Molecular and Cellular Life Sciences master's program

Writing Assignment

Biomolecular Mass Spectrometry and Proteomics group

Examiner: Dr Maarten Altelaar

Topic:

Post-translational modifications crosstalk: current knowledge and future perspectives

Student:

David Vargas-Diaz

Number 3885283

2014

Summary for general purpose

Proteins are cellular components product of gene expression with multiple and important functions in the cell environment and in the organism in general. Their role as building blocks and workers of the cell is widely recognized. Proteins are target of post-translational modifications (PTM), which are chemical modifications in specific aminoacids made by other proteins. Such PTMs may appear one at the time but also in combination with the same or other types of modifications, creating multiple versions of the same protein (PTM-isoforms). It is assumed that the cell benefits from this expand diversity of PTM-isoforms by assigning a different function or location to each one. The way in which two or more PTMs interact with each other to regulate protein activity is referred as PTM crosstalk. The present work reviewed recent literature in the topic of PTM crosstalk based on the hypothesis of an actual consistent pattern of protein modification that translates in precise regulation of the cellular processes.

From the current knowledge in the field of PTM crosstalk, it is known that post-translational modification can interact in many different ways depending of the type of PTM involved, the location of the protein and the desired effect on protein regulation, among others. As example, the addition of a phosphate in a specific aminoacid of many cytoplasmatic proteins may be necessary to attract the enzymes in charge of adding an ubiquitin (other type of PTM) in an adjacent aminoacid, forming together a specific signal for protein degradation by the respective machinery in the cell. The detection of these PTMs is done by various methods, being the measurement of the changes in the aminoacid mass after the addition or partial loss of one of such modifications the most commonly used. This is done by means of a technology called mass spectrometry, which have been adapted to measure with high accuracy the mass of proteins and its aminoacids. The increasing amount of information resulting from the study of PTMs requires the development of informatic tools for its analysis. The combination of experimental data and informatic analysis has been crucial for improving the understanding and validation of PTM crosstalk.

Regarding its practical importance, alterations related to the PTM crosstalk are found to be involved in many diseases, including cancer. The alterations may include, among others, mutations in aminoacids that cannot longer be modified by the same type of PTM or the inhibition of enzymes required for protein modification. In consequence, efforts are being made to use the knowledge on PTM crosstalk to guide therapies and recover the normal protein regulation. All the above supports the necessity to improve and innovate the methodology and technology involved in the understanding of PTM crosstalk and its practical applications.

PTMs and the concept of crosstalk

Post-translation modifications (PTM) are widely present in cellular proteins from nucleus to membrane and in evolutionary distant organism from bacteria to humans. They have been recognized as important regulators of protein function helping to achieve the high levels of complexity in signaling transduction and integration that more evolve organism required, by increasing protein diversity with the so called PTM isoforms and keeping the size of the genome proportionally invariant. The knowledge on PTMs is update daily and the data produce accordingly is growing almost exponentially, what has taken literature review and integrations efforts from a useful tool to a necessary activity. The present is one of such attempts to organize the current state of PTMs research, specifically the interaction between two or more modifications known as crosstalk, and the future endeavors are required to keep the field moving in the right direction. The literature chosen and the analysis of it were done under the hypothesis that PTMs can and do interact with among them through different mechanism to regulate the cellular activity by following a consistent and organize pattern.

As Hunter reviewed in his paper: “The age of crosstalk: phosphorylation, ubiquitination and beyond”, the experimental evidence supporting the effect of PTMs on well-known proteins involved in disease and signal transduction is strong and come from multiple sources and techniques [1]. He evidenced, by providing clear examples, the existence of basically two ways of interaction between PTMs:

(I) the positive crosstalk will happen when one modification facilitates addition or removal of the second modification, or both modification are mutually necessary regulate protein interactions or functions; (II) in the other hand, negative crosstalk occurs when PTMs compete for the same residue or one modifies the neighboring environment inhibiting the activity around the second modification by blocking its addition or removal. It is easy to recognize that the plethora of different interaction and mechanism for it to happen are not as simple as this classification may suggest however is a good start to organize and guide PTM research.

