Exploring protein phosphorylation by combining computational approaches and biochemical methods

Gonzalo Pérez-Mejías, Alejandro Velázquez-Cruz, Alejandra Guerra-Castellano, Blanca Baños-Jaime, Antonio Díaz-Quintana, Katiuska González-Arzola, Miguel Ángel De la Rosa, Irene Díaz-Moreno

Instituto de Investigaciones Químicas (IIQ), Centro de Investigaciones Científicas Isla de la Cartuja (cicCartuja), Universidad de Sevilla, Consejo Superior de Investigaciones Científicas (CSIC), Avda., Américo Vespucio 49, Sevilla 41092, Spain

1. Introduction

The complexity of the human proteomes goes far beyond the ca. 20,000 ORFs (open reading frames) contained in the genome. A myriad of processes give rise to the structural and functional variety of proteins without increasing the genetic background. Alternative promoters and splicing, as well as mRNA editing yield a five-fold increase in the transcriptome. Post-translational modifications (PTMs) represent another tool by which cells can expand the functional diversity of proteins. These modifications may occur at any step in the “life cycle” of a protein and alter either its function or stability. We can sort PTMs into five groups: 1) addition of complex molecules, 2) addition of proteins and peptides, 3) amino acid modifications, 4) addition of chemical groups and 5) cleavage of part of the protein sequence [1]. This work will focus on phosphorylation, which is the main PTM involved in cell regulatory processes [2–6].

Metabolic regulation by protein phosphorylation is notable for its rapid reversibility, an essential feature in cell signaling [7]. This PTM consists of the addition of a phosphoryl group either to a hydroxyl-containing residue through an ester bond, or to a nitrogen-containing one to form an amidate, this process is generally mediated by kinases. The energy of this phosphate ester linkage is similar to that of the ATP β–γ anhydride bond. In fact, under certain experimental conditions, it is possible for kinases to act in the opposite direction, thereby generating ATP. Thus, high ATP:ADP ratios forestall dephosphorylation by protein kinases [8,9].

The residues most prone to phosphorylation are Ser (ca. 60%), followed by Thr (ca. 25%) and Tyr (ca. 15%) [6,10]. Phosphorylation of other residues—such as histidine and arginine—is less stable and infrequent in eukaryotes [11,12], and is thus excluded from this work.

Phosphorylation sites can be found within intrinsically disordered regions (IDRs) or structured domains. Analyses suggest that serine and threonine phosphorylation takes place predominantly in IDRs, while tyrosine phosphorylation occurs both in IDRs and structured regions [13]. Phosphorylation induces both order to disorder and order to disorder transitions. An example of order to disor-
The substitution of the hydroxyl group of Ser, Thr and/or Tyr by a phosphate group affects the physicochemical and immunogenic properties of any protein, and is consequently used for its isolation and analysis. Purification methods for phosphorylated proteins usually take advantage of 1) changes in the isoelectric point (pI), 2) metal cation chelating capacity of phosphate groups, and 3) unique immunogenicity of isoforms [18].

The main purification techniques used for isolation of phosphorylated proteins are chromatofocusing (CF), ion exchange chromatography (IEX), immobilized metal ion affinity chromatography (IMAC), metal oxide affinity chromatography (MOAC) and immunoaffinity chromatography (IAC). CF and IEX are based on the downshift in protein pI caused by phosphorylation [19]. The magnitude of this shift depends strongly on amino acid composition, secondary structure and number of concurrently-phosphorylated residues [18,19]. There are several pI prediction algorithms which calculate the influence of different PTMs, including phosphorylation, but such theoretical approaches are not exempt from errors and their accuracy for phosphorylated proteins depends on neighboring phosphosides and protein folding [20,21].

