Proteomics

Phosphoproteomic Profiling of the Myocyte
Alistair V.G. Edwards, BS; Stuart J. Cordwell, PhD; Melanie Y. White, PhD

Protein phosphorylation underpins major cellular processes including energy metabolism, signal transduction, excitation-contraction coupling, apoptosis, and cell survival mechanisms and is thus critical to the myocyte. Targeted approaches, whereby a handful of phosphoproteins are investigated, can suffer from a relatively narrow view of cellular phosphorylation. In contrast, recent technical advances have allowed for the comprehensive documentation of phosphorylation events in complex biological environments, providing a deeper view of the “phosphoproteome.” A global, high-throughput characterization of the myocardial phosphoproteome, however, has not yet been achieved. Efficient analysis of phosphorylated proteins and their roles in a dynamic cellular environment requires high-resolution strategies that can identify, localize, and quantify many thousands of phosphorylation sites in a single experiment. Such an approach requires specific enrichment and purification techniques, developed to align with high-end instrumentation for analysis. Cutting-edge phosphoproteomics is no longer restricted to gel-based technology, instead focusing on affinity enrichment prior to liquid chromatography and mass spectrometry. We will describe the best current methods and how they can be applied, as well as the challenges associated with them. We also present current phosphoproteomic investigations in the myocyte and its subcompartments. Although the techniques and instrumentation required to achieve the goal of a myocardial phosphoprotein catalog in physiological and diseased states are highly specialized, the potential biological insight provided by such an approach makes phosphoproteomics an important new avenue of investigation for the cardiovascular researcher.

Key Words: myocardium • pathway analysis • phosphoproteins • mass spectrometry • phosphoproteomics
Phosphoproteomic Profiling of the Myocyte
Alistair V.G. Edwards, Stuart J. Cordwell and Melanie Y. White

*Circ Cardiovasc Genet.* 2011;4:575
doi: 10.1161/CIRCGENETICS.110.957787
*Circulation: Cardiovascular Genetics* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/4/5/575

Data Supplement (unedited) at:
http://circgenetics.ahajournals.org/content/suppl/2011/11/08/4.5.575.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation: Cardiovascular Genetics* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Circulation: Cardiovascular Genetics* is online at:
http://circgenetics.ahajournals.org//subscriptions/
Phosphoproteomic Profiling of the Myocyte

Alistair V.G. Edwards, BS; Stuart J. Cordwell, PhD; Melanie Y. White, PhD

Abstract—Protein phosphorylation underpins major cellular processes including energy metabolism, signal transduction, excitation-contraction coupling, apoptosis, and cell survival mechanisms and is thus critical to the myocyte. Targeted approaches, whereby a handful of phosphoproteins are investigated, can suffer from a relatively narrow view of cellular phosphorylation. In contrast, recent technical advances have allowed for the comprehensive documentation of phosphorylation events in complex biological environments, providing a deeper view of the “phosphoproteome.” A global, high-throughput characterization of the myocardial phosphoproteome, however, has not yet been achieved. Efficient analysis of phosphorylated proteins and their roles in a dynamic cellular environment requires high-resolution strategies that can identify, localize, and quantify many thousands of phosphorylation sites in a single experiment. Such an approach requires specific enrichment and purification techniques, developed to align with high-end instrumentation for analysis. Cutting-edge phosphoproteomics is no longer restricted to gel-based technology, instead focusing on affinity enrichment prior to liquid chromatography and mass spectrometry. We will describe the best current methods and how they can be applied, as well as the challenges associated with them. We also present current phosphoproteomic investigations in the myocyte and its subcompartments. Although the techniques and instrumentation required to achieve the goal of a myocardial phosphoprotein catalog in physiological and diseased states are highly specialized, the potential biological insight provided by such an approach makes phosphoproteomics an important new avenue of investigation for the cardiovascular researcher. (Circ Cardiovasc Genet. 2011;4:00-00.)