Along with recognition of protein regulation by some frequent pairs of PTMs as phosphorylation and ubiquitination, research has found some of the cellular systems use to read such combinations of modified residues and generate an effect. These mechanisms of crosstalk readout include the existence of proteins containing two modular binding domains for PTMs detection, ranging from proteins with two domains for the same PTM (e.g. twin SH2 domains) to adjacent domains for different PTMs in the same protein (e.g. PHD and bromodomain in the BPTF subunit of the NURF complex) [1]. Although such mechanism of simultaneous recognition of multiple PTMs exist, crosstalk may occur as sequential addition and removal of modifications each one requiring the previous one, also as events taking place in different cellular compartments, or by PTM regulation of any of the PTM enzymes.

The current knowledge around cellular methods for PTM recognition has been continuously reviewed and update by the scientific community. Deribe, Pawson and Dikic made a compressed while comprehensive analysis of 5 PTMs (i.e. phosphorylation, ubiquitination, acetylation, methylation and hydroxylation), the enzymes known to catalyze the reactions and some of the binding domains involved in the specific recognition of those PTMs [2]. More importantly for the final aim of the PTM crosstalk research, their review contextualizes PTMs in important examples of signal transduction as the cellular response to activation of the receptor for tumoral necrosis factor alpha, the activation and endocytosis of the receptor for epidermal growth factor, and the response to DNA damage. These three mechanisms of cellular response to different environmental and internal cues

are subject of extensive research since their discoveries years ago, however the contributions of the PTM research around their multiple components have considerable expand and improve their understanding by the scientific community. Although the examples of interactions and effects of PTMs in those cellular processes are multiple, the complexity and details of their actual contributions to the cell activities are changing constantly as the research on PTMs move rapidly forward with more PTMs, better technology and innovative informatics solutions to integrate data better (*see below*).

How to study PTMs

Research regarding detection of PTMs is published daily but the understanding of the biological function of such data is still poor. In the process of understanding the PTM landscape of proteins the energy have been particularly focus on finding consistent patterns that translate into function. The bases to look for such patterns come from accepted knowledge, for example the existence of regions of extensive modification in many proteins, the presence of cellular proteins with affinity for specific modification within short consistent sequences, and some clear examples of combinatorial PTMs that modified the function of well study proteins. An important step to solve this puzzle is to review and improve the methods to indentify PTMs, their mechanism of crosstalk and their biological relevance in the cellular context.

The experimental approaches most frequently used nowadays for detection of PTMs are based on antibodies, peptide libraries and mass spectrometry. Antibodies have been used for years in proteomics as well-understood tools to enrich for proteins and peptides containing PTMs. Lothrop *et al.* reviewed recent evidence on the effect upon antibodies affinity of PTMs after changes in the neighboring residues [3]. The effect may be generated by sequence variation involving unmodified residues down- or up-stream of the target, but also for the modification of original residues. This means that for high reliability in PTM-enrichments based on antibodies it is required to have at least one antibody for every possible combination of residues or modified residues around the PTM of

interest, what is an impractical goal to pursuit and clearly shows the limitations of the data acquired with this method.

In the case of peptides libraries, their importance rely in the present capacity to synthesize as many combinations of residues and modified residues as required to study their effect in protein-protein interactions as well as in recognition and modification by PTM enzymes [3]. However, this approach requires to take the protein of interest out of their biological environment to assess the interaction with the modified peptides. This generates the necessity of a suitable environment for protein function while in contact with the peptides and the imminent possibility of results irrelevant in the context of a cell because of the *in vivo* inaccessibility of the protein to the same peptide or because both components are located in different cellular compartments, among other causes of variation.

In reference to mass spectrometry, it has provided most of the current knowledge on PTMs and its benefits are widely known by interested in the field. Because of it, this analysis will focus on the limitations of the method as reviewed in [3, 4]. The first one is the effect of the dissociation method employed by the mass spectrometer to fragment the peptides, what allows the identification of the residues included in the sequence. Most PTMs had shown to be sensible to this process and their loss from the modified residue is not easily detected. To surpass this technical challenge methods are in constant development and improvement and they are showing promising results in cases as phosphorylation with the arrived of the electron transfer dissociation. The final consequence of the undetectable loss of modification has been the production of an enormous amount of PTM data up-to-date with incalculable but expected incompletes, unrelated to the sensitivity of the machine.

