseins:BSA test and CP2-2, CP4, CP5-3, CP5-5, and CP6-2 in 1:99 caseins:BSA shown incremental MS signal differences as expected when increasing amount of peptide mixture were enriched by the SIMAC. With the same loading amount, the MS signal of all casein phosphopeptides been investigated were higher when SIMAC were loaded with 1:49 caseins:BSA than loaded with 1:99 caseins:BSA as expected. Therefore, our results suggested SIMAC can only be used as semi-quantitative method with limited repeatability. Additionally, to this point, we concluded that there is no significant evidence showing one specific loading condition is preferable over other loading conditions.

3.3. SIMAC linearity and dynamic range

In the second experiment, four Invitrogen phosphopeptide standards (P1-P4) were spiked at 10-80 pmol, P5, P7 were spiked at 1-50 pmol, P6, P8 were spiked at 1-500 pmol into the sample, respectively. Satisfactory linear response was observed on the MS signal of $P4^{+3}$ (Fig 5.3B, R^2=0.98) when spiked from 10-80 pmol and $P8^{+3}$ (Fig 5.3F, R^2=0.99) spiked from 1-500 pmol for SIMAC enrichment. However, none of other standard phosphopeptides been investigated shown linear response over the spiked range, despite the overall trends of increasing MS signal intensity over spiked concentrations. The MS signals of $P1^{+2}$, $P1^{+3}$ (Fig 5.3A), $P2^{+3}$ (Fig 5.3B) and $P5^{+2}$ (Fig 5.3C) only showed linear trends in the first three concentration tiers but showed signal saturation spiked with high concentrations. MS signals of standard $P3^{+3}$ even level off after 20 pmol of loading amount. On contrary, multiple-phosphorylated peptides standard $P6^{+2}$ (Fig 5.3E) and $P7^{+3}$ (Fig 5.3D) showed low MS response at lower spiking amount (Fig 5.3C). Our data suggests that for different phosphopeptides, the linearity range of SIMAC-LC-MS varies from peptide to peptides. It is also noticeable that the signals observed showed significant differences between phosphopeptide species even though some of them were spiked with exactly the same amount in the same SIMAC repeat. However, it is unclear whether this is caused by different efficiency of SIMAC enrichment or of electrospray ionization between different phosphopeptide species.

4. Discussion

Metal-based phosphopeptides enrichment techniques such as IMAC and MOAC have been successfully employed for large-scale phosphoproteomics studies. However, IMAC and MOAC method development has primarily focused on the selectivity and sensitivity but not the quantitative attributes of the method. In a typical large-scale phosphoproteomics study, LC-based peptide fractionation usually generates multiple peptide mixture fractions with varying peptides amount and complexity from which phosphopeptides are enriched. Here, we evaluated whether dynamic characteristics such as loading amount and sample complexity can affect the performance of metal-based affinity chromatography.

Overall, the SIMAC method demonstrated limited repeatability in our hands between repeats regardless of loading amount and sample complexity. Thus, it is recommended that technical replicates be included at enrichment step to prevent detection of false changes. It is still unknown if the poor repeatability issue was the result of competitive binding of non-phosphorylated peptides to the SIMAC materials or just the result of severe ion suppression during electrospray with a more complex background. Nevertheless, to avoid both possible inhibition mechanisms, prefractionation is recommended to simplify the sample for both SIMAC enrichment and MS analysis. It is also notable that for phosphopeptide species such as CP4 and CP5-5 in the repeatability test, the difference of a MS signal due to 2-fold and 5-fold increase in loading amount falls into the same magnitude of SIMAC technical variance. Thus, small changes may not be quantifiable if SIMAC enrichment was

used with no technical replicates. The linearity test also has shown that different peptide species may have dramatically different linear relationships and linear range correlating the MS signal and their concentration. This means that not all phosphopeptides can be quantified at the same time from a complex sample in a real phosphoproteomics application by LC-MS coupled with enrichment techniques.

Although SIMAC procedures were designed to effectively enrich both mono-phosphorylated peptides and multiple- phosphorylated peptides [21], we experienced severe signal suppression of multiple-phosphorylated species such as CP3-2 from casein in Experiment 1 and P8 in Experiment 2. It is known that electrospray ionization is biased against multiple-phosphorylated peptides [7]. However, all multiple-phosphorylated peptides were detectable using Q-TOF if a casein digest or phosphopeptide standards were injected alone without a complex background (data not shown). Thus, we suspect that severe ion suppression in electrospray is at least in part responsible for the observed signal loss. Furthermore, we suspect that the severe ion suppression could be the result of non-phosphorylated peptides left-over from SIMAC when the enrichment began with low phosphopeptide abundance and a highly complex background. Glycolic acid was used to minimize non-phosphorylated peptides binding [18]. Despite the widespread implementation of ‘acidic peptide quenching’ agent, studies still report discrepant results on optimal enrichment conditions. Recent reports indicated that the addition of glycolic acid may hamper the selectivity [7, 24] and DHB can introduce bias against multiple-phosphorylated peptides [7]. In contrast, mono-phosphorylated peptides had higher overall signals in

our study.

5. Concluding Remarks

Our data suggest that overall the SIMAC procedure is insensitive to loading conditions and can be used as a semi-quantitative method with limited repeatability and linearity. We also show that the variance generated by the SIMAC procedure cannot be ignored, in particular without sufficient technical repeats. Therefore, careful evaluation of phosphopeptides enrichment techniques is recommended if they will be used in a quantitative phosphoproteomics study.

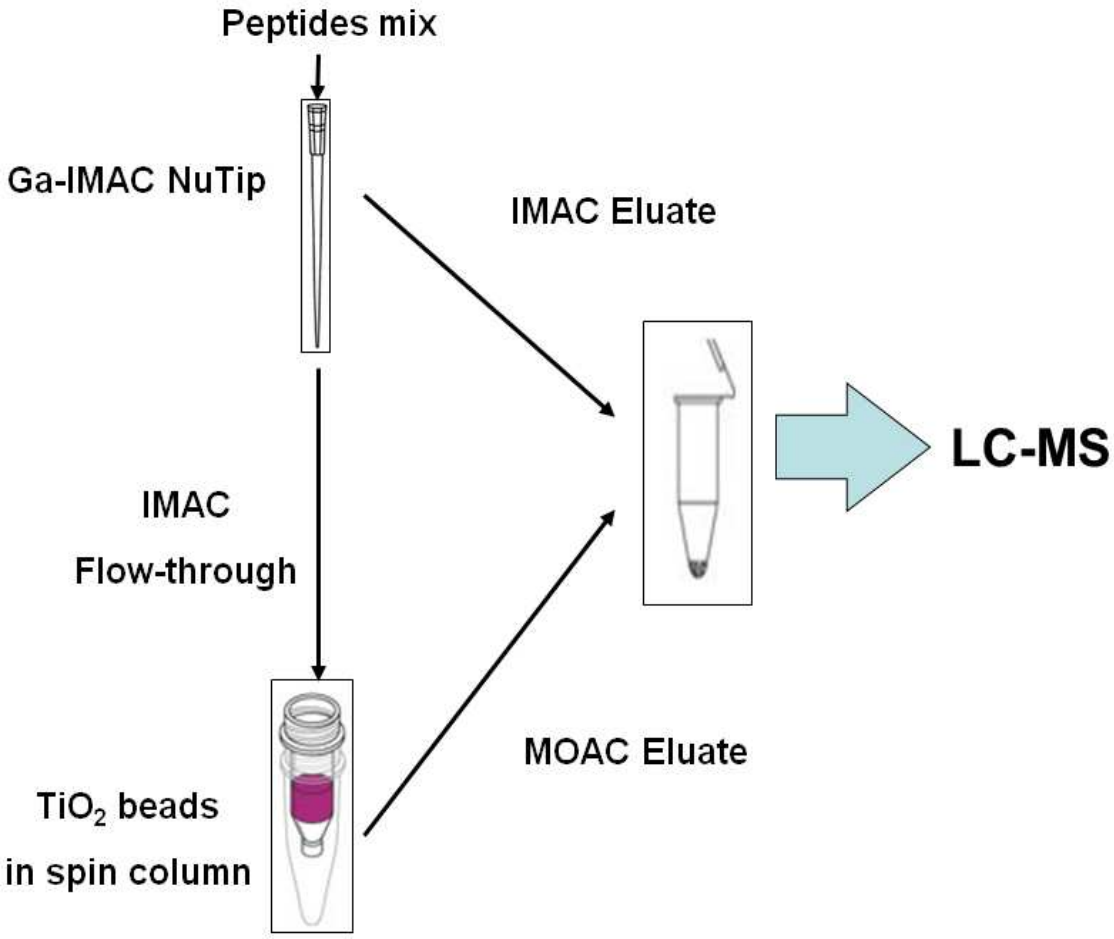
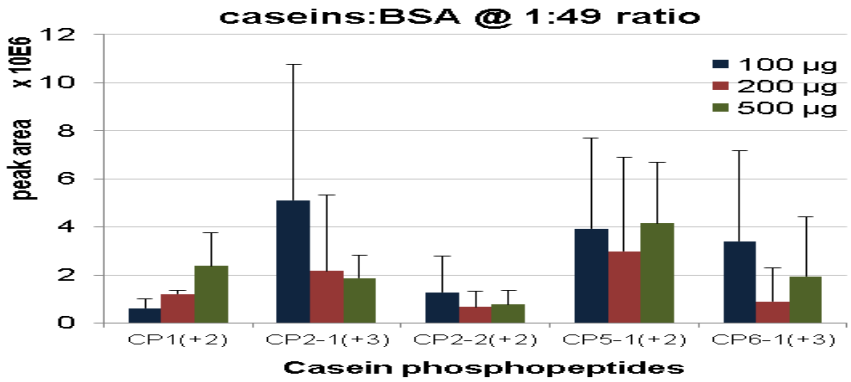
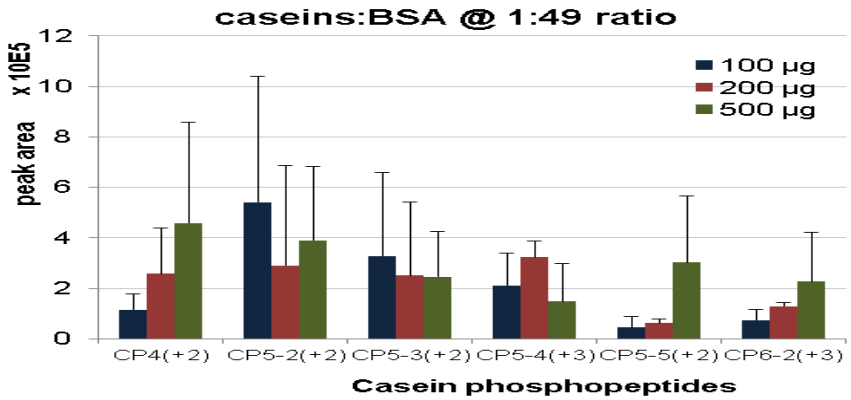


Fig 5.1, Schematic diagram of sequential IMAC workflow for phosphopeptides enrichment using Ga-IMAC and TiO2-MOAC. Eluates from both enrichment steps were combined prior to LC-MS analysis.

A



B



C

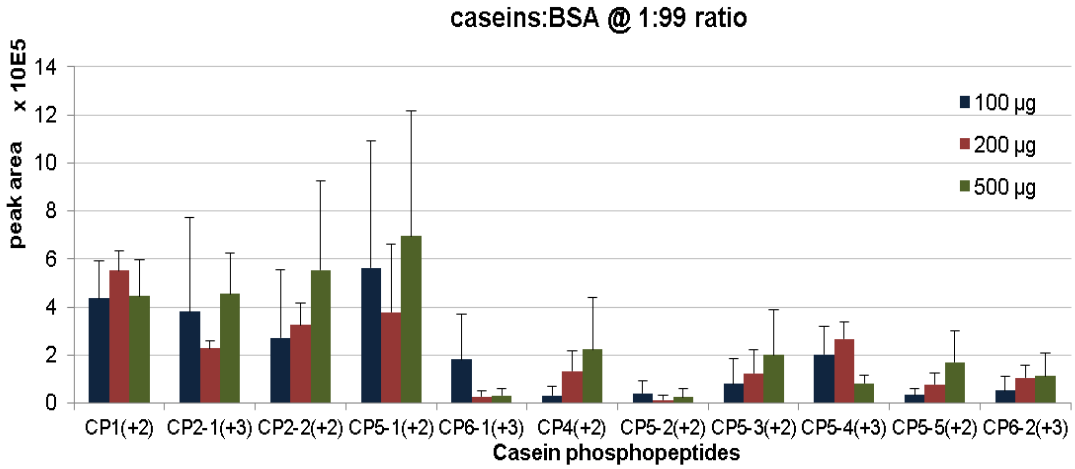


Fig 5.2, Repeatability test of SIMAC in experiment one. Plot A/B and C/D shown enrichment results from 1:49 and 1:99 caseins:BSA background, respectively. Each cluster of bars represents one casein phosphopeptide been quantified using LC-MS. Peak areas from replicate injections were averaged. The height of each bar represents the average peak areas of 3 enrichment replicates with standard deviation. Enrichment with 100, 200, 500 µg tryptic peptides were color-coded with blue, red and green.

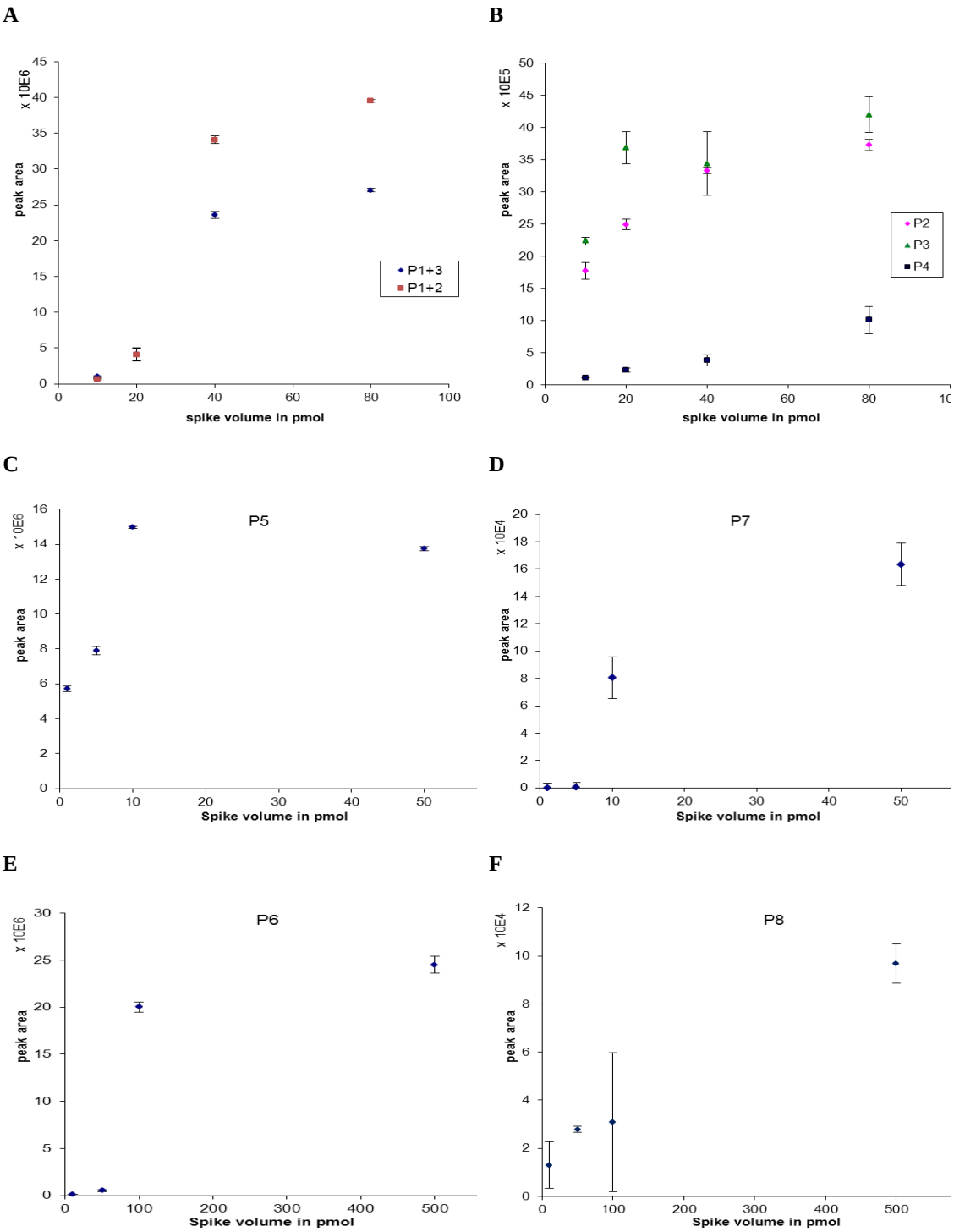


Fig 5.3, linearity test at different dynamic range. Each data points represent the averaged peak areas in scale from replicate injections.