Many phosphorylated protein purification strategies utilize the tendency of phosphate groups to chelate metal ions and oxides. IMAC consists of phosphoproteins binding to metal-functionalized resins, usually containing Fe3+, and its elution by either a pH gradient or competitive displacement by chelating ligands, typically imidazole [22,23]. To improve the selectivity and efficacy of IMAC, Kinoshita et al. [24] designed a phosphate chelating group, the so-called Phos-tag, which is an alkoxide-bridged binuclear metal (Zn2+ or Mn2+) complex. Phos-tag and its conjugated derivatives bind to phosphorylated residues in an amino acid-independent manner [25]. This technology has been extensively used for phosphoryl- 

tion detection by SDS-PAGE (sodium dodecyl sulfate–polyacrylamide gel electrophoresis) and western blot analyses [25–27], and has not yet been surpassed by other types of IMAC developed over the last few years [22,23,28]. As an alternative to this technique, MOAC employs metal oxides, most commonly TiO2, to selectively isolate phosphoproteins [29]. As with IMAC, phosphoproteins are complexed with metal oxides under acidic conditions; however, complete recovery of phosphorylated proteins in MOAC requires an alkaline pH gradient from 8.5 to 11.5 for elution of multiple phosphorylated isoforms [30].

When purifying phosphoproteins, IAC is more selective than IMAC or MOAC, though it still suffers from low specificity. IAC exploits the unique immunogenic properties of phosphorylated proteins, utilizing antibodies that detect specific phosphosides in a sequence-independent manner. Antibodies raised against kinase consensus sequences can also be used [31]. Currently, only anti-phosphorylated Tyr (pTyr) antibodies are widely used in phosphoproteomics, despite the lower abundance of phosphorylated Tyr residues, compared with Ser and Thr [32]. The relatively low specificity of anti-phosphorylated Ser (pSer) and anti-phosphorylated Thr (pThr) antibodies favors the use of more cost-efficient approaches such as IMAC/MOAC [31]. In fact, the best asset of IAC, i.e. the high affinity of antigen–antibody interactions, is offset by the high cost of antibody production [18]. Recent papers report combined purification strategies using IEX, TiO2-MOAC and/or IAC, either sequentially or as complementary techniques, to optimize protein enrichment into phosphoryroses [33–35].

Unlike anti-pTyr antibodies, the Src homology 2 (SH2) domain has recently emerged as an affinity reagent in the field of phosphoproteomics. SH2 modular domains are present in many signaling proteins such as growth factor receptor-bound and signal transducers and activators of transcription (STAT) protein family members. Single SH2 domain expression and covalent coupling to chromatographic resins allows the purification of phosphoryrosines [36,37]. For selectivity improvement, Kaneko et al. [38] designed high-affinity SH2 domains by directed evolution methods. They discovered that the substitution of three key residues in the SH2 binding site of the tyrosine kinase Fyn notably increases its affinity for Tyr-phosphorylated proteins.

Obtaining phosphoproteins from natural sources has limitations in terms of sample quantity and possible downstream applications. Many researchers dealing with endogenous phosphoproteins conduct quantitative and temporal phosphoproteomic profiling of the components of cellular signaling networks under certain stress conditions, including disease [39–41]. The methodology used in these studies includes chromatographic techniques for phosphoprotein isolation and enrichment so far described, as well as typical mass spectrometry analyses [37,42]. Nonetheless, natural sources of endogenous phosphoproteins are not optimal since large quantities are needed for structural and/or biophysical assays.

The simplest and most common strategy to produce high amounts of any protein is the use of overexpression vectors for bacterial (usually E. coli) or yeast systems. They allow the expression of the construct of interest, normally a gene encoding a fusion protein with a tag sequence—e.g. polyhistidine, glutathione S-transferase (GST), or maltose-binding protein—to facilitate purification by affinity chromatography (Ni2+–IMAC, GST-sepharose or amilose resin). Assuming that the expression and purification of the recombinant protein provides good yields, both in terms of quality and quantity of the sample, it is still necessary to phosphorylate in vitro the target residues using the appropriate kinase enzyme. However, there are very few examples in the literature of preparative methods for in vitro phosphorylation of proteins used in binding assays, i.e. to study the effect of protein phosphorylation on an interaction with a third protein [43,44]. In most papers, the in vitro kinase reactions are intended to characterize the phosphorylation event itself – to determine parameters such as kinetics, kinase substrate specificity and target-site identification [45–48].