Key Words: myocardium • pathway analysis • phosphoproteins • mass spectrometry • phosphoproteomics

Myocytes are highly dynamic cells that need to respond efficiently to alterations in the extracellular environment to maintain contractile function and thus perfusion of the body. One of the primary mechanisms by which myocytes respond to stimuli and maintain a stable intracellular environment is protein phosphorylation, whereby phosphate groups are added or subtracted through the action of protein kinases or phosphatases, respectively (although other modifications, particularly glycosylation, may perform similar roles). This process modifies the function or localization of the target protein to achieve a particular result, be it a direct functional consequence (eg, troponin I) or as part of a larger signaling cascade (eg, mitogen-activated protein kinases). Myocardial phosphoproteins regulate a range of processes, both physiological and pathophysiological, and our understanding of them is therefore central to our understanding of how myocyte function is controlled in health and disease.

For instance, the biological impact of myosin phosphorylation in the myocardium has been linked to changes in force generation, with other components of the contractile apparatus also displaying phosphorylation changes that have distinct effects on cardiac biology (eg, cardiac troponin complex). Protein phosphorylation cascades interweave to produce enormously complex networks of protein interactions (Figure 1). These interlocking pathways can be analyzed in isolation, but a loss of relevance and contextual information should be expected because of the proportion of the overall network being ignored. Accordingly, phosphoproteomic methodologies have been developed that aim to cast a net over the entire complement of proteins involved in cell signaling at a given time to achieve a global understanding of phosphorylation-regulated signaling. This involves the application of various techniques, described below, with the goal of producing a catalog of phosphorylation events. Thus far, the phosphoproteome of the myocyte remains poorly investigated, although the tools to study it do exist; at the time of this writing, there have been no global myocyte phosphoproteomic studies of this variety published, although there have been several that examined particular subcellular structures (eg, mitochondria). This situation is likely due to the expertise and instrumentation required to adequately catalog the myocyte phosphoproteome and will, no doubt, not remain the case given the enormous biological impact of protein phosphorylation and the insight into cell signaling that can be offered by phosphorylation analysis.

From the Discipline of Pathology (A.V.G.E., S.J.C.), School of Medical Sciences and School of Molecular Bioscience (S.J.C., M.Y.W.), The University of Sydney, New South Wales, Australia.

Guest Editor for this article was Jennifer E. Van Eyk, PhD.

Correspondence to Dr Melanie White, PhD, School of Molecular Bioscience, The University of Sydney, Bldg GO8, Maze Crescent, NSW 2006, Australia. E-mail melanie.white@sydney.edu.au

© 2011 American Heart Association, Inc.

Circ Cardiovasc Genet is available at http://circgenetics.ahajournals.org

DOI: 10.1161/CIRCGENETICS.110.957787
Challenges of Myocardial Phosphoproteomics

Ideal qualities of a phosphoproteomic methodology include the following:

- Unbiased enrichment of $\text{P}\text{Ser}$, $\text{P}\text{Thr}$, and $\text{P}\text{Tyr}$ residues
- Unbiased resolution of the phosphorylated residue regardless of the characteristics of the surrounding residues (i.e., considerations of peptide charge state and site localization)
- High reproducibility (allowing statistical quantification)
- High sensitivity (allowing identification of the lowest-abundance signal proteins and transcription factors)

A typical large-scale phosphoproteomics workflow is depicted in Figure 2. Briefly, a biological sample is obtained from which protein is extracted and digested to peptides; this complex mixture is then subjected to phosphopeptide purification/enrichment, followed by chromatographic fractionation with subsequent identification and characterization by mass spectrometry (MS). These data are then searched against various in silico databases to assign site localizations and kinase/phosphatase motif patterns and to determine biological significance. Each portion of this workflow and its specific challenges will be discussed in detail. Although there are methods of interrogating or validating phosphoproteomes besides those shown, they are not suited to high-throughput studies. In this category are techniques such as $^{32}\text{P}$ labeling and Western blotting, which may suit targeted, protein-centric workflows (and have much to offer in this respect) but which are labor intensive and impractical on a larger scale. Targeted studies of key proteins that regulate phosphorylation, or are themselves phosphorylated, are subject to less of the uncertainty that can hamper large-scale mass spectral projects as a function of the probabilistic searching methods used. For this reason, it is common for high-throughput workflows to be validated by use of the above-mentioned techniques.