An additional barrier is presented for the way in which the protein conformation and landscape is reconstructed by the bottom-up approach as a whole. In this widely used approach the information regarding the masses of the residues and modified residues are combined to the measurements of the unfragmented peptides to reconstructed the short sequences, which are later align to the data bases including protein sequence information coming from DNA readout. The

requirement from so many steps of fractionation and fragmentation and their subsequent reconstitution involves a sum of assumptions, predictions and opportunities for losing information, which makes the system dependent on statistic validation and constant experimental cross-validation [3, 4]. A solution to this problem is crucial for PTM crosstalk studies, as interaction is expected to happen between close positioned modification requiring high sequence coverage and reliable location of every modification. The best approach to study the PTM landscape decreasing the effect of protein destruction and reconstruction is studying complete undigested proteins, in what is call a top-down approach. Even a middle-down approach is preferred to protect crosstalk information, what is achieved by choosing non-tryptic enzymes to cut the protein in less frequent positions generating longer peptides. The implementation of those less destructive methods advance as fast as the technology developments allowed. For now, bottom-up approaches keep being necessary and practical although the accurate localization of detected PTMs is still a bottle neck, subject of the disadvantages above mentioned.

The specific mass variations used to identify each PTM are also subject of limitations that may be confounding and required attention to improve crosstalk understanding. Choudhary and Mann make emphasis in the detection of ubiquitination and disadvantages of using the GG motif (two consecutives glycines) as signature of its presence previous to trypsinization [4]. In this particular case the motif detected is also share by at least other two PTMs, the neural precursor cell-expressed developmental downregulated protein 8 (NEDD8) and the interferon-induced 17kDa protein (ISG15) [4]. Advances towards the independent identification of these three modifications it is important to classify accurately the interactions in which they are involved.

Besides the many obstacles that PTM crosstalk research still need to overcome in the technical aspect when studying single proteins, it has been seen that crosstalk may happen between residues in different although interacting proteins, presenting a remarkable opportunity for innovation considering that the

visualization of such multi-protein patterns are far from the current technical possibilities.

Current knowledge on PTM crosstalk

Research on post-translational modifications has been focus in its vast majority to the detection of one single PTM at the time, mainly because enrichment methods are necessary although most designed for one modification only. This single modification approach has delayed the detection and understanding of PTM crosstalk. The current knowledge on the topic comes from intensive review and integration of individual data from different PTMs on the same proteins, validated with *in vivo* and *in vitro* assays (e.g. knock-down, RNA interference, mutation of modified residues, inactivation of PTM enzymes), and supported in many cases by informatic tools. Recent literature containing in part or in total explicit analysis and results on PTM crosstalk is reviewed below.

Methylation and phosphorylation

Methylation occurs in lysine residues which may compete with other modifications as sumoylation, ubiquitination and biotinylation. It crosstalk with these and other modification have been recently reviewed by Erce and colleagues and their findings are discuss next [5]. Evidence suggests that methylation may negatively crosstalk with ubiquitination by means of the lysine residues to increase half-life of the cellular proteins, an effect seen with lysine acetylation. This type of crosstalk has been seen as well between methyl-lysines and other modifications commonly targeting different residues, as is the case of phosphorylation (i.e. serine, threonine and tyrosine, in eukaryotes). For example, methylation can be blocked by phosphorylation of the residue next to the lysine to be modified, affecting all functions that required the methylation as may be protein-protein interactions. Vice versa methylation can suppress phosphorylation while selecting for binding partners and protein localization. However, the interplay between these two PTMs

also occurs positively, like in the case of STAT6 which requires previous methylation before the phosphorylation can happen [5].

Along the examples supporting both types of crosstalk mentioned above, many studies have described a tight interaction between the two PTMs going from methyltransferase- and kinase-substrate networks to dynamic regulation of proteins localization and function. Moreover, crosstalk between methylation and other PTMs, including phosphorylation, can be directly associated to disease, as is the case involving mutations in the DNA damage-induced apoptosis network [5]. Such interaction in the context of an important cellular procees as apoptosis is supported by many independent studies connecting the key members of the network with the different PTM isoforms and the PTM enzymes involved in the process.