Table 5.1 list of phosphopeptide standards used in this study. Actual M/Z species observed in Q-tof were shown here in bold.

Name	**Sequence**	**# Phosphor**	**$(M+H)^{1+}$**	**$(M+2H)^{2+}$**	**$(M+3H)^{3+}$**
Invitrogen P33357 mixture standard					
NP1	DRVYIHPF	0	1046.54	523.77	349.52
NP2	DRVYIHPFHL	0	1296.69	648.85	432.90
NP3	GKGRGLSLSRFSWGA	0	1578.85	789.93	526.95
P1	DHTGFLpTEpYVATR	2	1669.67	**835.34**	**557.23**
P2	TRDIpYETDYYRK	1	1702.75	851.87	**568.25**
P3	VPIPGRFDRRVpTVE	1	1720.89	860.95	**574.29**
P4	DLDVPIPGRFDRRVpSVAAE	1	2192.09	1096.53	**731.37**
Glygen PHOSPHOSTD01 standard					
P5	WWGSGPSGSGGSpGGGK	1	1500.60	**750.80**	500.87
P6	WWGSGPSGSpGGSpGGGK	2	1580.58	**790.78**	527.53
P7	WWGSGPSpGSpGGSpGGGK	3	1660.53	830.77	**554.18**
P8	WWGSpGPSpGSpGGSpGGGK	4	1740.45	870.75	**580.82**

Table 5.2 list of phosphopeptides from Casein digestion. Actual M/Z species observed in Q-tof were shown here in bold.

Name	**Sequence**	**# Phosphor**	**Casein**	**$(M+H)^{1+}$**	**$(M+2H)^{2+}$**	**$(M+3H)^{3+}$**
CP1	(K)VNELSpK(D)	1	αS1	769.35	**385.18**	257.12
CP2-1	(K)VPQLEIVPNSpAEER (L)	1		1660.79	**830.90**	**554.27**
CP2-2	(K)YKVPQLEIVPNSpAEER(L)	1		1951.95	**976.48**	651.32
CP3-1	(K)DIGSpESTEDQAMEDIK(	1		1847.71	**924.37**	**616.58**

	Q) (K)DIGSESpTEDQAMEDIK(Q) (K)DIGSESTpEDQAMEDIK(Q)					
CP3-2	(K)DIGSpESpTEDQAMEDI K(Q) (K)DIGSpESTpEDQAMEDI K(Q) (K)DIGSESpTpEDQAMEDI K(Q)	2		1927.69	**964.35**	**643.24**
CP4	(K)NMAINPSpKENLCSTFC K(E)	1	αS2	2093.88	**1047.45**	698.64
CP5-1	(K)TVDMESpTEVFTK(K) (K)TVDMESTpEVFTK(K)	1		1466.61	**733.81**	489.54
CP5-2	(K)TVDMoESTpEVFTK (K)	1		1482.61	**741.81**	494.88
CP5-3	(K)TVDMESpTEVFTKK (T) (K)KTVDMESpTEVFTK (T)	1		1594.70	**797.85**	532.24
CP5-4	(K)TVDMoESpTEVFTK K(T) (K)KTVDMoESpTEVFT K(T)	1		1610.69	805.85	**537.57**

CP5-5	(K)KTVDMESpTEVFTK K(T)	1		1722.81	**861.91**	**574.94**
CP6-1	(K)FQSpEEQQQTEDELQDK (I)	1	β	2061.83	**1031.42**	**687.95**
CP6-2	(K)IEKFQSpEEQQQTEDEL QDK(I)	1		2432.05	1216.53	**811.35**

Table 5.3 Experiment 2 test linearity and dynamic range of SIMAC enrichment using 200 μg casein:BSA tryptic peptides (1:49) spiked phosphopeptides p33357 contains equal molar of P1-P4. The order of spiking series were randomized. Unit: pmol

Sample	P33357	P5	P6	P7	P8
#1	10	5	50	1	500
#2	80	10	500	50	100
#3	20	1	10	5	50
#4	40	50	100	10	10

Table 5.4 Enrichment steps and solutions used for SIMAC procedure

Steps	**IMAC**	**MOAC**
Binding Solution	150 μL 5% acetic acid, 5% ACN	300 μL 1M Glycolic acid, 5% TFA, 80% ACN
Washing Solution 1		
Washing Solution 2	150 μL 0.1% acetic acid, 60% ACN	200 μL 1% TFA, 80% ACN
Water wash	100 μL HPLC grade water	
Elute Soltuion 1	100 μL 0.3N NH_4OH	
Elute Soltuion 2	100 μL 0.3N NH_4OH in 60% ACN	

REFERENCES

[1] G. Muszynska, G. Dobrowolska, A. Medin, P. Ekmanand J.O. Porath, Model Studies on Iron(Iii) Ion Affinity-Chromatography .2. Interaction of Immobilized Iron(Iii) Ions with Phosphorylated Amino-Acids, Peptides and Proteins, *Journal of Chromatography* **604** (1) (1992), pp. 19-28.

[2] M.C. Posewitzand P. Tempst, Immobilized gallium(III) affinity chromatography of phosphopeptides, *Anal Chem* **71** (14) (1999), pp. 2883-2892.

[3] M. Feng S Fau - Ye, H. Ye M Fau - Zhou, X. Zhou H Fau - Jiang, X. Jiang X Fau - Jiang, H. Jiang X Fau - Zou, B. Zou H Fau - Gong *et al.*, Immobilized zirconium ion affinity chromatography for specific enrichment of phosphopeptides in phosphoproteome analysis, (1535-9476 (Print)) pp.

[4] Z. Yu, G. Han, S. Sun, X. Jiang, R. Chen, F. Wang *et al.*, Preparation of monodisperse immobilized Ti(4+) affinity chromatography microspheres for specific enrichment of phosphopeptides, *Anal Chim Acta* **636** (1) (2009), pp. 34-41.

[5] A. Sanoand H. Nakamura, Titania as a chemo-affinity support for the column-switching HPLC analysis of phosphopeptides: application to the characterization of phosphorylation sites in proteins by combination with protease digestion and electrospray ionization mass spectrometry, *Anal Sci* **20** (5) (2004), pp. 861-864.

[6] A. Sanoand H. Nakamura, Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides, *Anal Sci* **20** (3) (2004), pp. 565-566.

[7] U.K. Aryaland A.R. Ross, Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry, *Rapid Commun Mass Spectrom* **24** (2) pp. 219-231.

[8] M.B. Gates, K.B. Tomerand L.J. Deterding, Comparison of metal and metal oxide media for phosphopeptide enrichment prior to mass spectrometric analyses, *J Am Soc Mass Spectrom* **21** (10) pp. 1649-1659.

[9] A. Leitner, M. Sturm, O. Hudecz, M. Mazanek, J.H. Smatt, M. Linden *et al.*, Probing the phosphoproteome of HeLa cells using nanocast metal oxide microspheres for phosphopeptide enrichment, *Anal Chem* **82** (7) pp. 2726-2733.

[10] J.V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen *et al.*, Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, *Cell* **127** (3) (2006), pp. 635-648.

[11] B. Macek, I. Mijakovic, J.V. Olsen, F. Gnad, C. Kumar, P.R. Jensen *et al.*, The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis, *Mol Cell Proteomics* **6** (4) (2007), pp. 697-707.

[12] M.R. Larsen, T.E. Thingholm, O.N. Jensen, P. Roepstorffand T.J. Jorgensen, Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns, *Mol Cell Proteomics* **4** (7) (2005), pp. 873-886.

[13] F. Wolschin, S. Wienkoopand W. Weckwerth, Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC), *Proteomics* **5** (17) (2005), pp. 4389-4397.

[14] M.W. Pinkse, P.M. Uitto, M.J. Hilhorst, B. Oomsand A.J. Heck, Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns, *Anal Chem* **76** (14) (2004), pp. 3935-3943.

[15] Y. Ikeguchiand H. Nakamura, Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line

preconcentration on titania, *Analytical Sciences* **13** (3) (1997), pp. 479-483.

[16] H.J. Zhou, R.J. Tian, M.L. Ye, S.Y. Xu, S. Feng, C.S. Pan *et al.*, Highly specific enrichment of phosphopeptides by zirconium dioxide nanoparticles for phosphoproteome analysis, *Electrophoresis* **28** (13) (2007), pp. 2201-2215.

[17] E.W. Shin, J.S. Han, M. Jang, S.H. Min, J.K. Parkand R.M. Rowell, Phosphate adsorption on aluminum-impregnated mesoporous silicates: Surface structure and behavior of adsorbents, *Environmental Science & Technology* **38** (3) (2004), pp. 912-917.

[18] M.R. Jensen Ss Fau - Larsenand M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques, (0951-4198 (Print)) pp.

[19] M. Nousiainen, H.H. Sillje, G. Sauer, E.A. Niggand R. Korner, Phosphoproteome analysis of the human mitotic spindle, *Proc Natl Acad Sci U S A* **103** (14) (2006), pp. 5391-5396.

[20] S.B. Ficarro, M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz *et al.*, Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae, *Nat Biotechnol* **20** (3) (2002), pp. 301-305.

[21] T.E. Thingholm, O.N. Jensen, P.J. Robinsonand M.R. Larsen, SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides, *Mol Cell Proteomics* **7** (4) (2008), pp. 661-671.

[22] R. Kange, U. Selditz, M. Granberg, U. Lindberg, G. Ekstrand, B. Ek *et al.*, Comparison of different IMAC techniques used for enrichment of phosphorylated peptides, *J Biomol Tech* **16** (2) (2005), pp. 91-103.

[23] T. Pluskal, S. Castillo, A. Villar-Brionesand M. Oresic, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics* **11** pp. 395.

[24] N. Sugiyama, T. Masuda, K. Shinoda, A. Nakamura, M. Tomitaand Y. Ishihama, Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications, *Mol Cell Proteomics* **6** (6) (2007), pp. 1103-1109.

CHAPTER 6

QUANTITATIVE PHOSPHOPROTEOMIC ANALYSIS OF SIGNALING PATHWAY PERTURBATION BY GENISTEIN EXPOSURE: AN INITIAL STUDY ON PRIMARY CARDIOMYOCYTES

1. Introduction

Epistemological evidences have suggests that soy-derived phytoestrogens exhibit benefits for cardiovascular heath [1-5]. Genistein or GEN, the most abundant isoflavone synthesized by soybean, has been postulated to provide direct cytoprotection for cardiomyocytes against ischemic stresses [6-9]. Like other plant isoflavones, GEN is a weak estrogenic compound and can effectively binds to estrogen receptors (ER) at physiological relevant concentration (1-10 μM) [10-12]. This estrogenic property has been shown to contribute to the cardioprotective action of GEN [6, 7]. However, GEN at pharmacological relevant concentration (>10 uM) was also shown to be an effective pan-specific protein tyrosine kinases (PTK) inhibitor. It is postulated that the ER-independent PTK inhibition can possibly abolish the cardioprotection afforded by GEN. Genistein were used methodologically to block key PTK functions that mediate the cytoprotection pathways triggered by ischemic preconditioning procedure [13-17]. Genistein at 50 μM have shown to block PKC-mediated cardioprotection in the rabbit heart against the long ischemia via the

inhibition of downstream tyrosine kinase which appears to be a MAP kinase [13]. It is also shown that EGFR-mediated TK activation contributed to the cardioprotection in rat heart triggered by the ischemic/reperfusion preconditioning and thus can be blocked by GEN [14]. In a most recent study, GEN was also used as a receptor tyrosine kinase inhibitor to block the upstream signaling of zinc-induced cardioprotective Akt pathway in a cultured cardiomyocyte model [18]. Another study shown that 50 μM of GEN can effectively block pro-apoptosis Fas signaling triggered by hypoxia in murine ventricular myocytes [19]. Apart from these reports, most signaling cascades modulated by GEN in cardiomyocytes are largely unknown.

MS-based phosphoproteomics has become a valuable tool for characterizing protein phosphorylation in large-scale. To understand the impact of GEN on signaling pathways in cardiomyocytes, we introduced a high throughput phosphoproteomic approach based on phosphopeptides enrichment and label-free quantitation using electron transfer dissociation (ETD) mass spectrometry to globally identify the phosphorylated proteins and sites regulated by GEN in cardiomyocytes. To resolve the signaling alteration due to TK inhibitory effects and other ER-independent effects of GEN, primary ventricle myocytes isolated from male SD rat were incubated with GEN in the presence of ER antagonist ICI 182,780 before phosphoproteomics analysis.

2. Material and Methods

2.1. Isolation of rat primary cardiomyocytes

The use of animal in this experiment was approved by the CSU Institutional Animal Care and Use Committee in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals*. Primary cardiomyocytes were isolated from 6 month old male Sprague-Dawley rats weighing 300-350 g (n=3). Hearts were surgically removed according to the protocol described previously [20]. Excised hearts were subject to retrograde-perfusion of Joklik solution (Sigma, MO) containing 0.8 mg/ml type II collagenase (Worthington Biochemical, NJ). Following tissue digestion, the cell suspension was filtered and pelleted in a solution containing 10% bovine serum albumin (BSA, Sigma, MO), 111 mM NaCl, 5mM KCl, 10 mM NaH_2PO_4, 1 mM $MgSO_4$, 50 uM $CaCl_2$, 5mM HEPES, 60 mM taurine and 20 mM creatine. The myocytes pellet was then resuspended in Joklik solution with 6% BSA and gradually reintroduced $CaCl_2$ to final concentration of 1 mM. After incubation with $CaCl_2$ for 20 mins, cells collected by centrifuge (~300 rpm for 3 min) and resuspended in M119 medium supplemented with 10% fetal bovine serum (Sigma, MO), 10 mM butanedione monoxime (Sigma, MO), 2 mM glutamine (Invitrogen, CA), 100 U penicillin and 100 mg/ml streptomycin (Invitrogen, CA). Cells were then plated on 0.01% laminin-coated dishes followed by incubation at 37°C, 5% CO_2 for 4 hrs.

2.2. Treatment protocols

Cardiomyocytes from each animal were divided into 2 separate groups for treatments.

The treatment group was replenished with fresh supplemented containing 10 μM genistein with 10 μM ICI 182,780 (Tocris Bioscience, MO), a full estrogen receptor antagonist (GEN+ERA). The control (Con) group was treated with DMSO vehicle only. Myocytes were then incubated in at 37°C with 5% CO_2 for 15 min.