Producing Biological Samples for Phosphoproteomics

Given the transient and rapid nature of phosphorylation-based modifications, there are numerous considerations when it comes to choosing the origin of myocytes to be analyzed. One of the major factors influencing sample selection is the desired ability to partition those phosphorylation changes that pertain to the biological state of interest and those that are incidental to it. In this regard, use of human tissue, although obviously more clinically relevant, can be influenced by the manner in which it was acquired (biopsy or transplant), pharmaceutical status of the patient, and lifestyle influences. Alternatively, models of in vivo myocytes may circumvent this variability in clinical samples. The production of an effective model of a functional cardiomyocyte, however, is not a trivial exercise. Cardiomyocyte cell culture is one option, as is ex vivo tissue perfusion, yet both are limited by...
biological relevance compared with human tissue. Cardiomyocyte cell culture attempts to replicate a highly structured environment with well-defined routes of intracellular and intercellular communication; however, whether this is fully possible in this somewhat artificial environment is debatable. Simulations of biological events and their stimuli are also performed in a manner not necessarily representative of in vivo processes (eg, serum starvation before stimulation with high concentrations of growth factors). Ex vivo perfusion, although retaining some aspects of in vivo interactions, is limited by a lack of vagal parasympathetic innervations to regulate heart rate and by the constant degradation of the tissue associated with the preparation. Those conducting phosphoproteomic studies must be mindful of interference from acellular material (eg, extracellular structural proteins and growth medium constituents) when selecting a system, because the higher the degree of interfering material, the more difficult subsequent enrichments and, more importantly, biological interpretation will be. Similarly, highly abundant cellular proteins will interfere with analysis, especially if these are also phosphorylated. Preparative antibody pull-downs can be used to remove known problematic highly abundant proteins, but this can lead to unintentional loss of analyte (eg, Gundry et al7).

Another critical point in sample preparation for phosphoproteomics is the use of techniques to retain the in vivo phosphorylation pattern. This is typically achieved with inhibitors of protein kinases and phosphatases to limit phosphorylation and dephosphorylation as artifacts of sample handling. Common compounds include pervanadate to prevent tyrosine dephosphorylation8 and dasatinib to prevent phosphorylation.9 Studies of temporal signaling in the heart are few, but the frameworks have been laid down by high-throughput investigations of stimulated cells. Considerations here center around the time frame of sample collection with reference to the rapidity of phosphorylation, such that sampling must be performed at appropriate time points, typically not more than a few minutes apart and ideally at high temporal resolution in the early stages (<1 minute) of the stimulus in question.10–12