Acetylation, phosphorylation, methylation and ubiquitination

Lu *et al.* made a bioinformatic analysis of data on lysine acetylation and the crosstalk with other PTMs, their main finding are discussed next [6]. Acetylation has been study in a variety of organisms and the conservation of modifications among them can be used as a proxy to predict biological relevance. For example, data from human acetylation sites were found to be highly conserved in the mouse proteome (i.e. 92.9%) suggesting their functions are important for the survival of the cell [6]. From those proteins containing such conserved modification sites 79.0% are present outside the nucleus, which contrast with the idea that acetylation is an important regulator limited to histones [6]. When interrogated about crosstalk, data showed 1165 independent acetylated sites located within 10 residues of a phosposite, what is a highly significant result [6]. The interaction between this two PTMs is reinforce for the fact that acetylation neutralize the charge of modified lysines, what have been predicted to affect the phosphorylation of nearby residues (approx. 5 residues in distance). The cause of such effect is likely to be an alteration in kinase binding affinity. The same prediction approach was extended to methylation and ubiquitination sites, finding an effect in methylation sites 5 residues away of the acetylation and in ubiquitination sites

mostly 10 residues away but as far as 40. Should be noticed that the effect of acetylation over other PTMs and vice versa also occurs as competition or sequential events in the same residue, as lysine is a common target of many modifications. Information related to such interactions is still uncommon due to the requirement of a different experimental set-up (e.g. time-points).

These authors and others considered important to keep in mind the bias towards high abundant proteins inherent to mass spectrometry-based data, which in the present case translates into a possible overestimation of the conservation and positive pressure around the acetylation sites. Improvements in bioinformatics and statistics of proteomics data is making possible to include the effect of such bias in the algorithm of analysis, allowing better inferences while technology advancements solve the problem (see below for an example).

Phosphorylation, ubiquitination, SUMOylation and acetylation

An important and innovative effort combining experimental data and bioinformatic analysis was done by Beltrao and colleagues to analyzed phosphorylation, ubiquitination, SUMOylation and acetylation in organisms as different as humans, plants and fungus [7]. They made a phosphorylation-centric analysis considering that this is the type of PTM with more available information, and group the other three as lysine modifications. When the phosphosites in the protein sequence were compared with the other PTMs in study it was found a significant higher concentration of phosphorylated residues in lysine modified proteins compared to their concentration in the complete pool of proteins [7]. The possible bias towards abundant proteins, known to happen in mass spectrometry experiments, was taking in consideration in the statistical analysis of this co-existence showing a highly significant connection.

Further, the distance between PTMs was used as proxy to evaluate their possible interaction. It was established that phosphorylated versions of serine, threonine or tyrosine residues are significantly more abundant than their unmodified counterparts when closer than 15 positions to a modified lysine [7]. Additionally, to predict functional importance of the distance between PTMs the conservation of

the modification through the different organisms was considered, resulting in a higher conservation of phosphorylated aminoacids when they are closer than 15 residues to a modified lysine compared to average and random phosphosites [7]. Research combining experimental data, public data and bioinformatics are more plausible to generate realistic predictions and inferences than any other based on a single approach. Another recent example of such multi-source approach showing state-of-the-art characteristics is described in the next section.

The case of p53

When reviewing PTM crosstalk is impossible to skip the constant references to the better understood cases: tubulin, histones and p53. The histones modifications are not considered in the present work due to the apparent independency showed in the analysis of such results by the scientific community compared to the rest of the proteome. From the two remaining cases, p53 was chosen arbitrarily as a single protein example of the above mentioned interactions. Its multiple known interactions involve, but are not limited to, acetylation, ubiquitination and methylation. As Yang and Seto reviewed acetylation and ubiquitination in p53 are mutually exclusive but both are required for protein activation although involving different genes as targets [8]. Ubiquitination as usual may select the protein for cytosolic degradation and in contrary acetylation stabilized nuclear localization. Methylation is also found to crosstalk with acetylation in the context of p53. An example is the recruitment of acetylases by their chromodomain, promoting acetylation which can further form motif of recognition for other acetylases [8].