2.3. Protein extraction

Following treatments, cells were gently washed by 5 mL PBS twice each for 10 sec. Lysis solution (1 ml) containing 10 mM Tris-HCL (pH 8.0), 7 M urea, 2 M thiourea, 1X protease inhibitor cocktail (Sigma, MO), 1X HALT phosphatase inhibitor cocktail (Thermo Fisher Scientific, IL), 5 mM magnesium acetate and 1% CHAPS (Sigma) were then directly applied to the cells for quick protein extraction. Protein solutions were then subjected to ultrasonication for 5 min in ice bath. Protein samples were centrifuged at 17 000 x g and the supernatant was collected. Total protein concentration for each sample was determined using the RC-DC protein assay kit (Bio-Rad, CA) according to manufacturer's instruction. Bovine serum albumin (Sigma, MO) was used as protein assay standard.

2.4. Protein digestion and HILIC separation

A total of 2 mg of protein from each sample was directly reduced by 10 mM dithiothreitol for 1 h in lysis buffer and alkylated by 40 mM iodoacetamide for 30 min in dark. Alkylation was quenched by adding dithiothreitol to final concentration of 20 mM. Protein was then precipitated by 2DE Ready Prep clean up kits (Bio-Rad) according to manufacturer's protocol. The resulting was reconstituted by 600 μL 7 M

urea in 50 mM ABC stock with 0.17% ProteaseMAX (Promega, WI). Once the pellets were dissolved completely by votexing, 5 mL 50 mM ABC stock was added to the sample to dilute the urea. In order to break down the DNA and RNA, samples were first incubated with 250 U/ml of benzonase (Sigma, MO) with 1mM $MgCl_2$ for 1.5 h at 37 ºC. For protein digestion, 100 μg trypsin (Promega) was added to the sample. Digestion was carried out in 37 ºC for 4h and stopped by acidification using trifluoroacetic acid. Tryptic peptides were cleaned up by 500 mg C18 Sep-Pak SPE column (Waters), then dried by Speed Vac (Thermo Electron) and reconstituted in HILIC buffer B.

Peptides were fractionated using a PolyHYDROXYETHYL A column (4.6 mm × 200 mm, 5 μm particle size, 200 Å pore size) (PolyLC, Columbia, MD) on a 1050 HPLC system (Agilent, CA) at flow rate of 0.5 mL/min. To prevent column overloading, 2 mg of peptide was separated into 3 different each with approximately 667 μg peptides. A 60 min slow gradient designed for phosphopeptides separation using a combination of 0.1% TFA (Buffer A) and 0.1 %TFA in 85 % ACN (Buffer B) was created. The gradient started with 95% B for 5 min; then dropped to 70% B for 42.5 min; followed by 70-0% B for 5 min; then maintained at 0% B for 2.5 min; finally came back to 95% B within 2.5 min. The column was conditioned in 95% B for 20 min after each gradient and before the next run to ensure reproducibility. A total of 15 fractions were collected with 3.27 min intervals and a few (begin and end fractions) at 5 min intervals. Identical fractions from 3 replicate runs were combined, dried via Speed Vac, and reconstituted in IMAC loading buffer.

2.5. Phosphopeptides enrichment using SIMAC

Phosphopeptides from HILIC peptide fractions were enriched by sequential elution protocol from IMAC (SIMAC) modified from the procedure described by Thigholm et al [21]. This approach employed Ga(II)-IMAC Nutip (part No.TT2GAA, Glygen MD) as the first stage enrichment for multiple-phosphorylated peptides and then Titansphere TiO_2 beads (GL Science, Japan) as a second enrichment step to enrich mono-phosphorylated peptides from the flow-through of IMAC. During the IMAC enrichment, the Nutips were first equilibrated twice with 150 μL IMAC binding solution containing 5% acetic acid, 5% ACN and then loaded with peptide mixture in 150 μL IMAC binding solution. Eluate was collected and reloaded again for complete binding. Tips were then washed with 150 μL IMAC binding solution and 150 μL 0.1% acetic acid in 60% ACN. Tips were then washed by100 μL HPLC-grade water to remove acid. Phosphopeptides were then eluted by 100 μL elution solution 1 containing 0.3N NH_4OH and then elution solution 2 containing 0.3N NH_4OH in 60% ACN sequentially. Flow-through fractions from IMAC loading and all washing steps were combined and dried by Speed Vac for MOAC enrichment. For the MOAC enrichment, the 2 mg TiO_2 beads were aliquoted in spin column filter (Cat# M105010S, Boca Scientific FL) and equilibrated twice with 300 μL MOAC binding solution containing 1M glycolic acid, 5% TFA in 80% ACN. Prepared TiO_2 beads were then loaded with dried IMAC flow-through in 300 μL MOAC binding solution. Eluate was collected and reloaded again for complete binding. Two step of washing with 300 μL MOAC binding solution first and then 200 μL 1% TFA in 80% ACN

were performed followed by 100 μL water wash to remove acid. Phosphopeptides were then eluted by 100 μL elution solution 1 and 2 sequentially. Enrichment fractions from both IMAC and MOAC were then combined and dried and re-dissolved in 10 μL of 3% ACN, 0.1% formic acid for ETD-MS/MS analysis.

2.6. Protein identification by ETD MS/MS

Each SIMAC-enriched HILIC fraction was injected in triplicates onto the G4240-62010 LC-chip (Agilent) interfaced to a Agilent 6340 ion trap mass spectrometer (Agilent) equipped with an ETD source. LC separation was performed by an Agilent 1100 Series HPLC-Chip system with a 25 min gradient flowing at 450 nL/min starting with 97% solvent A (0.1% FA) and 3% solvent B (0.1% FA in 90% ACN) for 1 min and increase to 40% B over 11 min and from 40 to 80% B over the next 3 min, and maintained 80% B for 5 min, and finishing up by drop back to 3% B. Peptides eluent was analyzed by ion trap MS under the ETD only mode. The m/z response of the instrument was calibrated regularly with standards from manufacturer. Precursor scan and product ion spectra were acquired in centroid mode using data dependent data acquisition in MassHunter DataAnalysis (Agilent) with the following parameters: mass ranges for MS and MS/MS were m/z 250–4000 and 50–2000, respectively. Every 3 seconds, a MS spectrum was scanned, followed by two product ion spectra. The switching from MS to MS/MS is triggered by precursors with ion intensity >1000 counts with dynamic exclusion for 30 sec.

2.7. Spectrum analysis

Peak lists in mgf format were extracted from .d files by Mascot Distiller 2.3.2 (Matrix Science) for peak deisotope and charge state determination and then submitted to Mascot v2.3 (Matrix Science) in local server at CSU and X!Tandem Cyclone v 2010.06.01.6 for consensus search against a target-reverse concatenated International Protein Index (IPI , European Bioinformatics Institute, http://www.ebi.ac.uk/IPI) rat protein database (v3.70, 79158 sequences). The search parameters were set to allow for up to two missed cleavages, carboxyamidomethylation on Cys as fixed modifications, oxidation on Met, phosphorylation on Thr/Ser/Tyr as variable modifications, a MS and MS/MS mass tolerance of 2.5 Da and 0.7 Da, respectively. Searches were done using monoisotopic mode. Mascot search was set to use ETD-TRAP mode and X!Tandem was instructed to do search based on c, z and y ions. A consecutive second round X!Tandem search was automatically done with more variable modifications: oxidation on Trp and dioxidation on Met. All search hits from both Mascot and X!Tandem were compiled by Scaffold v3.00.03 (Proteome Software, OR) for peptide and protein probability calculation [22]. Peptide spectrum matches (PSMs) were first screened with a preliminary filter of 90% peptide probability and 80% protein probability in prior to following phosphorylation sites analysis. All qualified PSMs in each sample will be further loaded into ScaffoldPTM v1.0.3 (Proteome Software, OR) for phosphorylation site assignment. To assess the assignment ambiguity of a phosphorylation site, the Ascore developed by Sean Beausoleil et al were used to calculate the location probability of the phosphorylation site in the spectrum matched peptide sequence [23]. Qualified peptides matches with

phosphorylation site were filtered by 90% peptide probability and 90% Ascore probability and exported to Excel for further statistical analysis. The false discover rate (FDR) was calculated dividing the number of false hits by the number of all hits as summarized by table 6.1:

All target hits	= all hits above the filtering criteria
False positives (FP)	= Decoy hits
True positives (TP)	= All target hits - FP
FDR	= FP/(FP+TP)

Within each sample, spectrum gave the identification of the same phosphorylation site were counted and compared between 3 GEN+ERA samples and 3 Con samples using paired t-test to select significant ($p < 0.1$) changes of phosphorylation due to GEN+ERA treatment.

3. Results and Discussion

3.1. Phosphopeptides identification

Using preliminary filter of peptide 90% probability and 80% protein probability, 7296 PSMs were assigned with peptide FDR of 1.7%. These peptide IDs give identification of 229 proteins. Within those PSMs, only 890 PSM matched to phosphorylated peptides which gave 57 unique phosphosites identification from 49 phosphoproteins. However, it is notable that majority of those non-phosphorylated peptide IDs came from the abundant proteins such as myosin-6. In consistence with our conclusion from study 2, high abundant proteins also dominated over other proteins in this

spectral counting study. Out of 7296 total PSM in the whole experiment, the most abundant 10 proteins occupied 5324 (72.9%), in which 2727 (37.3%) PSMs came from myosin-6. Myosin-6 alone gave 41 phosphopeptide PSMs but 2686 non-phosphopeptide PSMs. We suspected that the presence of high abundant proteins, in particular those myofibril proteins may significantly interfere with the MS characterization of phosphopeptides in low abundant simply by occupying a large portion of instrumental sampling cycles and introducing severe ion suppression during the electrospray ionization. Therefore, we suspect that the sensitivity of our protocol on detecting signaling phosphoproteins in cardiomyocytes can be improved dramatically by immunodepletion of abundant proteins like myosin-6, actin, myosin light chain 3, tropomyosin alpha-1 chain, myoglobin and titin as a sample preparation step before the proteomics analysis. However, it is worth to point out that protein phosphorylation is also a ubiquitous PTM found in those contractile proteins and may have implications of key functional regulation mechanisms.

3.2. Efficiency of SIMAC enrichment

The sequential elution protocol from IMAC (SIMAC) was developed to achieve better enrichment efficiency for both mono- and multiple-phosphorylated peptides [21]. Out of 890 PSMs assigned to phosphopeptides in the whole experiment, 52 unique phosphopeptide species were identified including 40 mono- and 12 multiple-phosphorylated peptides, respectively. Consider the relative low abundance of multiple-phosphorylated peptides and the fact that they are more difficult to be characterized by MS, our data suggest SIMAC procedure did enriched both mono-

and multiple-phosphorylated peptides effectively. However, significant non-specific binding of non-phosphorylated peptides from abundant proteins indicated that the SIMAC protocol still lacks the selectivity when loaded with very complex sample. This conclusion is in consistence with our previous SIMAC test study. Although at this point we are unable to identify if the majority of non-phosphorylated peptides bind to IMAC or MOAC step, but considering the MOAC enrichment was carried out with the presence of 1 M glycolic acid as non-phosphorylated peptide excluder, we suspect that an augmented selectivity of IMAC can significantly improve the overall selectivity of SIMAC procedure.

3.3. ETD performance

Totally only 7296 PSMs were matched out of 1083104 spectra collected across the whole experiment which gave a relatively low overall identification rate (0.67%). One issue of spectra interpretation was the relatively low identification rate of +2 charged ions compare to +3 charge ions (Fig 6.1). When parsing through all spectra collected by the ion trap in the whole experiment, +2 and +3 ions both accounted for 20% of all charged species. However, after database searching, +2 and +3 charged ions generated 38.7% and 54.6% of the total successful identifications, respectively. As most peptides should form +2 or +3 charged ions during ESI, these statistics have clearly shown a systematic bias in ETD data interpretation against the +2 charged species. However, we are unable to identify whether this bias is due to the unfavorable fragmentation of +2 ions in ETD trap or differential interpretation of ions with different charges in search algorithms. It is noticeable that several recent reports

have suggested that ETD is an ineffective fragmentation method for peptide dications [24-26]. We suspect this issue can be solved by employ complementary CID collision mode either by alternating the ion trap in CID-ETD mix mode to get better characterization of +2 ions or add a supplementary CAD fragmentation during ETD for +2 ions described as the ETcaD protocol [26].

Another problematic issue with ETD data was the presence of dominant precursor species, specifically, the precursor ions itself $(M+nH)^{n+}$, charge reduced precursors $(M+nH)^{(n-1)+}$ in the MS/MS spectrum as exemplified in Fig 6.2. Such phenomenon has been described by Good, et al in the ETD MS/MS spectrum generated by a ThermoFisher's hybrid QLT-Orbitrap instrument [27]. This group developed a script to remove those dominant precursor peaks and have demonstrated much improved peptide identification rate using open mass spectrometry search algorithm (OMSSA) after precursor species reduction [27]. However, due to the proprietary nature of MS data generated in the Agilent 3D ion trap we are unable to try if a script can be written to remove the dominant precursors in the MS/MS spectra, nor we know how the Mascot and X!tandem remove those precursor species if there is any procedures in those algorithms.

3.4. Consensus database search performance

Most search engines including Mascot and X!tandem were originally written to process CID-based data and have been adapted to accept ETD-based data in recent development. However, it is difficult to assess the effectiveness of ETD data

interpretation using Mascot and X!tandem as their core algorithm is proprietary. Even using a consensus search strategy to combine search results from both Mascot and X!tandem, overall identification rate for the whole experiment only reach 0.67% indicating a poor performance of the search engines. However, we believe that low MS accuracy, and the presence of dominant precursor species in MS/MS spectra may partly contributed the discover power of search engines.

Compare the performance of both search engines, Mascot identified more PSMs and phosphopeptide PSMs than X!tandem even with 2nd round search (Fig 6.3). However, both search engines contribute significantly amount of PSMs that otherwise would not be detectable in the other search engine which highlighted the effectiveness of the consensus search strategy. Furthermore, we also found that both Mascot and X!tandem shown no bias towards phosphopeptides as phosphopeptide PSMs and other PSMs shown comparable Mascot ion score and X!tandem -log(e) score distribution as summarized by Fig 6.4.

3.5. Differential phosphorylation due to non-estrogenic action of GEN

Despite the difficulty of resolve protein phosphorylation in this study, we still observed 3 phosphorylation sites in 2 proteins inhibited by GEN treatment in the presence of ERA as summarized in table 6.2. The phosphorylation on the 117T site of ‘myosin-binding protein C, cardiac-type’ was completely abolished by GEN via non-estrogenic mechanism. Protein phosphorylation at 6722S and 6740S in ‘similar to titin isoform N2-A’ were down-regulated 2.9 and 2.7 fold by GEN via non-estrogenic

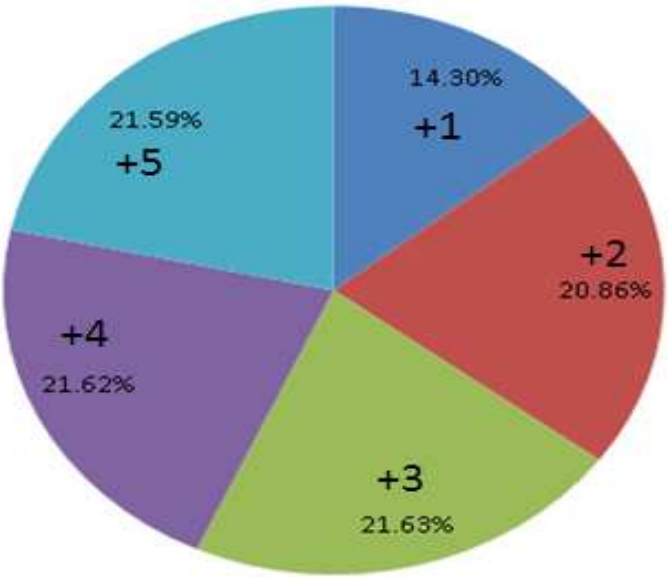
mechanism. All three identified phosphorylation sites are novel discoveries and have not been reported in protein phosphorylation database such as phosphor.ELM (http://phospho.elm.eu.org) [28]. However the biological functions related to those phosphorylation sites are still unclear.