**Phosphopeptide Enrichment Techniques**

Most current workflows use the peptide products of trypsin proteolysis for enrichment (although protein-level enrichment...
is also effective) and fractionation, with subsequent high-sensitivity detection by MS. Digestion with trypsin typically produces charged MS-compatible peptide ions, but other enzymes can be used depending on the desired qualities of the produced peptides. For example, phosphorylation sites in peptides with multiple acidic residues will be relatively more difficult to detect because of low ionization efficiencies; this problem may be circumvented by use of alternate enzymes with different cleavage specificities, which changes the composition of the relevant peptide and improves MS detection. After digestion, the most common enrichment techniques include affinity capture by titanium dioxide (TiO$_2$)\textsuperscript{13} and immobilized metal-ion affinity chromatography (IMAC).\textsuperscript{14} Numerous metals with high binding efficiency to phosphate (including iron\textsuperscript{15} and gallium\textsuperscript{16}) are chelated to nitriilotriacetic acid– or iminodiacetic acid–coated beads to form the IMAC stationary phase. IMAC, however, is limited by an incompatibility with several common biological reagents (including EDTA and alkaline metal salts), its preference for multiply phosphorylated peptides, and its decreased specificity relative to TiO$_2$.\textsuperscript{17} By comparison, TiO$_2$ is tolerant to reagents unsuitable for use in IMAC-based enrichment strategies; however, it is thought that multiphosphorylated peptides interact too strongly with TiO$_2$ for complete retrieval. The complementary nature of TiO$_2$ and IMAC has led to their sequential application; often, IMAC-unbound flow-throughs are enriched by TiO$_2$.\textsuperscript{17,18} There are many other enrichment materials available (eg, zirconium dioxide\textsuperscript{19}), each with its own specificity and therefore slight bias. This is not always a negative; indeed, some investigators have found uses for these biases as separation techniques, as exemplified by the use of IMAC beads to separate multiple and singly phosphorylated peptides.\textsuperscript{17}

A common issue with chromatographic enrichment strategies is the specificity for phosphorylated peptides. Because phosphorylation introduces a net negative charge to the peptide, contamination from nonphosphorylated peptides is most often observed when the sequence contains long chains of charged residues.\textsuperscript{20} As a result, loading and elution conditions for these techniques are very specific. Acidic loading buffers show reduced nonspecific interactions from charged, nonphosphorylated peptides, and very alkaline elution conditions allow for ion exchange to occur, therefore ensuring maximal release from the charged chromatographic material.

Enrichment of phosphopeptides is not inherently biased toward serine, threonine, or tyrosine events, but the less frequent and low-abundance nature of tyrosine phosphorylation means that specific enrichment strategies may be needed to reliably and reproducibly detect it (this dearth of phosphotyrosine events is somewhat counteracted by their relatively easier MS fragmentation pattern; see section: Mass Spectrometry). Even though it accounts for only a small proportion of phosphorylated residues (<5%), tyrosine phosphorylation is biologically important, as exemplified by receptor tyrosine kinases, the first step in many signaling cascades.\textsuperscript{21} As such, phosphotyrosine residues are often investigated separately or as an adjunct to a global phosphoproteomic study by use of an antibody-based approach.\textsuperscript{22} Monoclonal antibodies directed against phosphotyrosine are effective and readily available. The same is not true of antiphosphoserine or antiphosphothreonine, because the relatively less characteristic structures of these residues decreases the specificity of the antibodies. Antibody-based methods are limited by the requirement for relatively large amounts of protein (typically on the order of 10 mg, compared with a global phosphoproteomic experiment run with <1 mg of protein) and are subject to similar levels of interference from nontarget peptides as is observed with affinity-based chromatographic methods. An interesting corollary of these antibody-based experiments is that as more of them are performed, with more sensitive MS techniques and yet uniformly large sample amounts, investigators are able to probe ever deeper into the low-abundance region of the phosphotyrosine proteome, such that a recent publication reported more than 1100 unique sites.\textsuperscript{23} This result raises the question of whether this level of tyrosine phosphorylation is still 1% to 2% of a much larger cellular phosphoproteome, as dogma suggests, or a much larger proportion of total phosphorylation (the PHOSIDA database\textsuperscript{24} currently lists \( \approx 24,000 \) human phosphorylation sites).