From the previous it is easily inferred that lysine residues are an important target of PTM enzymes and they are fairly involved in crosstalk. Therefore, it is expected that removal of exchange of such residues will interfere with regulation of p53 activities. The outcome of a set of *in vivo* experiments challenging the importance of lysines in p53 was the total loss of regulation, both positive and negative, by lysine-modifications (e.g. acetylation, methylation and ubiquitination) [8].

Developments for the analysis of PTM crosstalk

Recently, the method development specifically aimed to detect and analyze PTM crosstalk is starting to produce results. Three of these efforts are discussed below, two of them related to mass spectrometry-based detection and one for bioinformatic integration of available data.

Mass spectrometry-based detection of PTM crosstalk

Phosphorylation and ubiquitination are widely found in almost all cellular activities and their interaction may be as simple as the activation of ubiquitin ligases by phosphorylation. However the best example of actual crosstalk between these two PTMs is the so called phosphodegrons. These phosphodegrons described the process of ubiquitination as consequence of phosphorylation at specific residues in the same protein, which is a mechanism assumed to regulate proteins activated by phosphorylation in an irreversible manner. Until recently most of the research related to PTMs was made studying one single modification at the time and peptides rather than full proteins, what by definition make impossible to detect crosstalk. Crosstalk hypothesis, as the discussed previously, are the result of bioinformatic analysis of separate datasets or from regulating one modification and measuring the effect in other PTMs, which in both cases is full of assumptions and estimations.

In response to this single approach, researchers as Swaney *et al.* are developing methods to sequentially isolate peptides containing at least two modifications of interest [9]. One of such approaches Swaney and colleagues requires the culture of cells with His-tagged ubiquitin followed by isolation of the ubiquitin-modified proteins as first step [9]. Next, proteins are digested and phosphorylated peptides are selected, as well as those containing the glycine-glycine (GG) signature seen after tryptic digestion of ubiquitinated residues. Such approach allowed the identification of peptides containing single modifications as well as a combination of both. However, many of the mono-modified peptides may actually be a longer sequence in the same protein separated by enzymatic digestion.

Swaney *et al.* also developed a method based on direct tryptic digestion of all cellular proteins and their chromatographic separation based on charge states (i.e. +1 for phosphorylation; +2 for simultaneous phospho-, GG-peptides; and +3 for GG-peptides) [9]. This second approach allowed the identification of ubiquitinated peptides that may not be easily enriched by His-tag affinity while maintaining the use of endogenous ubiquitin by the biological model in study. Also, it requires less manipulation of the sample decreasing the inevitable loss of material by any additional experimental step. The difference in methodology was enough to produce an 18-fold increase in unique ubiquitinated phosphopeptides selection [9]. Should be notice that charge-based approaches require close co-localization of both PTMs. Moreover, both methods include a bias towards the tryptic peptides containing both PTMs in their short sequence.

Further analysis of the mentioned results showed that phosphorylated residues differ between ubiquitinated and non-ubiquinated peptides, suggesting that close location of both modifications has a specific function and it is not merely coincidental. Comparative studies of the modified peptides against their yeast counterparts demonstrated preferential conservation of modified sites in peptides with both PTM rather than sites of mono-modified peptides, which supports the hypothesis of a functional interaction between co-ocurring phosphorylation and ubiquitination. Additional analysis also provides directionality to the crosstalk, suggesting ubiquitination is preceding phosphorylation in most cases.

Published at the same time as the previous, Mertins and colleagues at The Broad Institute in Massachusetts developed a method of serial enrichment involving three PTMs (i.e. phosphorylation, ubiquitination and acetylation), also based in a mass spectrometry analysis [10]. The method use tryptic peptides and starts with a metal-affinity selection of phophopeptides, followed by antibody-based selection of GG peptides from the first supernatant, and finally antibody-based selection of acetylated peptides from the second supernatant. From 7.5 mg of protein, 20800 phospho-, 15408 GG-, and 3190 acetylated sites were found when search individually [10]. When the three products of purification were analyzed and searched for the presence of two or three PTMs at the time, only ca. 0.3% of all

peptides contain two different modifications and none showed three simultaneously. Such percentage of di-modified peptides represent a much lower number than the obtained by Swaney and colleagues for co-ubiquitinated and phosphorylated peptides (*see above*).