4. Concluding Remarks

In this project, a phosphoproteomic approach based on HILIC peptide fractionation, SIMAC phosphopeptides enrichment, and label-free quantitation using electron transfer dissociation (ETD) mass spectrometry was tested to understand impact of GEN on the global protein phosphorylation in rat cardiomyocytes. However, only 57 unique phosphosites from 49 phosphoproteins were identified in the whole experiment which gave us very limited resolving power to see the phosphoproteome alteration in response to GEN treatment. Several potential pitfalls of this phosphoproteomics application on cardiomyocytes have been identified. The presence of high abundant proteins, in particular the myofibril proteins in the sample strongly limit the phosphoproteomics power to identify phosphorylation events which is in low abundance and low stoichiometry. With such complex sample background, the risk of having non-specific binding of unphosphorylated peptides in SIMAC procedure and electrospray ion suppression is conceivable high. Enrichment using SIMAC still shown low selectivity when loading samples with such noisy background and further optimization, in particular the IMAC step is recommended. However, our data suggest the SIMAC have successfully enriched both mono- and

multiple-phosphorylated peptides as expected. In order to characterize more phosphopeptides species, it is also recommended to use complementary CID fragmentation in addition to current ETD mode for future studies.

A



B

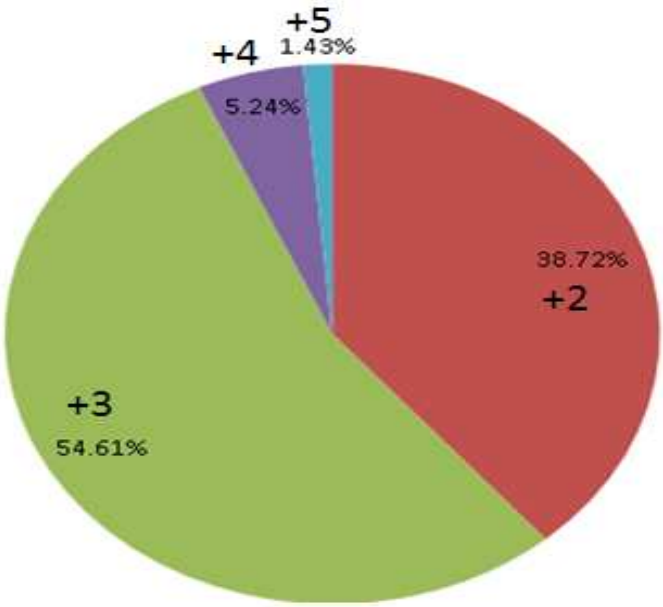
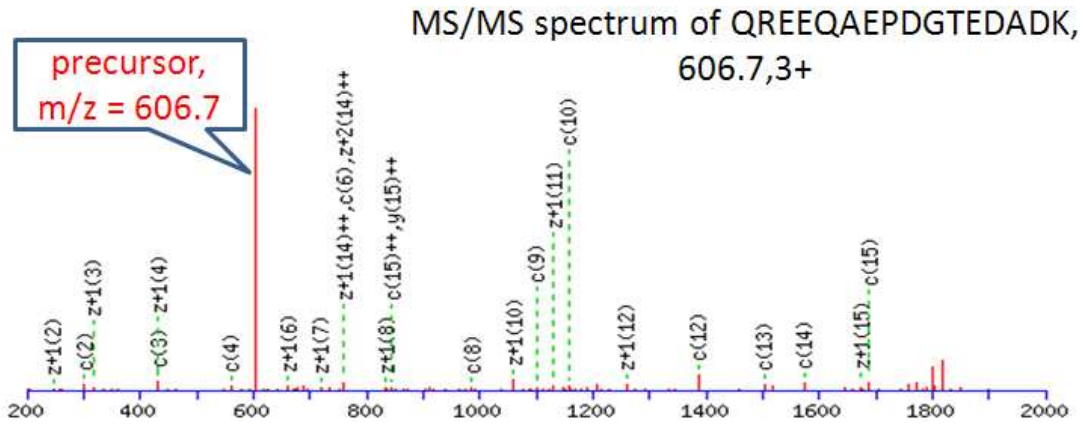


Fig 6.1 Distribution of charge states of all ions in the raw MS spectra (A) and that among those identified PSMs (B) across the whole experiment

A



B

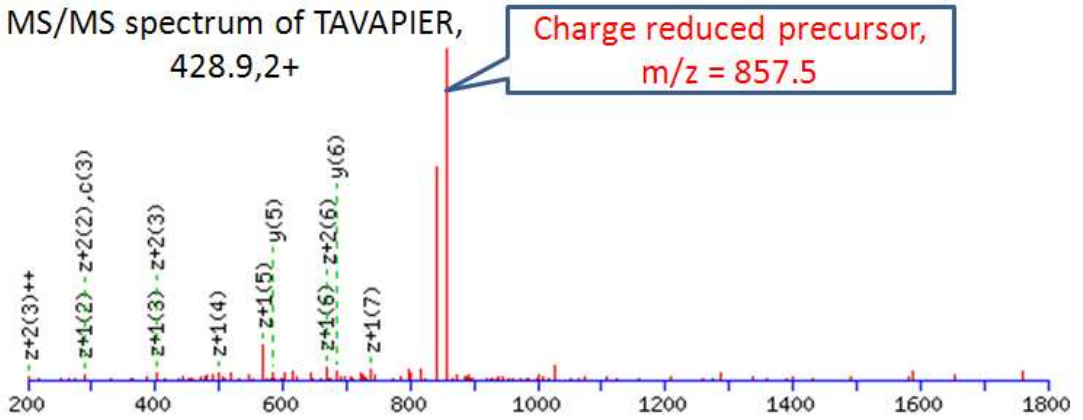
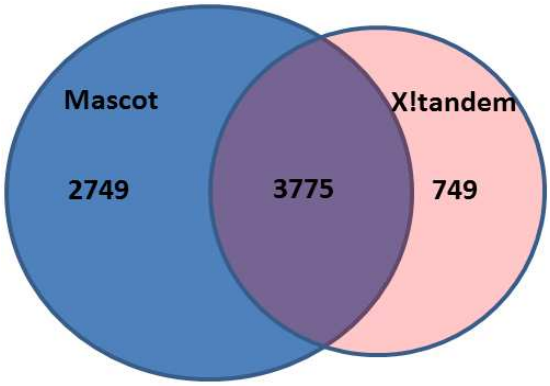


Fig 6.2 Example of ETD spectra showing the presence of dominant precursor peak (A) and charge reduced precursor peak (B) in ETD MS/MS spectrum.

A



B

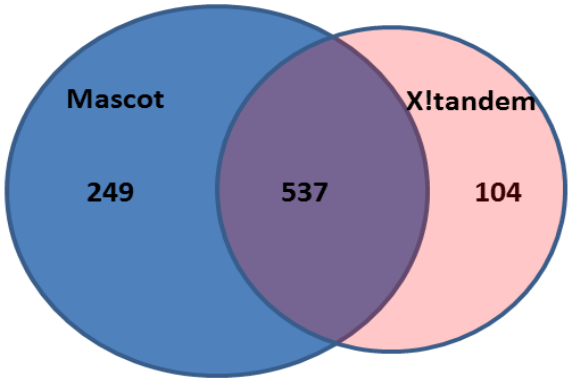


Fig 6.3 Number of total PSMs (A) and number of phosphopeptides PSM (B) identified by Mascot and X!tandem using 90% peptide probability and 80% protein probability filter.

A

B

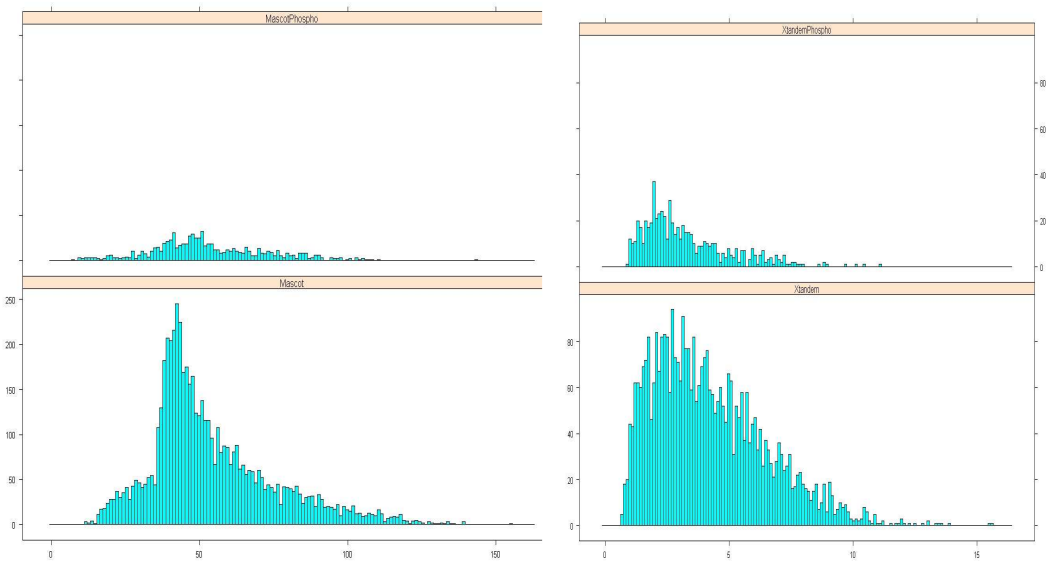


Fig 6.4 Frequency histogram of Mascot ion score (A) and X!tandem -log(e) score (B) for all PSM pass the 90% peptide probability and 80% protein probability filter. PSMs gave phosphopeptide IDs from both search engines (the upper panels), shown no difference in score distribution from all other PSMs (lower panels).

Table 6.2, Protein phosphorylation identified in response to GEN+ERA treatment. Peptide sequences were listed with # following the phosphorylated amino acid. CON and GEN+ERA column shown the spectral counting result for the corresponding phosphorylation sites in control and GEN+ERA treated group. Three numbers each represent the number of spectra found in samples from rat 1,2,3.

Accession	Protein	peptide	CON	GEN+ERA	p-value
IPI00870316	myosin-binding protein C, cardiac-type	AESAVAPTSMEAPET#PK	5,2,4	0,0,0	0.03
IPI00388754	similar to titin isoform N2-A	AVS#PTETKPTEK	6,8,5	0,6,1	0.04
IPI00388754	similar to titin isoform N2-A	VKSPETVKS#PK	21,35,12	0,12,11	0.08

REFERENCE

[1] A. Menotti, A. Keys, H. Blackburn, D. Kromhout, M. Karvonen, A. Nissinen *et al.*, Comparison of multivariate predictive power of major risk factors for coronary heart diseases in different countries: results from eight nations of the Seven Countries Study, 25-year follow-up, *J Cardiovasc Risk* **3** (1) (1996), pp. 69-75.

[2] X. Zhang, X.O. Shu, Y.-T. Gao, G. Yang, Q. Li, H. Li *et al.*, Soy food consumption is associated with lower risk of coronary heart disease in Chinese women, *The Journal of Nutrition* **133** (9) (2003), pp. 2874-2878.

[3] K.D. Setchelland A. Cassidy, Dietary isoflavones: biological effects and relevance to human health, *J Nutr* **129** (3) (1999), pp. 758S-767S.

[4] M.G. Hertog, D. Kromhout, C. Aravanis, H. Blackburn, R. Buzina, F. Fidanza *et al.*, Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study, *Arch Intern Med* **155** (4) (1995), pp. 381-386.

[5] A. Keys, A. Menotti, C. Aravanis, H. Blackburn, B.S. Djordevic, R. Buzina *et al.*, The seven countries study: 2,289 deaths in 15 years, *Prev Med* **13** (2) (1984), pp. 141-154.

[6] R. Tissier, X. Waintraub, N. Couvreur, M. Gervais, P. Bruneval, C. Mandet *et al.*, Pharmacological postconditioning with the phytoestrogen genistein, *J Mol Cell Cardiol* **42** (1) (2007), pp. 79-87.

[7] P. Zhai, T.E. Eurell, R.P. Cotthaus, E.H. Jeffery, J.M. Bahrand D.R. Gross, Effects of dietary phytoestrogen on global myocardial ischemia-reperfusion injury in isolated female rat hearts, *Am J Physiol Heart Circ Physiol* **281** (3) (2001), pp. H1223-1232.

[8] L. Al-Nakkash, B. Markus, K. Bowden, L.M. Batia, W.C. Prozialeckand T.L. Broderick, Effects of acute and 2-day genistein treatment on cardiac function and

ischemic tolerance in ovariectomized rats, *Gend Med* **6** (3) (2009), pp. 488-497.

[9] B. Deodato, D. Altavilla, G. Squadrito, G.M. Campo, M. Arlotta, L. Minutoli *et al.*, Cardioprotection by the phytoestrogen genistein in experimental myocardial ischaemia-reperfusion injury, *Br J Pharmacol* **128** (8) (1999), pp. 1683-1690.

[10] M. Maggiolini, D. Bonofiglio, S. Marsico, M.L. Panno, B. Cenni, D. Picard *et al.*, Estrogen receptor alpha mediates the proliferative but not the cytotoxic dose-dependent effects of two major phytoestrogens on human breast cancer cells, *Mol Pharmacol* **60** (3) (2001), pp. 595-602.

[11] G.G. Kuiper, J.G. Lemmen, B. Carlsson, J.C. Corton, S.H. Safe, P.T. van der Saag *et al.*, Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta, *Endocrinology* **139** (10) (1998), pp. 4252-4263.

[12] M.W. Carter, W.W. Smart, Jr.and G. Matrone, Estimation of estrogenic activity of genistein obtained from soybean meal, *Proc Soc Exp Biol Med* **84** (2) (1953), pp. 506-508.

[13] C.P. Baines, L. Wang, M.V. Cohenand J.M. Downey, Protein tyrosine kinase is downstream of protein kinase C for ischemic preconditioning's anti-infarct effect in the rabbit heart, *J Mol Cell Cardiol* **30** (2) (1998), pp. 383-392.

[14] I.F. Benter, J.S. Juggi, I. Khan, M.H. Yousif, H. Canatanand S. Akhtar, Signal transduction mechanisms involved in cardiac preconditioning: role of Ras-GTPase, Ca2+/calmodulin-dependent protein kinase II and epidermal growth factor receptor, *Mol Cell Biochem* **268** (1-2) (2005), pp. 175-183.

[15] R.M. Fryer, J.E. Schultz, A.K. Hsuand G.J. Gross, Pretreatment with tyrosine kinase inhibitors partially attenuates ischemic preconditioning in rat hearts, *Am J Physiol* **275** (6 Pt 2) (1998), pp. H2009-2015.

[16] R.M. Fryer, J.E. Schultz, A.K. Hsuand G.J. Gross, Importance of PKC and tyrosine kinase in single or multiple cycles of preconditioning in rat hearts, *Am J Physiol* **276** (4 Pt 2) (1999), pp. H1229-1235.

[17] S. Okubo, Y. Tanabe, K. Takeda, M. Kitayama, S. Kanemitsu, R.C. Kukreja *et al.*, Pretreatment with tyrosine kinase inhibitor attenuates the reduction of apoptosis 24 h after ischemic preconditioning, *Jpn J Physiol* **54** (2) (2004), pp. 143-151.

[18] S. Lee, G. Chanoit, R. McIntosh, D.A. Zvaraand Z. Xu, Molecular mechanism underlying Akt activation in zinc-induced cardioprotection, *Am J Physiol Heart Circ Physiol* **297** (2) (2009), pp. H569-575.

[19] M. Shilkrut, G. Yaniv, R. Asleh, A.P. Levy, S. Larischand O. Binah, Tyrosine kinases inhibitors block Fas-mediated deleterious effects in normoxic and hypoxic ventricular myocytes, *J Mol Cell Cardiol* **35** (10) (2003), pp. 1229-1240.