**Fractionation**

The second stage of a pre-MS enrichment strategy is an effective chromatographic fractionation technique. This is a conceptually simple operation whereby peptides are loaded onto a column of solid stationary phase (eg, long-carbon-chain material such as C$_{18}$) with which they interact and on which they are therefore retained. A mobile phase containing a solvent (eg, acetonitrile), which reduces the strength of the peptide-column interaction, is then run over the stationary phase using an increasing gradient of solvent. The result is a gradual elution of peptides, based on their physicochemical properties (eg, hydrophobicity), in a manner dependent on the degree of initial interaction. The degree of fractionation ultimately defines the number of eventual identifications by reducing the analyte complexity in the mass spectrometer and providing more analytic “space” in which individual analytes can be subjected to fragmentation. This is primarily a function of available MS analysis time. Therefore, the key element of successful phosphoproteomics is the use of coupled alternative chromatographic separations, based on different principles, with the aim of providing consistent but low-complexity mixtures for each MS scan.

Several separations can be performed effectively on phosphopeptides, although strong cation exchange (SCX) and hydrophilic interaction liquid chromatography (HILIC) are the most commonly used. SCX chromatography involves peptide binding to a negatively charged resin with elution by an increasing gradient of salt (mobile phase), such that peptides with relatively lower charge and therefore weaker interaction with the column elute earlier than those with strong interaction.\textsuperscript{22} Peptides carrying a negatively charged phosphate group are therefore enriched in earlier elutions. SCX is primarily a charge-state–based separation, whereas HILIC, by contrast, is based purely on hydrophobicity in a directly orthogonal fashion to standard reversed-phase C$_{18}$
liquid chromatography (RP-LC) separation. When HILIC is used, peptides are loaded in high concentrations of organic solvent (typically acetonitrile; mobile phase) with peptide hydrophilicity determining retention time. Phosphopeptides, which have increased hydrophilic characteristics, are retained on the column and eluted in decreased concentrations of organic solvent. The high-salt mobile phase used for SCX is not directly compatible with the RP-LC traditionally used before MS. As a result, additional cleaning steps are required, which can lead to sample loss. This postfractionation cleanup is not required for HILIC, because the mobile phase is directly compatible with RP-LC.

Mass Spectrometry
MS strategies for phosphoproteomics, when performed in a high-throughput manner, are based on fragmentation of peptide ions, which will provide sequence information that allows rapid identification and site elucidation. The rate-limiting step is the number of MS/MS (tandem mass spectrometry) acquisitions that can be performed for each MS (mass) scan, which in itself is defined by the number of analytes released from the RP-LC separation per unit of time (generally in milliseconds). With these considerations in mind, an MS capable of as many fragmentation events as possible per unit of time is preferable, because even long gradients and 2 levels of chromatographic separation may not provide the resolution needed for total proteome coverage. There are many mass spectrometers available, but a typical choice would combine an ion trap (to provide high-speed scanning and enable MS$^3$ fragmentation; see below) with a high-resolution mass analyzer (eg, Fourier transform-MS) to allow for reduced search space and increased identification rates. As a rule, liquid chromatography MS experiments are run with RP-LC separation immediately preceding the MS phase such that in addition to all foregoing separations, peptides undergo a final level of separation. This is the last stage at which peptides can be separated, and any peptides that are coeluted from reverse-phase resin will compete for analysis time.

When peptides undergo activation and fragmentation to produce an interpretable mass spectrum, loss of labile modifications such as phosphorylation (but including glycosylation and others) tends to provide a more energetically favorable fragmentation pathway than peptide backbone amide bond (N-C1) cleavage (Figure 3). Therefore, instead of a series of $b$- and $y$-ions, the result of conventional dissociation techniques (such as collision-induced dissociation) is often a diagnostic loss-of-phosphate peak observed at a mass 98 Da below the intact parent ion but with few sequence-informative peaks present (Figure 4). This in turn leads to an inability to sequence and identify the peptide. To avoid this problem, several options are available; for example, it is possible to use alternative fragmentation techniques that do not rely on collision events to fragment the peptide. These methods include electron-driven dissociation strategies such as electron transfer dissociation or electron capture dissociation, which lead to fragmentation at the N-C\textsubscript{H} backbone bond, creating $c$- and $z$-ions without preferential modification cleavage. The result is that the modification is retained on the residue in question, and sequence information is still obtained. Alternatively, collision-induced dissociation can be used to perform MS$^3$ (MS/MS/MS), whereby on detection of the dephosphorylated peptide in MS/MS (mass−98 mass units, as above), a third round of fragmentation can be performed to produce peptide backbone information from this nonphosphorylated ion. Depending on the number of phosphate groups on the peptide and the number of ions available, this MS$^3$ strategy can be repeated as many times as required until all modifications are removed and sequence data can be extracted. The caveat of such an approach is the longer time frame required to perform MS$^3$, which reduces the chance of any coeluting peptides receiving full analysis. Peptides that contain phosphotyrosine do not exhibit this neutral loss of phosphate and so do not require any specific MS strategies to detect and identify them.