Important differences to consider between both methods are the samples analyzed, which in the case of Mertins were human cell lines compared to yeast cells in Swaney´s case [10, 11]. The dynamic range that need to be overcome in the case of human cell lines is greater, what may required a higher efficiency of affinity purification to gather the same amount of modified peptides. The different outcome may also suggest that co-modification of peptides in human cells occurs less frequently inside tryptic peptides than in the yeast proteome. It should be notice that conservation of PTM sites doesn´t exclude that the residues around may change inducing new points of enzymatic digestion. Other plausible explanations to keep in mind are the complexity of the PTM landscape in each protein and it increase in human samples compared to yeast, what alters the interactions of such peptides in processes of selection. Even if the regulation of the detected sites remains constant through evolution, the state of the PTM landscape in the cell at the moment of collection will differ if both samples are not synchronized. In summary, the above mentioned points make clear that experiments implementing different biological models and conditions should not be compared; and crosstalk studies should include different enzymes for protein digestion to improve coverage and reduce bias towards one kind of peptide.

Bioinformatic integration of available data to infer crosstalk

The vast number of proteins interacting at the same time in a single cell makes technically challenging to proof *in vitro* or *in vivo* biological relevance of the increasing data on PTMs. As a solution, comprehensive efforts have been made on bioinformatic analysis and modeling of public available data, with a frequent use of evolutionary conservation of single and co-occurring PTMs as proxy to predict biological function. Minguez *et al.* compiled data from 8 different eukaryotes comprising more than 100 000 modified sites from any of 13 abundant PTMs (e.g.

phosphorylation, ubiquitination, glycosylation), and analyzed them with a novel algorithm contemplating the evolutionary distance as a parameter for scoring conservation results of every residue [11].

Their analysis brings to light many relevant results with statistical significance. One with particular importance for the design of future experiments is the preference of most PTMs for a specific location in the cell, functionality or protein region [11]. It should be inferred from the previous that searching for every PTM in a complete cell approach is inefficient and misleading. Novel experiments must analyze the PTM landscape within cellular compartments to first understand their role in a simpler context, and subsequently correlate the results with the widely available data from complete cells. They also observed as few as 20% of proteins containing two or more PTMs types [11]. Which suggest a considerable challenge for the sensitivity and enrichment capacity of the current techniques, considering the low possibility to detect two modifications in the same sequence within the already small fraction of residues modified at one time. The previous provides an additional layer of difficulty for the study of protein crosstalk along with the many already mentioned above. Such challenge requires constant innovation and creativity as first step to solve the puzzle.

Further, Minguez and colleagues discovered that each PTM type associates with at least other 6 types and particular pairs of these modifications co-evolve together [11]. Some of those pairs are found to be related to different cellular processes or locations. For example, when O-linked glycosylation and SUMOylation co-evolve the proteins involved are specifically connected to cell differentiation and tissue development [11]. The examples given, including the latttest, support the existence of patterns of co-occurring modifications functionally relevant for the cell so they are preserved through evolution. It was also found an important association between phosphorylation, acetylation and ubiquitination in nucleus or cytoplasm, particularly when associated to protein degradation, interaction already discussed that validates the work by Minguez *et al.* [11].

The existence of a stable and consistent pattern of co-modification in proteins is additionally supported by other conclusions of this multi-species computational

analysis, including high conservation of specific pairs of PTMs in certain protein domains. Protein domains are known for being involved in particular functions and locations. Moreover, short motifs including conserved pairs of PTMs are consistently present as indicative of crosstalk. The suggested physical interaction between close PTMs may be important for PTM enzymes recognition, formation of protein complexes and other interactions.

Databases for PTM crosstalk

As it becomes evident from the previous sections, modern science is providing more data that what can be analyzed or understood. This makes imperative to advance understanding before or in parallel to the design of new experiments, in order to avoid misuse of resources and to guide the development of technology. Because of it, databases are emerging as a need and a valuable tool to handle such amount of data at the time that help to proof and create hypothesis.