[20] K.L. Hamilton, S. Guptaand A.A. Knowlton, Estrogen and regulation of heat shock protein expression in female cardiomyocytes: cross-talk with NF kappa B signaling, *J Mol Cell Cardiol* **36** (4) (2004), pp. 577-584.

[21] T.E. Thingholm, O.N. Jensen, P.J. Robinsonand M.R. Larsen, SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides, *Mol Cell Proteomics* **7** (4) (2008), pp. 661-671.

[22] A. Keller, A.I. Nesvizhskii, E. Kolkerand R. Aebersold, Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search, *Anal Chem* **74** (20) (2002), pp. 5383-5392.

[23] S.A. Beausoleil, J. Villen, S.A. Gerber, J. Rushand S.P. Gygi, A probability-based

approach for high-throughput protein phosphorylation analysis and site localization, *Nat Biotechnol* **24** (10) (2006), pp. 1285-1292.

[24] S.J. Pitteri, P.A. Chrisman, J.M. Hoganand S.A. McLuckey, Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: reactions of doubly and triply protonated peptides with SO2*, *Anal Chem* **77** (6) (2005), pp. 1831-1839.

[25] S.J. Pitteri, P.A. Chrismanand S.A. McLuckey, Electron-transfer ion/ion reactions of doubly protonated peptides: effect of elevated bath gas temperature, *Anal Chem* **77** (17) (2005), pp. 5662-5669.

[26] D.L. Swaney, G.C. McAlister, M. Wirtala, J.C. Schwartz, J.E. Sykaand J.J. Coon, Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors, *Anal Chem* **79** (2) (2007), pp. 477-485.

[27] D.M. Good, C.D. Wenger, G.C. McAlister, D.L. Bai, D.F. Huntand J.J. Coon, Post-acquisition ETD spectral processing for increased peptide identifications, *J Am Soc Mass Spectrom* **20** (8) (2009), pp. 1435-1440.

[28] F. Diella, S. Cameron, C. Gemund, R. Linding, A. Via, B. Kuster *et al.*, Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins, *BMC Bioinformatics* **5** (2004), pp. 79.

Chapter 7

Conclusion Remarks and Future Directions

1. Project Significance and Contributions

Genistein (GEN) has been postulated to provide direct cytoprotection for cardiomyocytes against adverse stress. However, the molecular mechanism of such benefit of GEN is still uncharacterized. This dissertation aimed to use functional proteomics approaches to identify key downstream effectors of GEN which may mediate the cardioprotection. The merit of using proteomics is two-fold. First, it allowed us to comprehend the extent and the complexity of molecular action of GEN in cardiomyocytes from a global angle. Second, it allowed us to identify novel molecular targets of GEN that may otherwise be out of the scope of conventional reductionist investigations. Proteomics approaches are valuable for generating new hypotheses from which new research direction can be rationaled to identify cardioprotective action of GEN.

In our first study, we investigated the concentration-dependent proteome changes in cultured HL-1 cardiomyocytes in response to GEN treatments. We found at a

physiologically relevant concentration (1 μM), the expression of heat shock proteins and anti-apoptotic proteins was up-regulated. However, at a pharmacological concentration (50 μM), GEN down-regulated several glycolytic proteins and antioxidant enzymes. These concentration-dependent protein expression trends indicate a biphasic action of GEN that at low dose it can potentially protect the cardiomyocytes while at high dose it may make cardiomyocytes more susceptible to energy depletion and apoptosis during ischemic insults.

In our second study, we investigated the proteome changes in rat primary cardiomyocytes in response to ER-dependent and ER-independent actions of GEN. Our proteomics data suggested that GEN may have a critical influence on myocardial energy metabolism, in particular by up-regulation of glycolytic proteins via ER binding. We also found a novel anti-cardioprotective gene the soluble expoxide hydrolase can be down-regulated by GEN in an ER-dependent manner. Independent of ER binding, we identified novel molecular targets in steroidogenesis and estrogen signaling pathways that were down-regulated by GEN.

Although most proteomics data from this project need to be further validated, we believe several interesting hypotheses can be drawn based on the major discovery from our proteomics study to explain the cardioprotective role of GEN. First, we suspect that GEN provides cardioprotection only at low concentration but not at high pharmacological concentration. Second, at low concentration, GEN confers cardioprotection via the binding of ER. Third, at high concentration, GEN may

compromise the glycolytic pathway in an ER-independent fashion, making the cardiomyocytes more susceptible to energy depletion under ischemic stress. Fourth, at high concentration, GEN may trigger pro-apoptotic event in an ER-independent fashion making the cardiomyocytes more susceptible apoptosis under ischemic stress. To our best knowledge, this project is the first attempt to investigate the cellular effects of GEN in cardiomyocytes from a systems biology perspective.

In addition to interesting biological discoveries, this dissertation also focused on the development and optimization of proteomic applications to study protein expression in cardiomyocytes. Conventional 2DE protein separation was coupled with a two-stage hydrophilic and hydrophobic protein prefractionation method to provide broad coverage of cardiac proteome. Over 2300 protein spots were visualized by 2D gel. Further, a quantitative shotgun proteomic protocol was established coupling newly introduced HILIC separation and 8plex iTRAQ quantitation technology. A consensus-iterative searching strategy, that increased the sensitivity for protein identification was established with an independent house-written script to facilitate the iTRAQ quantitation. The whole protocol can be used for multiple parallel iTRAQ experiments to accommodate more samples. Pathway analyses including GeneGO MetaCore and Ingenuity IPA were applied to enrich pathways and cellular functions overrepresented by the differentially expressed protein influenced by GEN treatment in our experiment. These *in silico* pathway analyses have confirmed pathways that were significantly influenced by GEN such as anti-apoptosis, glycolysis, fatty acid metabolism pathways. We also took the next step to further explore the

phosphorylation patterns impacted by the GEN by developing a HILIC-SIMAC phosphopeptides fractionation and enrichment protocol.

2. Challenges and Unsolved Problems

Despite being used for cardiovascular research for nearly two decades, proteomics methodology still faces great challenges when applied to complex samples such as mammalian cells or tissues. Both mouse and rat genomes contain around 22,000 protein coding genes according to the gene ensemble database (www.ensembl.org). Additionally, the presence of differential slicing and post-translational modification makes the proteome even more complex. Such a large number of protein species has made it almost impossible to cover the whole proteome completely using the current 2DE or shotgun protocols. Therefore, further improvement on the resolving capacity of gel-based or LC-based separation techniques should result in expanded lists of proteins for analysis. Other than having huge variety of protein species, mammalian proteomes were also characterized by the huge differences in copy numbers among different proteins. Current 2DE-MS and shotgun-MS based protocols are usually biased against low abundant proteins many of which are involved with key signaling and regulatory processes. Our shotgun proteomics data and phosphoproteomics data suggested that highly abundant proteins, in particular the myofibril proteins such as myosin can dramatically decrease the discovery power of proteomics methods. In our experiments this was more problematic in primary cardiomyocytes than in the

immortalized cardiomyocyte line. Therefore, it is highly recommended for future proteomics studies on cardiomyocytes, that an immunodepletion step be established to prefractionate the abundant proteins from the rest of the less abundant proteins. Mammalian proteomes also contain large numbers of protein isoforms which in these experiments caused some incidences of identification ambiguity. Both 2DE and shotgun proteomics workflows typically identify proteins only with partial sequence coverage and also depend on protein databases which still contain inaccurate sequence information especially on protein isoforms. Such identification ambiguity may restrict the specificity of subsequent functional interpretation. It is always challenge to differentiate protein isoforms which rely on the identification of unique peptides especially in shotgun experiments. We believe a thorough separation of peptide species can improve the chance to detect unique peptides to distinguish protein isoforms.

The dynamic range of protein expression alteration in cells can be huge. However, our experience indicating the dynamic range of proteomics quantitation using 2DE or iTRAQ was compressed. Gel-based quantitation suffers from high variability, uneven image background, spots co-migration and low dynamic range of staining protocol, in particular the silver stain protocol. Shotgun based iTRAQ technology also has the problem with co-elution of similar m/z species in LC-MS analysis. All those disadvantages may interfere with quantitation and generate weak statistics to discover biological differences between samples. Therefore, it is vital to use other method such as ELISA, western blotting to confirm the protein expression changes quantified from

a proteomics study. In recent years, MS-based validation strategy described as multiple reaction monitoring technology (MRM) has been evaluated as an alternative approach to ELISA or western blotting. The MRM typically operates on a triple quadrupole mass spectrometry and have advantages of high dynamic range, high specificity, high throughput and no need for antibodies.

Identification and quantitation of protein phosphorylation in large-scale is a challenging task. Key technical difficulty we encountered was in simplifying the proteome prior to phosphopeptide enrichment. Our experience suggests that simply separating the whole lysate digestion into 15 fractions may be insufficient for effective enrichment. Based on these experiences, we recommended using immunodepletion to remove abundant contractile proteins prior to phosphopeptide enrichment. Second, we found that the SIMAC enrichment can only be used in a semi-quantitative study with limited selectivity, reproducibility and linearity. We consider all other types of IMAC or MOAC based enrichment techniques may also have poor quantitative performance as non-specific binding is a ubiquitous side-effect in all enrichment protocols. However, we believe with a simplified loading background, one might be able to improve the quantitative performance of metal based enrichment. Third, the choice of mass spectrometry for phosphopeptides is critical as no instrument type or collision mode has shown absolute superiority over others. ETD instrument was still considered one of the most promising strategies to analyze phosphopeptides as it prevents the neutral loss of phosphor groups from the peptides during peptide backbone fragmentation. However, the interpretation of ETD

data is challenging and the understanding of the knowledge about the fragmentation behavior of peptides in ETD is still incomplete. Further, current database search engines are designed to interpret CID data with recent adaptation for ETD data. However, very little is known about their performance on ETD data interpretation so far.

A final challenge in proteomics research is the data interpretation. Proteomics *per se* only measures the relative expression level of multiple protein species. These –omics type of data usually are noisy and contain many unexpected discoveries such as protein never been characterized before. Even proteins that have been extensively studied are often involved in multiple cellular processes that may make the biological interpretation of proteomic data unspecific. Therefore, in addition to thorough literature reviews on each identified protein, a meta-analysis approach using MetaCore or IPA pathway analysis package was carried out to highlight important pathways in which differentially expressed proteins play biological roles. However, one limitation to these softwares is that they were initially written for mRNA microarray data which usually has much wider coverage of gene expression profile than proteomics. Our experience suggested that proteomics data are relative sparse and usually generate unspecific pathway analysis results with weak statistics. Moreover, both literature review on single protein or meta-analysis approach can only serve to generate researchable hypotheses for future investigation. Therefore, proteomics analyses should ideally be accompanied by mechanistic studies to test the hypotheses generated.

3. Future Directions

Given what we have learned from the current proteomics projects, mechanistic study can be launched to investigate the GEN cardioprotection in the following direction:

1, confirm in which concentration/dose range does the GEN provide cardioprotection against hypoxic/ischemic stress. Proteomics approach can further help to locate key cardioprotective proteins modulated by GEN during the ischemic/hypoxic stress.

2, identify whether ER binding is a key intermediate step for GEN cardioprotection against hypoxic/ischemic stress. Proteomics approach can also help to select key cardioprotective proteins modulated by GEN during the ischemic/hypoxic stress via ER-dependent mechanism.

3, identify whether glycolytic activity of cardiomyocytes can be inhibited by high concentration of GEN. Other type of pan-specific PTK inhibitor and isoflavones with no PTK inhibitory effects such as daidzein can be used for comparison. Radiolabeling can be used for glycolytic flux analysis. Possible expression changes of glycolytic proteins can be measured via MRM approach.

4, identify whether pro-apoptotic pathway in cardiomyocytes can be activated by high concentration of GEN in an ER-independent fashion. Again, other type of pan-specific PTK inhibitor and daidzein can be used for comparison.

5, identify phosphorylation pattern changes due to the ER-independent action of GEN in cardiomyocytes. Improvement on phosphopeptides enrichment technique and mass spectrometry data interpretation has to be made in order to identify and quantify protein phosphorylation successfully. Alternatively, since GEN is a pan-specific PTK inhibitor, it will also be interesting to use Tyr-specific immunoprecipitation to focus the phosphoproteomics survey on the direct impact of GEN on tyrosine kinase signaling. Differentially phosphorylated sites can also be confirmed by MRM approach.

Other than those specific aims for the future studies, it will also be interesting to study the proteome changes triggered by other isoflavone, soy phytoestrogen extract or isoflavone cocktails in cardiomyocytes. Differential proteomics can also be taken to contrast the GEN responses in cardiomyocytes with gender differences.

Abbreviations and Acronyms

2DE	2-dimensional electrophoresis
ABC	ammonium bicarbonate
ACN	acetonitrile
AMT	accurate mass and retention time tag
BSA	bovine serum albumin
CID	collision induced dissociation
CVD	cardiovascular disease
DDA	data dependent acquisition
DHB	2,5-dihydroxybenzoic acid
DTT	dithiothreitol
ER	estrogen receptor
ERA	estrogen receptor antagonist
Erk1/2	p42/p44 extracellular signal-regulated kinases
ERLIC	electrostatic repulsion hydrophilic interaction chromatography
ESI-MS/MS	electrospray ionization tandem mass spectrometry
ETD/ECD	electron transfer/capture dissociation
FA	formic acid
FDR	false discovery rate
GEN	genistein
GO	gene ontology
HCD	higher-energy C-trap dissociation
HILIC	hydrophilic interaction liquid chromatography
I/R	ischemia/reperfusion
IAA	iodoacetamide
ICAT	isotope-coded affinity tags
IEF	iso-electric focusing
IMAC	immobilized metal affinity chromatography
IPG	immobilized pH gradient
IPI	international protein index
IT	ion trap mass spectrometry
iTRAQ	isobaric tag relative and absolute quantitation

m/z	mass to charge ratio
MALDI-MS/MS	matrix assisted laser desorption/ionization ionization tandem mass spectrometry
MAPKs	mitogen-activated protein kinases
MDLC	multidimensional liquid chromatography
MOAC	metal oxide affinity chromatography
MSA	multi-stage activation
MudPIT	multidimentional protein identification technology
PAC	phosphoramidate chemistry
PI3K	phosphatidylinositol-3-OH kinase
PQD	pulsed-Q dissociation
Q-tof	quadrupole time of flight mass spectrometry
RPLC	reverse phase liquid chromatography
SAX	strong anion exchange
SCX	strong cation exchange chromatography
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC	stable isotope labeling by amino acids in cell culture
SIMAC	sequential immobilized metal affinity chromatography
TFA	trifluoroacetic acid
TK	tyrosine kinase
TMT	tandem mass tag

Appendix I

Table of differentially expressed proteins in Chapter 3

Differentially altered proteins identified in 1μM and 50 μM GEN treated cardiomyocytes compared with control cells. Proteins were grouped according to their major biological functions. Spots quantification was evaluated as the averaged fold-changes ± pooled standard deviation based on percent spot volume (% voli = voli/∑voln), NS means no significant changes. The prefix of spot ID, the ‘phi’ or ‘pho’, designates the spot origin from hydrophilic or hydrophobic gel.