Another common fragmentation approach is multistage activation (also known as pseudo-MS$^3$), which involves simultaneous selection and activation of the putative phosphopeptide ion to produce a composite MS$^2$/MS$^3$ spectrum.
that contains elements of both the 98-Da neutral loss spectrum and the fragmentation of that neutral loss peak. Compared with MS³, multistage activation does not require additional scanning events and therefore maintains optimal peptide-sampling conditions. Alternatively, a combination of electron transfer dissociation and collision-induced dissociation known as ETcaD can also be used to effectively add an element of collisional activation to an electron transfer...
discovery rate that is assessed with a decoy (eg, reverse or of false-discovery rates. Most investigators quote a false-
data sets. One development aimed at reducing this is the use into published data, especially posttranslational modification of a threshold introduces a high degree of variation searches must assign some threshold at which hits are deemed a result. This is a significant issue and one that has not yet thereof, is quantified in a manner that varies somewhat between search engines, and each peptide-spectrum match is displayed with a score that is proportional to the “fit.” This process appears simple and works well for unmodified peptides; however, it is not always perfectly suited to peptides that contain modifications such as phosphorylation, and the peptide-spectrum matches produced are often low-scoring as a result. This is a significant issue and one that has not yet been fully addressed; the outcome is essentially that all searches must assign some threshold at which hits are deemed untrustworthy and accept the false-positive and false-negative effect that this has. Given the variety of possible ways to search data and the stringencies to use when doing so, the choice of a threshold introduces a high degree of variation into published data, especially posttranslational modification data sets. One development aimed at reducing this is the use of false-discovery rates. Most investigators quote a false-discovery rate that is assessed with a decoy (eg, reverse or randomized) database to estimate the proportion of randomly matching peptides to address the issue of false-positives. Arguably an issue of more importance for phosphoproteomics, however, is site localization.

The precise site of phosphorylation is critical in regulating protein function. Therefore, it is important to assign a phosphorylation event to a specific residue within an identified phosphopeptide. This is not always easy, because mass spectra can be unclear and difficult to annotate unambiguously, even with sophisticated algorithms combined with manual inspection. This is especially true of modified peptide spectra. To address this point, various computational options have been produced. An example is Ascore (ambiguity score); another is the posttranslational modification score. Using “site determining ions,” each of these aims to assign a confidence score to a proposed phosphorylation site using computational “fit” models that predict the possibility of each potential site, such that the investigator can be quantitatively more certain that the phosphorylation is on one residue and not another. If the MS data are poor, they can use low-intensity ions to extract information that may have been missed by the standard search engines; if the data are good, on the other hand, their contribution may not need to be as substantial. Not all phosphorylation sites are in areas that contain multiple possible sites; naturally, these can be assigned unambiguously simply on the basis of peptide mass or diagnostic neutral loss peaks. However, for a trypsin-digested whole cell lysate (based on the Rat International Protein Index database), there are ~513 000 peptides that contain only 1 serine or threonine residue (phosphotyrosine can be ignored here because it typically remains intact), compared with ~629 000 with multiple possible sites of labile phosphorylation, which would therefore require site assignment. This is another factor to consider with regard to the unambiguous detection and resolution of the site of phosphorylation with reference to the residues in the immediate vicinity.