Recently, databases have evolved from merely repositories of topic-specific information to complex multimedia efforts to integrate and present data in new and better ways. A great example of such resourceful databases is PTMcode, created by Minguez and colleagues to improve understanding of PTM crosstalk by gathering and organizing data based on 5 criteria: (I) co-evolution of modified residues among different eukaryotes; (II) modifications competing for the same residue; (III) protein regions accumulating a high density of PTMs, so called hotspots; (IV) actual proximity of modifications in the folded protein; and (V) manual annotation of interactions based on published literature [12]. This approach is innovative in many ways, for example it presents data from 7 PTM-databases and makes use of at least other 5 to obtain, classify or organize data. Furthermore, it imports pre-existing 3D protein models while provides visualization by known applications as Jmol. In summary, it takes advantage of well-developed efforts by the scientific community and integrates them in a PTM-centric perspective with a creative interface.

Tools as PTMcode are what is required to accelerate understanding of PTM crosstalk and what modern science is all about. It is successful in providing new

insights in the field while using the data already available, instead of adding new pieces to a puzzle that has become too complex and big to solve with independent efforts. The integration of structural information with PTM localization in PTMcode generates not only a visual summary of the current knowledge in PTM landscape of a large set of proteins but also allows the prediction of physical interactions, as competition for protein niche, when distance between modification are closer than 4.69 Å supporting the existence of crosstalk.

PTM crosstalk as a code

Although PTM crosstalk is a strongly supported by examples as varied and abundant as the mentioned above, the existence of a code that allows a systematic readout of the existent modifications and their related functions is still a discussion in the field. The multiple considerations for avoiding the use of such a term are summaries by Sims and Reinberg [13], and cited frequently in articles referring to the topic. Three mains arguments are given by these authors. First, examples are given of how the same PTM may produce different and even opposite effects when present, as positive and negative control of the gene expression by methylation. Second, it is argued that the PTM function depends in the cellular context in which occurs, lacking consistency. And third, it is seen that even when the same PTM is present in the same context the effect may or may not be present, resulting in unpredictability. These arguments against the PTM code are based on evidence obtained from single PTM analysis, and as have been already discussed above PTM effect seems to work in a concerted manner as crosstalk more than single independent events. The introduction of the concept of PTM crosstalk may explained why the same PTM does not have the same effect every time is present, because the state of modification of the neighboring residues is probably different, which change the protein-protein interactions and the affinity for proteins partners. Also, for a correct fine-tune of the cellular processes the effect of PTMs and their crosstalk do not need to be consistent in all cellular contexts, they only need to be consistent every time the same cellular context is presented.

In the light of the recent discoveries it appears that the difference with the genetic code, predictable and consistent by codifying information in nucleotide triplets, is the distance from which is dissect, not the existence of one and the inexistence of the other. The genetic code only requires a continuous sequence of nucleotides to be understood, while the PTM code may required the simultaneous detection of the cellular context and the modification landscape of all the proteins involved at one time point. It can also be understood as a gain in consistency when PTM crosstalk is observed from far, considering protein partners and enzymes available, and adjusting the code to the cellular compartment of interest.

The future of PTM crosstalk

To achieve a better understanding of PTM crosstalk and its consequences for the regulation of cellular processes it is necessary to change from the classic single modification to an integral approach. This integral approach needs to consider as many components of the network of regulation as possible and being able to assign the results to a specific cellular context. Steps towards such goal should consist of technology development and also changes in the algorithm of analysis.

In the technical aspect, network research is a concept growing in the mass spectrometry field, where novel methodologies of detection are developed constantly. Two of the most important advances that may facilitate PTM crosstalk research are the possibility of labeled proteins metabolically or chemically accompanied by highly accurate detection in the mass spectrometer, and the development of targeted approaches. Labeling may help to differentiate the changes in the PTM landscape after certain stimulus, or allow the analysis of the modification state by cellular compartment [14, 15]. Targeted mass spectrometry is useful to increase the depth of the analysis, increasing sensitivity by specific selection of the protein to analyze [14]. Such approach also improves consistency of results and allows reliable quantitation [14]. Analysis of complete networks of regulation based on PTM crosstalk, in different cellular context and clinical relevant states, is a realistic aim in the foreseeable future techniques as technological advances are combined with the right experimental designs.