Spot ID	Protein identification	IPI IDs	unique peptides	MW	GEN 1 vs Control	GEN 50 vs Control	Biological Process
Stress response and protein folding machinery							
phi_32	Heat-shock protein beta-1 (HspB1) (Heat shock 27 kDa protein) (HSP 27) (Growth-related 25 kDa protein) (P25) (HSP25)	IPI00128522	6	23057	NS	3.98±0.01	anti-apoptosis/response to stress/apoptosis and survival FAS signaling/signal transduction ESR-1 pathway
phi_63			5		NS	1.36±0.04	
phi_65			3		1.33±0.03	NS	
pho_32			2		-1.36±0.09	NS	
pho_28	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP)	IPI00319992	6	72492	NS	1.53±0.02	Signal transduction ESR 1 pathway/anti-apoptosis/cellular response to glucose starvation/negative regulation of caspase activity
phi_51	Alpha-synuclein (Non-A beta component of AD amyloid) (Non-A4 component of amyloid precursor) (NACP)	IPI00115157	3	14476	NS	1.30±0.06	cellular response to oxidative stress/fatty acid metabolic process/mitochondrial ATP synthesis coupled electron transport
phi_07	Stress-70 protein, mitochondrial precursor (75 kDa glucose-regulated protein) (GRP 75)	IPI00133903	3	73768	2.06±0.03	NS	protein folding/response to stress
phi_08			7		NS	-2.10±0.01	
phi_20			5		NS	2.96±0.02	
phi_74			2		1.62±0.03	NS	
pho_01	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)	IPI00323357	7	71055	1.57±0.01	2.12±0.02	protein folding/chaperone cofactor-dependent protein refolding/ response to stress
phi_31			11		NS	-2.17±0.03	
phi_37			3		-1.70±0.02	-5.26±0.02	
phi_53			2		1.75±0.03	NS	
phi_62			9		NS	1.61±0.06	
phi_27	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock protein 60) (HSP-60) (Mitochondrial matrix protein P1) (HSP-65)	IPI00308885	19	61088	NS	4.42±0.04	Protein folding/activation of caspase activity/anti-apoptosis
pho_27			14		NS	1.64±0.02	
phi_03	Phosphatidylethanolamine-binding protein 1 (PEBP-1)	IPI00137730	3	20988	NS	2.25±0.06	aging/response to oxidative stress
phi_56			2		NS	-1.64±0.02	
phi_40			3		NS	-2.29±0.04	
phi_02	Nascent polypeptide-associated complex subunit alpha, muscle-specific form (Alpha-NAC, muscle-specific form)	IPI00121297	2	221379	-2.22±0.03	NS	Protein transport/transcription/transcription regulation

pho_06	Protein NDRG1 (N-myc downstream-regulated gene 1 protein) (Protein Ndr1)	IPI00125960	5	43437	1.38±0.03	NS	Response to metal ion/response to hypoxia
Cellular redox balance							
pho_29	Peroxiredoxin-4 (EC 1.11.1.15) (Prx-IV) (Thioredoxin peroxidase AO372) (Thioredoxin-dependent peroxide reductase A0372) (Antioxidant enzyme AOE372)	IPI00116254	4	31261	-1.31±0.02	NS	cell redox homeostasis/oxidation reduction
phi_11	Glutathione synthetase (EC 6.3.2.3) (Glutathione synthase) (GSH synthetase) (GSH-S)	IPI00127691	2	52442	NS	3.65±0.02	Glutathione biosynthetic process/response to cadmium ion
phi_24	Protein disulfide-isomerase precursor (EC 5.3.4.1) (PDI) (Prolyl 4-hydroxylase subunit beta)	IPI00133522 /IPI00122815	2	57507	NS	-3.39±0.02	cell redox homeostasis
phi_54	Protein disulfide-isomerase A3 precursor (EC 5.3.4.1) (Disulfide isomerase ER-60) (ERp60) (58 kDa microsomal protein) (p58) (ERp57)	IPI00230108	2	57042	NS	-6.00±0.02	cell redox homeostasis/positive regulation of apoptosis
phi_78			6		NS	-1.56±0.02	
Apoptosis							
pho_17	Voltage-dependent anion-selective channel protein 2 (VDAC-2) (mVDAC2) (mVDAC6)	IPI00122547	4	32340	NS	1.36±0.06	anion transport/survival regulation of apoptosis
pho_18			3		1.52±0.03	NS	
pho_21	BAG family molecular chaperone regulator 2 (BCL2-associated athanogene 2) (BAG-2)	IPI00130304	6	23630	NS	-1.72±0.02	Apoptosis
phi_66	Cathepsin D precursor	IPI00111013	2	45381	NS	1.61±0.04	Proteolysis
pho_08	Import inner membrane translocase subunit TIM50, mitochondrial precursor	IPI00111045	5	39980	1.38±0.04	NS	mitochondrial membrane organization/protein amino acid dephosphorylation/protein transport/transmembrane transport
Energy metabolism							
phi_79	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial precursor	IPI00130804	4	36437	NS	1.52±0.02	fatty acid metabolic process
phi_52	Phosphoglycerate kinase 1	IPI00555069	2	44907	NS	-1.59±0.06	glycolysis/phosphorylation/gluconeogenesis
phi_69			4		NS	-1.99±0.02	
phi_72			5		NS	-1.47±0.02	

phi_73			3		1.27±0.02	NS	
pho_30	Pyruvate dehydrogenase E1 component alpha subunit, somatic form, mitochondrial precursor (EC 1.2.4.1) (PDHE1-A type I)	IPI00337893	3	43888	1.59±0.02	2.34±0.03	glycolysis/oxidation reduction
phi_70	Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1)	IPI00462072	7	47453	NS	-1.72±0.06	glycolysis
phi_26	Pyruvate kinase isozyme M2	IPI00407130	2	58420	NS	-3.06±0.05	glycolysis
phi_71			8		NS	-1.63±0.03	
phi_42	Triosephosphate isomerase	IPI00467833	4	27038	NS	-5.01±0.02	Glycolysis/pentose-phosphate shunt/gluconeogenesis/fatty acid biosynthetic process
phi_81	Ribose-5-phosphate isomerase (EC 5.3.1.6) (Phosphoriboisomerase)	IPI00113408	3	26098	-1.39±0.02	NS	pentose-phosphate shunt
phi_19	Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A) (Glutamate oxaloacetate transaminase 1)	IPI00877205	5	46488	NS	2.33±0.01	cellular amino acid metabolic process/fatty acid homeostasis
phi_64	Electron transfer flavoprotein subunit beta (Beta-ETF)	IPI00121440	3	27834	NS	1.65±0.02	electron transport chain/transport
phi_35	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial precursor (EC 1.6.5.3) (EC 1.6.99.3) (NADH-ubiquinone oxidoreductase 24 kDa subunit)	IPI00169925	9	27640	NS	-3.59±0.04	electron transport chain/transport
pho_35			3		NS	-1.24±0.04	
DNA integrity and RNA transcription/translation							
pho_04	Heterogeneous nuclear ribonucleoprotein H (hnRNP H)	IPI00133916	4	49454	NS	1.58±0.01	RNA splicing/mRNA processing
pho_07			7		1.48±0.04	NS	
pho_38	Elongation factor 1-beta (EF-1-beta)	IPI00320208	3	24849	NS	1.44±0.06	translational elongation
phi_55	Prohibitin (B-cell receptor-associated protein 32) (BAP 32)	IPI00133440	2	29859	NS	-1.51±0.02	DNA replication
pho_19			2		NS	-2.19±0.02	
pho_12	60S acidic ribosomal protein P0 (L10E)	IPI00314950	5	34366	NS	1.44±0.07	ribosome biogenesis/translational elongation
pho_05	RuvB-like 2 (EC 3.6.1.-) (p47 protein)	IPI00123557	10	51252	NS	1.74±0.03	DNA recombination/DNA repair/regulation of transcription
pho_15	Proliferating cell nuclear antigen	IPI00113870	3	29108	-2.19±0.05	NS	DNA replication/regulation of DNA

	(PCNA)						replication/intracellular protein transport
phi_77	Heterogeneous nuclear ribonucleoprotein A/B (hnRNP A/B) (CArG-binding factor-A) (CBF-A)	IPI00117288	4	30926	-1.97±0.02	NS	positive regulation of gene-specific transcription/transcription
Cytoskeletal remodeling mobility/contractility							
phi_18	Vimentin	IPI00227299	6	53712	NS	2.08±0.02	intermediate filament-based process
phi_57			12		NS	1.61±0.05	
pho_41			2		NS	-1.68±0.03	
pho_13			4		NS	-2.00±0.05	
pho_14	Tropomyosin-1 alpha chain (Alpha-tropomyosin)	IPI00123316/IPI00830701	9	32718	1.32±0.09	NS	cardiac muscle contraction/positive regulation of heart rate by epinephrine/ventricular cardiac muscle tissue morphogenesis
pho_16			6		NS	-1.61±0.03	
phi_58	Tropomyosin beta chain (Tropomyosin 2) (Beta-tropomyosin)	IPI00874728	3	32931	NS	-1.47±0.03	muscle contraction
phi_10	Actin-related protein 2/3 complex subunit 5 (ARP2/3 complex 16 kDa subunit)	IPI00399943	4	16335	NS	2.27±0.03	Regulation of actin filament polymerization
pho_10			2		1.52±0.07	NS	
pho_24	Myosin regulatory light chain 9	IPI00750595	4	19898	NS	1.79±0.11	motor activity/calcium ion binding
phi_21	Tubulin beta-5 chain	IPI00117352	2	50095	-2.58±0.02	NS	microtubule-based movement/protein polymerization/spindle assembly
phi_12	Myosin-6 (Myosin heavy chain 6) (Myosin heavy chain, cardiac muscle alpha isoform) (MyHC-alpha)	IPI00129404	2	224225	NS	6.47±0.01	cardiac muscle fiber development/regulation of ATPase activity/regulation of heart rate/ regulation of the force of heart contraction/sarcomere organization
pho_26			8		1.49±0.04	NS	
phi_34	Myosin light polypeptide 4 (Myosin light chain 1, atrial/fetal isoform) (MLC1A) (MLC1EMB)	IPI00331411/IPI00318901	2	21260	NS	-2.18±0.02	N/A
phi_47	Myosin light polypeptide 6 (Smooth muscle and nonmuscle myosin light chain alkali 6) (Myosin light chain alkali 3) (Myosin light chain 3) (MLC-3)	IPI00354819	2	17090	NS	1.39±0.06	muscle filament sliding
pho_43	Calponin-3	IPI00119111	4	36577	NS	-1.47±0.03	actomyosin structure organization
phi_75	Vinculin (Metavinculin)	IPI00405227	8	117215	NS	-1.80±0.02	cell adhesion
Protein phosphorylation and cellular signaling							

phi_39	Low molecular weight phosphotyrosine protein phosphatase	IPI00134135	2	18636	2.14±0.02	2.34±0.02	protein amino acid dephosphorylation
pho_20	Growth factor receptor-bound protein 2 (Adapter protein GRB2) (SH2/SH3 adapter GRB2)	IPI00119058	3	25336	NS	-2.59±0.02	MAPKKK cascade/Ras protein signal transduction/cell differentiation
pho_39	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein 1) (KCIP-1) (SEZ-2)	IPI00116498	2	27925	-3.67±0.02	1.98±0.03	protein targeting/anti-apoptosis/signal transduction/mRNA metabolic process
pho_42	COP9 signalosome complex subunit 4 (Signalosome subunit 4) (SGN4) (JAB1-containing signalosome subunit 4)	IPI00131871	3	46541	NS	1.53±0.04	Signal transduction
				others			
phi_49	Uncharacterized protein C15orf38 homolog	IPI00461011	2	65696	NS	1.33±0.02	N/A
pho_31	Ubiquitin carboxyl-terminal hydrolase isozyme L5	IPI00124938	4	37878	-1.75±0.02	NS	ubiquitin-dependent protein catabolic process
pho_36	COMM domain-containing protein 3 (Bmi-1 upstream gene protein) (Bup protein)	IPI00227640	4	22308	NS	-1.98±0.03	N/A

Appendix II

Protein identification reports for Chapter 3

Mascot identification reports for differentially expressed proteins identified from HL-1 cardiomyocytes in response to low and high concentration of GEN treatment. Each Mascot report was generate using the Select Summary (protein hits) format. Each PDF file was named after the spots location where the ‘phi’ and ‘pho’ means protein spots from hydrophilic and hydrophobic fraction, respectively.

Available online: http://www.dropbox.com
User name: zeyusun@engr.colostate.edu
Password: proteomics
File name: Appendix 2. rar
Note: Appendix 2 rar package contains all Mascot protein identification in PDF format. Use WinRAR or WinZip for file extraction.

Appendix III

Table of GO processes relavent to differentially expressed proteins in Chapter 3

GeneGO enriched GO processes that associated with the differentially expressed proteins in response to 1 μM (Part A) and 50 μM Genistein (Part B) treatement. Detailed description of GO processes can be found in GeneGO metacore database (http://www.genego.com/genego_lp.php). A hypergeometric distribution probability test was carried out to decide if the experimentally identified proteins (objects) were involved with a particular biological pathway in the GeneGO MetaCore database. The p-value represents the odds of having a given pre-built network be associated with the list of experimentally- identified proteins simply by chance.

Part A

#	Processes	pValue
1	regulation of ATPase activity	1.234E-07
2	sarcomere organization	1.234E-07
3	small molecule catabolic process	2.331E-07
4	organelle organization	3.815E-07
5	muscle filament sliding	4.387E-07
6	ventricular cardiac muscle tissue morphogenesis	4.764E-07
7	actin-myosin filament sliding	4.764E-07
8	regulation of heart rate	5.165E-07
9	ventricular cardiac muscle tissue development	5.165E-07
10	myofibril assembly	5.591E-07
11	actin-mediated cell contraction	6.520E-07
12	actomyosin structure organization	9.341E-07
13	cardiac ventricle morphogenesis	1.142E-06
14	cardiac muscle tissue morphogenesis	1.298E-06
15	muscle tissue morphogenesis	1.470E-06
16	positive regulation of heart rate by epinephrine	1.583E-06
17	cytoskeleton organization	1.682E-06
18	actin filament-based movement	1.759E-06
19	anti-apoptosis	1.823E-06
20	glucose catabolic process	2.088E-06

Part B

#	Processes	pValue
1	muscle filament sliding	9.583E-20
2	actin-myosin filament sliding	1.227E-19
3	actin-mediated cell contraction	3.122E-19
4	actin filament-based movement	5.800E-18
5	regulation of ATPase activity	8.730E-17
6	muscle contraction	3.822E-16
7	muscle system process	2.632E-15
8	glycolysis	8.670E-13
9	actomyosin structure organization	1.001E-12
10	glucose catabolic process	5.283E-12
11	actin filament-based process	7.591E-12
12	ventricular cardiac muscle tissue morphogenesis	2.270E-11
13	ventricular cardiac muscle tissue development	2.626E-11
14	hexose catabolic process	2.758E-11
15	myofibril assembly	3.030E-11
16	cardiac muscle tissue development	3.178E-11
17	monosaccharide catabolic process	3.967E-11
18	small molecule catabolic process	4.854E-11
19	cardiac ventricle morphogenesis	1.094E-10

20	cardiac muscle tissue morphogenesis	1.377E-10
21	muscle tissue morphogenesis	1.721E-10
22	generation of precursor metabolites and energy	2.088E-10
23	alcohol catabolic process	2.256E-10
24	muscle organ morphogenesis	3.920E-10
25	cellular carbohydrate catabolic process	4.435E-10
26	cardiac muscle contraction	5.032E-10
27	glucose metabolic process	7.174E-10
28	cardiac ventricle development	8.902E-10
29	muscle tissue development	9.046E-10
30	regulation of system process	9.497E-10
31	heart contraction	1.039E-09
32	heart process	1.039E-09
33	cardiac chamber morphogenesis	2.152E-09
34	carbohydrate catabolic process	2.248E-09
35	tissue development	2.629E-09
36	muscle organ development	2.783E-09
37	positive regulation of ATPase activity	2.901E-09
38	cardiac chamber development	3.331E-09
39	hexose metabolic process	5.048E-09
40	cardiac muscle fiber development	8.572E-09