At the completion of the data-searching stage, the investigators now have a list of identified phosphorylated peptides that they are confident are present within the sample. The final challenge in phosphoproteomics, and perhaps the greatest, is that of turning large-scale phosphorylation data into meaningful systems biology information. There is an ever-expanding set of software tools to do exactly this, but they are all limited by one factor: The information in the database is at protein-level resolution, whereas the data that investigators aim to fit into these databases are generally at peptide or site resolution. Furthermore, where relative quantification has been conducted, information does not exist on whether an increase (or decrease) in a particular phosphopeptide/site is activatory or inhibitory. This means that for almost all large-scale studies, the effect of phosphorylation at one residue versus another is unknown, and all that can be done is to map phosphorylation sites onto a network of known protein interactions (which is not necessarily based on phosphorylated forms of those proteins). Databases of protein interactions such as KEGG (Kyoto Encyclopedia of Genes and Genomes) and Ingenuity Pathway Analysis can be used only to assess signaling flux; that is, if one pathway or network is displaying significantly more phosphorylation activity than another, then it is likely that this pathway is propagating biological information. The effect of this information transfer cannot truly be predicted unless the pathway and all phosphorylation sites have been characterized. Yet although many pathways are becoming better understood, most large-scale phosphoproteomics experiments currently produce far more never-before-seen sites than those for which we have some functional understanding; thus, we are limited to the well known and must extrapolate to, or hypothesize about, those that are unknown.

A substantial amount of effort in phosphoproteomics is devoted to quantitation; obviously, a binary “present or absent” phosphorylation assessment is simplistic. Because of the “cascading” nature of signaling pathways, a biological effect can be produced without phosphorylating every member of a protein pool, and in these cases, it is the relative degree of phosphorylation that is critical. Fortunately, this is comparatively easy to investigate. A mass tag (eg, dimethyl or iTRAQ) is added to all peptides from a given state or condition, which is detected in an abundance-dependent manner to allow for relative quantification in an MS setting (this can be extended to absolute quantification if a standard is included). Phosphorylation events that depart from those observed in control states (either in absence/presence or in degree) can then be mapped onto pathways in exactly the same way as above. In cell culture models, stable isotope labeling (SILAC) can be applied in place of chemical labeling. This isotopic labeling approach avoids complica-
tions related to chemical effects of added tags, such as systematic underestimation\(^4^2\) or charge state effects,\(^4^3\) and removes the need for an extra labeling step.

An important consideration in phosphopeptide quantitation is the interplay between protein abundance and the relative degree of phosphorylation, the determination of which requires 2 discrete measurements. The first of these is the site occupancy rate, or the proportion of phosphorylatable sites that are in fact phosphorylated for a particular peptide species. Estimation of site occupancy for high-throughput approaches in particular is a difficult task given that it requires both detection and accurate measurement of the phosphopeptide and nonphosphorylated counterpart and can only strictly be applied to singly phosphorylated peptides. The biological impact of differential site occupancy is not yet clear, but because it appears to be a tightly regulated phenomenon,\(^4^4\) and given the need for fine control of the complex overlapping pathways, it appears likely to have an important role. Second, overall protein abundance, for which alterations to phosphorylation need to be quantified with relation to the remaining nonphosphorylated protein (eg, increasing phosphorylation combined with decreasing abundance), needs to be determined. Although the impact of these situations will likely need to be established on a case-by-case basis, a 2-stage approach whereby the phosphoproteome and the entire proteome are analyzed concurrently may provide clues to the extent of pathway activation/deactivation in the context of proteome-level changes. To date, such an approach has yet to be undertaken, possibly because of the magnitude of such a task.