The ultimate goal of research in PTM crosstalk and it network of regulation should go beyond the natural curiosity implicit in every scientific quest. It is the possibility to improve understanding and treatment of clinical relevant conditions what should move the field forward and motivates the assignment of resources. Recently, the importance of PTMs and their crosstalk have been evidence in cancer network studies. Creixell and colleagues use current literature to support the connection between oncogenic mutations and alteration in the fine-tune regulation of cellular networks [16]. Evidence of such concept comes from mutated PTM enzymes as drivers of malignant transformation and the high probability that deregulation of any given protein may be related to modification of its PTM landscape, considering the many examples and references discussed above [16]. Network medicine is emerging as a concept to provide sequential treatment against the affected component of the cellular network and the possible resistance to be developed because of redundancy in the crosstalk and changes in the effect of PTMs according to the cellular context.

It is clear that the field of PTM crosstalk is still in its infancy. Many hypothesis need to be proved and many more need to be developed. Innovation is required, in technology, methodology and experimental design. Concepts as rich, complex and novel as the PTM crosstalk required the will to analyze and review new and accepted knowledge with new perspectives. To conclude, it is consider that the present analysis of current literature in the topic of post-translational modifications crosstalk provided supportive evidence and discussion around its original hypothesis: PTMs can and do interact among them through different mechanism to regulate the cellular activity by following a consistent and organize pattern; although ultimate evidence remains elusive.

Reference

1. Hunter T. The age of crosstalk: phosphorylation, ubiquitination and beyond. Molecular Cell 2007; 28: 730-738.
2. Deribe YL, Pawson T and Dikic I. Post-translational modifications in signal integration. Nature Structural & Molecular Biology 2010; 17: 666-672.

3. Lothrop AP, Torres MP and Fuchs SM. Deciphering post-translational modification codes. FEBS Letters 2013; 587, 1247-1257.
4. Choudhary C and Mann M. Decoding signaling networks by mass spectrometry-based proteomics. Nature Reviews – Molceular Cell Biology 2010; 11: 427-439.
5. Erce MA, Pang CNI, Hart-Smith G and Wilkins MR. The methylproteome and the intracellular methylation network. Proteomics 2012; 12: 564-586.
6. Lu Z, Cheng Z, Zhao Y and Volchenboum S. Bioinformatic analysis and post-translational modification crosstalk prediction of lysine acetylation. Plos One 2011; 6: e28228.
7. Beltrao P, Albanese V, Kenner LR, Swaney DL, Burlingame A, Villén J, Lim WA, Fraser JS, Frydman J and Krogan NJ. Systematic functional prioritization of protein posttranslational modifications. Cell 2012; 150: 413-425.
8. Yang X and Seto E. Lysine acetylation: codified crosstalk with other posttranslational modifications. Molecular Cell 2008; 31: 449-461.
9. Swaney DL, Beltrao P, Starita L, Guo A, Rush J, Fields S, Krogan NJ and Villén J. Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. Nature Methods 2013; 10, 676-682.
10. Mertins P, Qiao JW, Patel J, Udeshi ND, Clauser KR, Mani DR, Burgess MW, Gillette MA, Jaffe JD and Carr SA. Integrated proteomic analysis of post-translational modifications by serial enrichment. Nature Methods 2013; 10, 634-637.
11. Minguez P, Parca L, Diella F, Mende DR, Kumar R, Helmer-Citterich M, Gavin A, van Noort V and Bork P. Deciphering a global network of functionally associated post-translational modification. Molecular Systems Biology 2012; 8: 599.
12. Minguez P, Letunic I, Parca L and Bork P. PTMcode: a database of known and predicted functional associations between post-translational modifications in proteins. Nucleic Acids Research 2013; 41, D306-D311.
13. Sims RJ and Reinberg D. Is there a code embedded in proteins that is based on post-translational modifications? Nature Reviews, Molecular Cell Biology 2008; 9: 1-5.
14. Bensimon A, Heck AJR and Aebersold R. Mass spectrometry-based proteomics and network biology. Annual Reviews in Biochemistry 2012; 81: 379-405.
15. Rhee H, Zou P, Udeshi ND, Martel JD, Mootha VK, Carr SA and Ting AY. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. Science 2013; 339: 1328-1331.
16. Creixell P, Schoof EM, Erler JT and Linding R. Navigating cancer networks attractors for tumor-specific therapy. Nature Biotechnology 2012; 30: 842-847.