41	regulation of the force of heart contraction	1.278E-08
42	cellular component assembly involved in morphogenesis	1.872E-08
43	sarcomere organization	2.019E-08
44	monosaccharide metabolic process	2.232E-08
45	regulation of muscle contraction	3.482E-08
46	regulation of heart contraction	3.482E-08
47	muscle cell development	3.840E-08
48	regulation of hydrolase activity	4.058E-08
49	catabolic process	5.589E-08
50	gluconeogenesis	7.213E-08
51	regulation of muscle system process	8.726E-08
52	small molecule metabolic process	9.361E-08
53	anatomical structure morphogenesis	9.625E-08
54	muscle structure development	1.138E-07
55	regulation of heart rate	1.225E-07
56	actin cytoskeleton organization	1.501E-07
57	regulation of nucleotide catabolic process	1.536E-07
58	regulation of purine nucleotide catabolic process	1.536E-07
59	alcohol biosynthetic process	1.608E-07
60	regulation of biological quality	1.650E-07
61	tissue morphogenesis	1.942E-07

62	hexose biosynthetic process	2.161E-07
63	cellular component organization at cellular level	2.488E-07
64	cerebellar Purkinje cell layer development	2.699E-07
65	heart morphogenesis	2.815E-07
66	ATP metabolic process	2.860E-07
67	cellular component organization	2.996E-07
68	cellular component organization or biogenesis at cellular level	4.730E-07
69	cellular carbohydrate metabolic process	5.238E-07
70	cellular component organization or biogenesis	5.659E-07
71	heart development	7.200E-07
72	monosaccharide biosynthetic process	7.538E-07
73	cell projection organization	7.684E-07
74	cellular component morphogenesis	8.839E-07
75	blood circulation	9.613E-07
76	response to external stimulus	9.621E-07
77	circulatory system process	9.831E-07
78	alcohol metabolic process	1.178E-06
79	atrial cardiac muscle tissue morphogenesis	1.420E-06
80	atrial cardiac muscle tissue development	1.420E-06
81	pyruvate metabolic process	1.466E-06
82	myosin filament assembly	1.949E-06

83	elastic fiber assembly	1.949E-06
84	myosin filament assembly or disassembly	1.949E-06
85	regulation of cellular component size	2.138E-06
86	cerebellar cortex development	2.258E-06
87	cellular component biogenesis	2.323E-06
88	cellular component assembly at cellular level	2.881E-06
89	exocytosis	3.258E-06
90	regulation of cellular catabolic process	3.486E-06
91	cytoskeleton organization	3.654E-06
92	cellular component assembly	3.822E-06
93	response to wounding	4.116E-06
94	extracellular matrix assembly	4.279E-06
95	muscle cell differentiation	4.522E-06
96	regulation of nucleotide metabolic process	5.130E-06
97	regulation of protein folding in endoplasmic reticulum	5.338E-06
98	positive regulation of heart rate by epinephrine	5.338E-06
99	anti-apoptosis	5.781E-06
100	response to chemical stimulus	6.076E-06
101	cardiac myofibril assembly	6.561E-06
102	protein refolding	6.561E-06
103	regulation of anatomical structure size	6.812E-06

104	multicellular organismal development	6.888E-06
105	anatomical structure development	7.175E-06
106	monocarboxylic acid metabolic process	7.184E-06
107	organ development	7.808E-06
108	carbohydrate metabolic process	8.168E-06
109	negative regulation of caspase activity	8.213E-06
110	cellular carbohydrate biosynthetic process	8.455E-06

Appendix IV

Table of identified peptides in Chapter 4

Complete list of all peptides identified for 509 quantified proteins in Chapter 4. Database search parameters of both Mascot and X !Tandem were included.

Available online: http://www.dropbox.com
User name: zeyusun@engr.colostate.edu
Password: proteomics
File name: Appendix 4. xls

Appendix V

Table of differentially expressed protein in Chapter 4

Differentially altered proteins identified using pair wise t-test. All protein IDs were categorized into 5 groups according to the expression patterns implicating possible ER-dependent or –independent mechanism by which GEN triggers the expression alternation. Note, the MW refer to molecular weight of the protein, and GO annotation on biological process was extract from the public Uniprot database, NS means no significant changes was observed. The peptides/spectra column was format as the unique number of peptide ID from run A + run B| the total number of assigned spectra from run A + run B.

Protein	Accession #	MW	Peptides/Spectra	Biological Process	Gen/Con	Gen+ERA/ Con	Gen/ Gen+ERA
ER-dependent only							
Carnitine O-palmitoyltransferase 2, mitochondrial	IPI00195593	74,094.50	7+8|53+60	fatty acid beta-oxidation, long-chain fatty acid transport	1.08	NS	1.08
Mitochondrial carnitine/acylcarnitine carrier protein	IPI00205413	33,137.30	2+2|42+71	transmembrane transport	-1.21	NS	-1.27
RCG20659, isoform CRA_b	IPI00766463	38,086.30	1+1|3+5	response to stress	-1.31	NS	-1.37
ER-dependent, possibly with ER-independent in synergism							
Fructose-bisphosphate aldolase A	IPI00231734	39,334.50	11+12|119+107	glycolysis, protein homotetramerization, response to estrogen stimulus, response to heat, response to hypoxia, response to lipopolysaccharide,response to nicotine	1.08	NS	NS
alpha-enolase	IPI00464815	47,111.00	12+12|209+173	glycolysis	1.22	NS	NS
myosin-6	IPI00189809	223,508	141+134|9578+8589	muscle myosin complex, myofibril,myosin filament, perinuclear region of cytoplasm	-1.04	NS	NS
ATP synthase subunit b, mitochondrial	IPI00196107	28,851.50	9+7|121+85	ATP synthesis coupled proton transport	-1.07	NS	NS
Elongation factor 1-α 2	IPI00325281	50,436.60	6+5|48+92	respons to inorganic substance	-1.13	NS	NS
Epoxide hydrolase 2	IPI00195735	62,323.80	1+1|2+1	aromatic compound catabolic process, linoleic acid metabolic process, positive regulation of blood pressure, prostaglandin production involved in inflammatory response, response to toxin, sensory perception of pain	-1.38	NS	NS
RCG62645, isoform CRA_b	IPI00372407	30,873.50	1+1|1+1	N/A	-1.41	NS	NS
T-complex protein 1 subunit delta	IPI00337168	58,083.00	1+1|2+3	chaperone mediated protein folding requiring cofactor	-2.07	NS	NS
ER-independent and ER-dependent in antagonism							
3-hydroxyacyl-CoA dehydrogenase type-2	IPI00886470	28343.9	9+6|141+173	cell aging, oxidation-reduction process, protein homotetramerization, tRNA processing	NS	-1.05	1.05
ER-independent only							
Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	IPI00395281	68,181.00	5+5|57+63	electron transport chain, response to oxidative stress, transport	1.14	1.14	NS
similar to AHNAK nucleoprotein isoform 1	IPI00207069	27,442	1+1|14+5	N/A	-1.17	-1.21	NS
MACRO domain-containing protein 1	IPI00203232	35,290.20	2+1|4+2	N/A	-1.59	-1.54	NS
ER-independent, possibly with ER-dependent in antagonism							
Stomatin-like protein 2	IPI00203528	38396	1+1|2+1	N/A	NS	1.53	NS
Protein	IPI00778558	114029.4	1+1|2+1	isoleucyl-tRNA aminocylation	NS	1.23	NS

similar to isochorismatase domain containing 2 isoform 1	IPI00764444	24,651.70	2+3|4+18	N/A	NS	1.10	NS
similar to Nebulette	IPI00565256	116173.6	2+3|23+20	NA	NS	1.05	NS
60S ribosomal protein L13	IPI00230916	24,292.50	1+1|2+1	translation	NS	1.05	NS
similar to titin isoform N2-B	IPI00554003	***1409058***	81+76|594+847	N/A	NS	-1.05	NS
Isoform Mitochondrial of Fumarate hydratase, mitochondrial	IPI00231611	54,446.30	9+10|174+217	fumarate metabolic process, malate metabolic process, tricarboxylic acid cycle	NS	-1.11	NS
Phosphoglycerate mutase 1	IPI00421428	28,814.80	2+2|14+22	glycolysis	NS	-1.11	NS
Heat shock protein 75 kDa, mitochondrial	IPI00369217	80,445.90	1+1|1+1	protein folding	NS	-1.25	NS
Trimeric intracellular cation channel type A	IPI00189667	33,396.10	3+2|39+73	ion transport, potassium transport, transport	NS	-1.61	NS
hypothetical protein LOC690102	IPI00569279	13,981.80	1+1|7+3	nucleosome assembly	NS	-1.87	NS

Appendix VI

MZmine2 LC-MS data processing protocol

This protocol was written for SIMAC LCMS data processing using MZmine2. Data was collected from Agilent 6150 Q-tof. For data collected from other sources, it is highly remembered to check the following aspect of raw data: 1, mass accuracy, 2, chromatogram baseline and noise, average chromatogram peak span and intensity, 3, mass spectrum baseline and noise, overall peak intensity. 4, retention time range where most of the peptides elute out.

1 Data loading
Format: mzData, mzXML, mzML

2 Crop filter: RT 0:40 – 14 min

3 Chromatogram builder setting:
Centroid mode
Noise level: 2000 counts
Filtering: None
Chromatogram Construction: highest data point
Minimum time Span: 0:15 min,
Minimum height 3000 counts,
m/z tol: 0.02

4 chromatogram deconvolution:
Use the local minimum search algorithm
Chromatographic threshold: 70%,
Search minimum in RT range: 0.1
Minimum relative height: 5%,
Minimum absolute height: 100
Minimum ratio of peak top/edge: 2,
Remove source peak list after filtering? yes

5 Isotopic peaks grouper setting:
m/z tol: 0.02
RT tol: 0:10
Monotonic shape? No
Maximum charge= 5
Representative isotope: highest intensity
Remove source peak list after filtering? yes

6 Peak list rows filter setting:
Minimum peak in a row: 1
Minimum peaks in an isotope pattern: 1
Min/Max m/z: 400/1300
Min/Max RT: 0:40/14:00 min
Remove source peak list after filtering? Yes

7 RANSAC aligner setting,
m/z tol: 0.02 da
RT tol: 1:00 (before alignment)
RT tol after correction: 0:15 min
RANSAC iteration: 10
Minimum number of points: 50%
Threshold value: 0:15 min
Linear model? No
Require same charge state? Yes

Appendix VII

Table of identified peptides in Chapter 5

This table contains the complete list of all peptides identified from tryptic caseins and BSA mixture in Chapter 5.

sequence	z	m/z	score	ΔR1-R2	%SPI	RT(min)	Protein
(K)HIQKEDVPSER(Y)	3	446.5659	16.22	8.38	90.4	3.33	α-S1-casein
(K)SCQAQPTTMAR(H)	2	625.7847	13.25	4.28	79.7	3.53	κ-casein
(K)AVPYPQR(D)	2	415.7299	13.95	8.7	94.3	3.76	β-casein
(K)VLPVPQK(A)	2	390.7546	11.4	3.77	75.8	4.2	β-casein
(K)FQsEEQQQTEDELQDK(I)	2	1031.419	20.9	20.9	74.9	4.3	β-casein
(K)FQsEEQQQTEDELQDK(I)	3	687.9475	16.09	16.09	80.5	4.31	β-casein
(K)TVDmEsTEVFTK(K)	2	741.8061	17.33	6.89	84.5	4.49	α-S2-casein
(K)AEFVEVTK(L)	2	461.7479	15.08	15.08	76.4	4.64	Serum albumin
(K)VIPYVR(Y)	2	373.7308	15.35	7.94	97.1	4.99	α-S2-casein
(K)TVDMEsTEVFTKK(T)	2	797.8547	14.88	7.88	79.1	5.35	α-S2-casein
(R)NAVPITPTLNR(E)	2	598.3423	15.14	4.75	89.6	5.52	α-S2-casein
(K)DIGsEsTEDQAMEDIK(Q)	2	964.3484	18.41	18.41	63.5	5.64	α-S1-casein
(K)DIGsEsTEDQAMEDIK(Q)	3	643.2349	13.88	5.9	76.6	5.66	α-S1-casein
(K)TVDMEsTEVFTK(K)	2	733.8091	15.01	8.55	72	5.82	α-S2-casein
(K)VPQLEIVPNsAEER(L)	2	830.8999	18.29	18.29	91.4	6.85	α-S1-casein
(K)VPQLEIVPNsAEER(L)	3	554.2693	15.2	15.2	75.6	6.85	α-S1-casein
(K)LVNELTEFAK(T)	2	582.3188	17.7	11.06	89.5	7.57	Serum albumin
(K)YKVPQLEIVPNsAEER(L)	3	651.3224	14.28	14.28	70.8	7.64	α-S1-casein
(K)YKVPQLEIVPNsAEER(L)	2	976.4783	16.31	10.48	89.2	7.68	α-S1-casein
(K)FALPQYLK(T)	2	490.2807	12.41	3.81	78.6	8.49	α-S2-casein
(K)LGEYGFQNALIVR(Y)	2	740.4046	13.89	13.89	61.4	9.16	Serum albumin
(R)YLGYLEQLLR(L)	2	634.3556	20.3	7.78	95.2	11.82	α-S1-casein
(R)FFVAPFPEVFGK(E)	2	692.8673	23.68	23.68	98.7	12.73	α-S1-casein
(R)DMPIQAFLLYQEPVLGPVR(G)	3	729.3947	20.93	15.65	88.9	14.11	β-casein
(R)DMPIQAFLLYQEPVLGPVRGPFPIIV(-)	3	970.5386	16.48	16.48	85.3	16.24	β-casein
(K)HIQKEDVPSER(Y)	3	446.5663	14.85	6.3	84.6	3.32	α-S1-casein
(K)AVPYPQR(D)	2	415.7293	16.02	10.27	96.1	3.74	β-casein
(K)FQsEEQQQTEDELQDK(I)	2	1031.42	23.65	23.65	82.8	4.26	β-casein
(K)FQsEEQQQTEDELQDK(I)	3	687.9476	16.41	13.01	80	4.29	β-casein
(K)TVDmEsTEVFTK(K)	2	741.8057	22.2	9.72	94.1	4.44	α-S2-casein
(K)VIPYVR(Y)	2	373.7316	13.84	5.75	94.9	4.96	α-S2-casein
(K)TVDMEsTEVFTKK(T)	2	797.8559	17.14	5.58	90	5.31	α-S2-casein
(R)NAVPITPTLNR(E)	2	598.3432	16.94	7	91.3	5.48	α-S2-casein
(R)NAVPITPTLNR(E)	3	399.2308	11.69	5.12	76.4	5.55	α-S2-casein
(K)DIGsEsTEDQAMEDIK(Q)	2	964.3483	21.2	21.2	70.7	5.58	α-S1-casein
(K)DIGsEStEDQAMEDIK(Q)	3	643.2352	15.4	6	82.3	5.6	α-S1-casein
(K)TVDMEsTEVFTK(K)	2	733.809	18.32	9.88	78.2	5.77	α-S2-casein
(K)VPQLEIVPNsAEER(L)	3	554.2697	14.58	14.58	73.6	6.79	α-S1-casein
(K)VPQLEIVPNsAEER(L)	2	830.9	19.52	19.52	97.4	6.8	α-S1-casein
(K)VPQLEIVPNsAEER(L)	2	830.9001	11.08	11.08	60.8	7.09	α-S1-casein
(K)LVNELTEFAK(T)	2	582.3152	17.59	9.74	84.5	7.53	Serum albumin
(K)YKVPQLEIVPNsAEER(L)	2	976.4808	15.99	15.99	69.4	7.63	α-S1-casein
(K)FALPQYLK(T)	2	490.2827	16.47	8.98	84.7	8.44	α-S2-casein