As a final step of analysis, consensus sequences can be used to perform motif analysis that can identify kinase/phosphatase-specific regulation with the use of open source tools including NetPhosK\(^4^5\) and MotifX\(^4^6\) (Figure 5). These provide a means of implicating those kinases or phosphatases that may have been active in the biological setting through identification of overrepresented sequences that surround phosphorylation sites, which can then be matched with kinase recognition sequences.

**Application to Biology**

Phosphoproteomics experiments remain in the minority in the cardiovascular area. More common are those that use phosphorylation-specific stains and 2-dimensional gel electrophoresis\(^4^7,4^8\) or \(^3^2^P\) labeling. It is our opinion that these techniques have limited application to high-throughput studies. They are, however, ideally suited to the examination of smaller or individual segments of pathways in greater detail; this is likely to become an important function as the amount of phosphoproteomic data increases and there is a greater demand for information relating to the function of particular sites.

Nevertheless, for discovery proteomics, high-throughput is necessary. A recent example of the application of these methods is the work of Deng et al,\(^4^9\) who used electron transfer dissociation and collision-induced dissociation fragmentation to map phosphorylation sites in cardiac mitochondria. They effectively applied all of the above-mentioned stages of a phosphoproteomic workflow, substituting a gel-based separation for a chromatographic separation, before using TiO\(_2\) enrichment and localizing 236 phosphorylation sites. With the use of bioinformatics tools, the investigators were able to align the observed phosphoproteins with the electron transport chain and the Krebs cycle, among other pathways. It is informative to note that approximately 90% of these sites were not recorded previously in databases, an observation that underscores the lack of site-specific information available to investigators.

In a similar vein, Boja et al\(^5^0\) also used liquid chromatography MS/MS to determine phosphorylation changes in response to several stimuli, achieving quantification of phosphorylation sites from porcine heart mitochondria and identifying 56 in total. This study also used TiO\(_2\) as the enrichment medium, although the investigators chose not to apply any further fractionation. By comparison with the above-mentioned study by Deng et al,\(^4^9\) a substantial drop in identifications was observed. The 2 studies are otherwise very similar, and it seems likely that the lack of both an orthogonal separation technique and an alternative fragmentation method will have contributed greatly to the lower recovery of phosphopeptide identifications. However, myocardial phosphoproteomics is understudied relative to other tissue types (eg, liver [5000 sites]\(^2^2\)) and especially compared with cell culture studies (eg, Olsen et al\(^4^4\)).

**Future Applications**

Novel fragmentation technologies that are in their infancy could have a substantial impact on phosphoproteomics; a way to obtain sequence information within a single scan with an intact phosphate moiety while avoiding the drawbacks of current technologies (eg, the required anion reaction time for electron transfer dissociation) would significantly improve the ability to penetrate the phosphoproteome. Similarly, more specific and simpler enrichment techniques will reduce the amount of MS time spent sequencing coeluting nonphosphopeptides. These enrichment improvements are constantly taking place, with each successive publication bringing the methods further forward, and they are also being combined with novel liquid chromatography techniques to further enhance fractionation.

The area in which phosphoproteomics can most be improved, however, is bioinformatics. More effective ways to turn phosphopeptide observations into biology are needed, although for the time being it does appear that protein-level databases are the best available option. Databases that contain a large proportion of all phosphoproteomic data (ie, phos-
phorylation sites) do exist (eg, PhosphoSite™), and these allow the veracity of a particular observation to be assessed, but they need to be linked to biological information more effectively. These additions to the phosphoproteomics arsenal will enable more effective and cohesive contributions to be made to our understanding of myocyte phosphorylation signaling, as well as more global integration of these contributions into a meaningful biological systems analysis.

Sources of Funding
This work was supported by the National Health and Medical Research Council of Australia (NHMRC 571002; Dr Cordwell). Dr White is a National Health and Medical Research Council of Australia CJ Martin postdoctoral fellow (464905). A.V.G. Edwards is the recipient of an Australian postgraduate award.

Disclosures
None.

References