(R)YLGYLEQLLR(L)	2	634.3543	17.03	4.24	88.4	11.78	α-S1-casein
(R)FFVAPFPEVFGK(E)	2	692.8679	22.58	22.58	97.8	12.71	α-S1-casein
(R)DMPIQAFLLYQEPVLGPVR(G)	3	729.3939	19.95	15.84	90.1	14.09	β-casein
(R)DMPIQAFLLYQEPVLGPVR(G)	2	1093.589	18.01	18.01	85.4	14.17	β-casein
(R)DMPIQAFLLYQEPVLGPVRGPFPIIV(-)	3	970.5362	14.78	14.78	82.3	16.24	β-casein
(K)HIQKEDVPSER(Y)	3	446.5658	16.03	8.28	90.6	3.29	α-S1-casein
(K)SCQAQPTTMAR(H)	2	625.7841	13.12	13.12	72.2	3.5	κ-casein
(K)AVPYPQR(D)	2	415.7298	16.03	10.57	95.9	3.74	β-casein
(K)FQsEEQQQTEDELQDK(I)	2	1031.422	24.22	16.96	89.6	4.27	β-casein
(K)FQsEEQQQTEDELQDK(I)	3	687.9485	15.21	11.58	77.9	4.29	β-casein
(K)TVDmEStEVFTK(K)	2	741.8059	12.61	4.47	67.5	4.44	α-S2-casein
(K)AEFVEVTK(L)	2	461.7488	12.78	4.63	73.9	4.63	Serum albumin
(K)VIPYVR(Y)	2	373.7314	13.69	6.03	93.9	4.95	α-S2-casein
(K)TVDMEsTEVFTKK(T)	2	797.8553	15.49	5.66	71.3	5.29	α-S2-casein
(R)NAVPITPTLNR(E)	2	598.3426	11.7	2.32	86.1	5.47	α-S2-casein
(K)DIGsEsTEDQAMEDIK(Q)	2	964.3505	17.81	17.81	69.7	5.58	α-S1-casein
(K)DIGSEstEDQAMEDIK(Q)	3	643.2356	15.06	6.49	79.1	5.61	α-S1-casein
(K)DIGsESTEDQAMEDIK(Q)	2	924.3613	11.45	11.45	64.6	5.65	α-S1-casein
(K)TVDMEsTEVFTK(K)	2	733.8096	15.34	6.61	72.3	5.76	α-S2-casein
(K)VPQLEIVPNsAEER(L)	2	830.9007	20.2	20.2	100	6.79	α-S1-casein
(K)VPQLEIVPNsAEER(L)	3	554.27	15.25	15.25	75.2	7	α-S1-casein
(K)YKVPQLEIVPNsAEER(L)	2	976.4794	19.11	12.65	92.7	7.64	α-S1-casein
(K)FALPQYLK(T)	2	490.2845	16.48	7.71	81.7	8.43	α-S2-casein
(R)YLGYLEQLLR(L)	2	634.3555	20.52	7.06	94	11.78	α-S1-casein
(R)YLGYLEQLLR(L)	3	423.2383	13.13	13.13	71.8	11.9	α-S1-casein
(R)FFVAPFPEVFGK(E)	2	692.8688	23.45	23.45	100	12.71	α-S1-casein
(R)DMPIQAFLLYQEPVLGPVR(G)	3	729.3938	19.33	13.59	88	14.07	β-casein
(R)DMPIQAFLLYQEPVLGPVRGPFPIIV(-)	3	970.5394	14.91	14.91	79.1	16.26	β-casein
(K)KTVDmEsTEVFTK(K)	3	537.5738	11.57	10.04	65.1	2.64	α-S2-casein
(K)KTVDmEStEVFTK(K)	3	537.5725	11.31	8.77	63.6	2.64	α-S2-casein
(K)KTVDmEStEVFTK(K)	3	537.5768	10.6	9.84	62.6	2.65	α-S2-casein
(K)KTVDMEStEVFTK(K)	3	532.2421	15.21	7.72	60.1	3.13	α-S2-casein
(K)KTVDMEStEVFTK(K)	3	532.2431	12.36	7.69	67.9	3.12	α-S2-casein
(K)TVDmEsTEVFTKK(T)	3	537.5737	11.98	9.53	64.2	2.64	α-S2-casein
(K)TVDmEStEVFTKK(T)	3	537.5729	13.64	13.47	66.3	2.66	α-S2-casein
(K)NMAINPsKENLCSTFCK(E)	3	698.6386	9.31	9.31	60.8	5.9	α-S2-casein
(K)NMAINPsKENLCSTFCK(E)	3	698.6375	8.82	8.82	60.9	5.89	α-S2-casein
(K)NMAINPsKENLCSTFCK(E)	2	1047.446	6.34	0.16	65.3	5.96	α-S2-casein
(K)KTVDMEsTEVFTKK(T)	3	574.943	14.34	14.34	66.8	4.14	α-S2-casein
(K)IEKFQsEEQQQTEDELQDK(I)	3	811.3593	12.39	12.39	65.8	4.13	β-casein
(K)IEKFQsEEQQQTEDELQDK(I)	3	811.3595	8.35	8.35	63.5	4.15	β-casein

Appendix VIII

Table of identified phosphopeptides in Chapter 6

This table contains the complete list of all phosphopeptides identified by ETD-MS/MS from tryptic rat cardiomyocyte whole cell lyate in Chapter 6.

Sample	Protein	Peptide	Modifications	Spectrum counts
Con1	similar to proteasome (prosome, macropain) activator subunit 4	SVWGVSLVPRGQPRVETtAADTK	T18	1
	Isoform 1 of Tropomyosin alpha-1 chain	AISEELDHALNDMTsI	S15	10
		AISEELDHALNDMtSI	T14	11
	30 kDa protein	sPPNPENIAPGYSGPLK	S1	9
		IDGSNLEGGSQQAPStPPNtPDPR	T20	2
		IDGSNLEGGSQQAPStPPNTPDPR	T16	3
	77 kDa protein	ASSEGTQGSVsPK	S11	4
		ASSEGTQGsVSPK	S9	1
		DPSLDTNSSLATPsPSPEAR	S14	19
		DPSLDTNSSLATPSPsPEAR	S16	5
		DPSLDTNSSLAtPSPSPEAR	T12	15
	71 kDa protein	ASSEGTQGSVsPK	S11	4
		ASSEGTQGsVSPK	S9	1
		DPSLDTNSSLATPsPSPEAR	S14	20
		DPSLDTNSSLATPSPsPEAR	S16	5
		DPSLDTNSSLAtPSPSPEAR	T12	16
	Troponin I, cardiac muscle	ADEsSDAAGEPQPAPAPVR	n-term acetyl, S4	7
		ADESsDAAGEPQPAPAPVR	n-term acetyl, S5	1
		RRssANYR	S3 , S4	1
	similar to titin isoform N2-A	VKsPETVKsPK	S3 , S9	21
		AVsPTETKPTEK	S3	6
		TRPRsPsPVSSER	S5 , S7	3
		AVSPtETKPTEK	T5	2
		SRPQPAEEYEDDTERRsPTPER	S17	2
	35 kDa protein	ASSEGTQGSVsPK	S11	4
		ASSEGTQGsVSPK	S9	1

		DPSLDTNSSLATPsPSPEAR	S14	20
		DPSLDTNSSLATPSPsPEAR	S16	5
		DPSLDTNSSLAtPSPSPEAR	T12	16
	Heat shock protein HSP 90-beta	IEDVGsDEEDDSGKDK	S6	7
	myosin-binding protein C, cardiac-type	TSDsHEDAGTLDFSSLLK	S4	7
		RTSDsHEDAGTLDFSSLLK	S5	4
		AESAVAPTSMEAPEtPK	T15	5
		RTsDSHEDAGTLDFSSLLK	S3	1
	myozenin 2	sPPNPENIAPGYSGPLK	S1	9
		IDGSNLEGGSQQAPStPPNtPDPR	T20	2
		IDGSNLEGGSQQAPStPPNTPDPR	T16	3
Con2	30 kDa protein	sPPNPENIAPGYSGPLK	S1	18
	Myosin regulatory light chain 2, ventricular/cardiac muscle isoform	LEGGSsNVFSMFEQTQIQEFK	S6	15
		LEGGSSNVFsMFEQTQIQEFK	S10	1
		LEGGsSNVFSMFEQTQIQEFK	S5	1
	tropomyosin 1 alpha chain isoform i	AISEELDHALNDMtSI	T14	19
		AISEELDHALNDMTsI	S15	23
		AIsEELDHALNDMTSI	S3	1
	similar to titin isoform N2-A	VKsPELVASHPK	S3	42
		VKsPETVKsPK	S9	35
		AVsPTETKPTEK	S3	8
		AVSPtETKPTEK	T5	6
		RVKsPELVASHPK	S4	3
		RRTPsPDyDLYYYR	S5 , Y8	1
		SLGDIsDEELLLPIDDYLAMK	S6	1
		sPELVASHPK	S1	1
		TRPRsPsPVSSER	S5 , S7	1

		YSsPPAHVK	S3	1
	Heat shock protein HSP 90-beta	IEDVGsDEEDDSGKDK	S6	11
	myosin-binding protein C, cardiac-type	RTsDSHEDAGTLDFSSLLK	S3	17
		RTSDsHEDAGTLDFSSLLK	S5	8
		TSDsHEDAGTLDFSSLLK	S4	3
		AESAVAPTSMEAPEtPK	T15	2
		RtSDSHEDAGTLDFSSLLK	T2	1
	myozenin 2	sPPNPENIAPGYSGPLK	S1	18
Con3	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	YHGHsMSDPGVsYR	S12	15
		YHGHsmSDPGVsYR	S5	20
		YHGHsmSDPGVSyR	Y13	2
		YHGHSmsDPGVSyR	S7	1
		YGMGTsVER	S6	2
	Isoform 1 of Tropomyosin alpha-1 chain	AISEELDHALNDMTsI	S15	16
		AISEELDHALNDMtSI	T14	5
	Troponin I, cardiac muscle	RRSsANYR	S4	28
		RRssANYR	S3	13
	similar to titin isoform N2-A	VKsPETVKsPK	S9	12
		VKsPELVASHPK	S3	18
		RVKsPELVASHPK	S4	6
		AVsPTETKPTEK	S3	5
		AVSPtETKPTEK	T5	2
		SRPQPAEEYEDDTERRsPTPER	S17	7
		SRPQPAEEYEDDTERRSPtPER	T19	3
		TRPRsPsPVSSER	S5 , S7	2
	similar to Pyruvate dehydrogenase E1 component alpha subunit, somatic form, mitochondrial precursor (PDHE1-A	YHGHsMSDPGVsYR	S5	20
		YHGHsmSDPGVSyR	Y13	2

	type I) isoform 1	YHGHSMsDPGVsYR	S12	19
		YGMGTsVER	S6	4
		YHGHSmsDPGVSyR	S7	1
	myosin-binding protein C, cardiac-type	RTSDsHEDAGTLDFSSLLK	S5	9
		AESAVAPTSMEAPEtPK	T15	4
		AESAVAPTSmEAPEtPK	M10 oxi, T15	3
		RTsDSHEDAGTLDFSSLLK	S3	3
		TSDsHEDAGTLDFSSLLK	S4	1
	113 kDa protein	RVKsPELVASHPK	S4	7
		AVSPtETKPTEK	T5	2
		AVsPTETKPTEK	S3	5
		TRPRsPsPVSSER	S5 , S7	2
		SRPQPAEEYEDDTERRsPTPER	S17	7
		SRPQPAEEYEDDTERRSPtPER	T19	3
		VKsPETVKsPK	S3 , S9	12
		VKsPELVASHPK	S3	6
ERA+GEN1	Isoform 1 of Tropomyosin alpha-1 chain	AISEELDHALNDMtSI	T14	14
		AISEELDHALNDMTsI	S15	3
	30 kDa protein	sPPNPENIAPGYSGPLK	S1	3
	similar to titin isoform N2-A	AVSPtETKPTEK	T5	6
		VKsPELVASHPK	S3	2
	Heat shock protein HSP 90-beta	IEDVGsDEEDDSGKDK	S6	3
	myosin-binding protein C, cardiac-type	RTSDsHEDAGTLDFSSLLK	S5	4
		RTsDSHEDAGTLDFSSLLK	S3	1
		RtSDSHEDAGTLDFSSLLK	T2	1
	113 kDa protein	AVSPtETKPTEK	T5	6
		VKsPELVASHPK	S3	2

	141 kDa protein	RTSDsHEDAGTLDFSSLLK	S5	4
		RTsDSHEDAGTLDFSSLLK	S3	1
		RtSDSHEDAGTLDFSSLLK	T2	1
	myozenin 2	sPPNPENIAPGYSGPLK	S1	3
ERA+GEN2	myosin-6	QREEQAEPDGTEDADKsAyLmGL	S17 , Y19 , M21 oxi	8
		KLAEQELIETSERVQLLHsQ	S19	1
		NKDPLNETVVGLYQKssLK	S16 , S17	1
	Isoform 2 of Basigin	GSGsHLNDKDK	S4	7
	30 kDa protein	sPPNPENIAPGYSGPLK	S1	8
		IDGSNLEGGSQQAPStPPNTPDPR	T16	3
	Troponin I, cardiac muscle	RRSsANYR	S4	7
	similar to Chromobox protein homolog 1	KADsDSEDKGEESKPK	S4	11
	similar to titin isoform N2-A	VKsPELVASHPK	S3	13
		VKsPETVKsPK	S9	12
		AVsPTETKPTEK	S3	6
		AVSPtETKPTEK	T5	1
		SRPQPAEEYEDDTERRsPTPER	S17	4
		SRPQPAEEYEDDTERRSPtPER	T19	1
		TRPRsPsPVSSER	S5 , S7	4
	Heat shock protein HSP 90-beta	IEDVGsDEEDDSGKDK	S6	9
	similar to Chromobox protein homolog 1	KADsDSEDKGEESKPK	S4	11
	113 kDa protein	VKsPELVASHPK	S3	1
		VKsPETVKsPK	S3 , S9	11
		AVsPTETKPTEK	S3	6
		SRPQPAEEYEDDTERRsPTPER	S17	4
		TRPRsPsPVSSER	S5 , S7	4
		AVSPtETKPTEK	T5	1

		SRPQPAEEYEDDTERRSPtPER	T19	1
	myozenin 2	sPPNPENIAPGYSGPLK	S1	8
		IDGSNLEGGSQQAPStPPNTPDPR	T16	3
ERA+GEN3	30 kDa protein	sPPNPENIAPGYSGPLK	S1	14
		IDGSNLEGGSQQAPStPPNTPDPR	T16	4
		IDGSNLEGGSQQAPSTPPNtPDPR	T20	3
	77 kDa protein	DPSLDTNSSLAtPsPSPEAR	T12	8
		DPSLDTNSSLATPsPsPEAR	S16	2
		DPSLDTNSSLATPsPSPEAR	S14	11
	71 kDa protein	DPSLDTNSSLAtPsPSPEAR	T12	8
		DPSLDTNSSLATPsPsPEAR	S16	2
		DPSLDTNSSLATPsPSPEAR	S14	11
	Troponin I, cardiac muscle	RRSsANYR	S4	16
	similar to titin isoform N2-A	RVKsPELVASHPK	S4	21
		VKsPELVASHPK	S3	17
		VKSPETVKsPK	S9	11
		YSsPPAHVK	S3	2
		AVsPTETKPTEK	S3	1
		AVSPtETKPTEK	T5	1
		AVSPTEtKPTEK	T7	1
		SRPQPAEEYEDDTERRsPtPER	T19	1
		SRPQPAEEYEDDTERRsPTPER	S17	5
		TRPRsPsPVSSER	S5 , S7	1
	35 kDa protein	DPSLDTNSSLAtPsPSPEAR	T12	8
		DPSLDTNSSLATPsPsPEAR	S16	2
		DPSLDTNSSLATPsPSPEAR	S14	11
	Heat shock protein HSP 90-beta	IEDVGsDEEDDSGKDK	S6	15

myosin-binding protein C, cardiac-type	RTsDSHEDAGTLDFSSLLK	S3	14
	RTSDsHEDAGTLDFSSLLK	S5	14
	TSDsHEDAGTLDFSSLLK	S4	4
	TsDSHEDAGTLDFSSLLK	S2	2
	tSDSHEDAGTLDFSSLLK	T1	1