Baudin and Müller [45–48] took a different approach and in bacteria co-expressed both a kinase domain and a STAT protein for in vivo phosphorylation. Using this approach, they purified Tyr-phosphorylated STAT proteins, which were used to assess the binding of STAT proteins to specific DNA sequences by electrophoretic mobility shift assay (EMSA). This successful strategy could potentially be applied to other kinase-phosphoprotein pairs.
Although, if any given kinase phosphorylates more than one residue, it will be necessary to resort to point (e.g. Tyr to Phe) and/or deletion mutants of the target protein to study the effect of each phosphoresidue separately.

The lack of information on specific kinases for the vast majority of the identified phosphorylation targets demands the adoption of other methodological approaches. Thus, the use of phosphomimetic mutants, either classic variants (e.g. Ser/Thr/Tyr to Asp or Glu) or those incorporating non-canonical amino acids (ncAAs), is a good alternative for studying the structure and interactions of phosphorylated proteins.

### 1.2. Classical phosphomimetic variants using standard amino acids

As discussed above, the use of endogenous phosphorylated protein is frequently not feasible because of its low yield when purified from cell extracts or the absence of information on specific kinases. In this case, the easiest approach is the use of phosphomimetic proteins. Since phosphorylation is a post-translational event, cells lack native tRNAs to carry phospho-residues and to transfer them to nascent polypeptides. Classically, Ser- and Thr-phosphorylation have been mimicked by mutations to aspartate and glutamate. These substitutions provide the negative charge of each phosphoresidue and are of a similar volume. Therefore, they constitute the simplest methodology to introduce canonical amino acids [49,50]. Glu is a more suitable phosphomimetic replacement as it matches the phosphate oxygens of pSer and pThr, while Asp is less suitable as its chain is shorter. However, a major drawback of these mutations is they fail to feature the hydration layer and formal charges of phosphate [7].

Despite their utility, residue substitutions are not always possible. In some cases, it is necessary to introduce a pair of acidic residues to generate the double charge and mimic the behavior of the natural phosphoprotein [51–53]. However, this strategy may lead to local stress on protein structure due to side chain repulsion.

Glutamate is often used to mimic pTyr [54,55]. However, the glutamate residue is unable to mimic either the charge or the volume of pTyr. Indeed, there are examples in the literature wherein Tyr phosphorylation could not be emulated by this mutation [56,57]. Alternatively, several ncAAs can be used to mimic the abovementioned properties of the phosphate group (Table 1).

### 1.3. Incorporation of unnatural amino acid to obtain phosphomimetic proteins

The use of ncAAs in recombinant protein expression enables modification of the chemical properties of one or more residues [58]. Moreover, the use of technologies to selectively and cotranslationally incorporate phosphomimetic amino acids into proteins can increase sample preparation yields to enable biophysical analysis. There are two current options for introducing ncAAs into polypeptides: genetic code expansion and reprogramming. Both strategies rely on codon reassignment. The expansion approach targets stop or quadruplet codons, whereas reprogramming involves reassigning sense codons to ncAAs [59]. The most popular approach utilizes the evolved tRNA and aminoacyl-tRNA synthetase (aa-RS) pair system, which consists of a suppressor tRNA, normally targeting the low-usage amber (TAG) triplet, and an orthogonal aa-RS recognizing the tRNA bearing a specific ncAA.

Thus, recombinant expression of aa-RS/tRNA pairs and the protein construct of interest, together with the addition of the ncAA to the culture medium, constitute the basis for phosphomimetic protein biosynthesis [60–62]. The aa-RS/tRNA pair system has successfully been used with a wide variety of cells, both prokaryotic and eukaryotic, including yeast, plants and animals [58]. However, the effectiveness of ncAA incorporation into proteins is variable.

| Phosphorylated residues | Unnatural amino acid | Citation |
|-------------------------|----------------------|---------|
| Phosphothreonine        | 4'-Carboxymethoxy-3'-phosphonophenylalanine | [85] |
|                          | N-{[4-(2-{2-[3-(2-Bromo-acetylaminomethyl)-propionylamino]-3-hydroxypropionylamino})ethyl]-phenyl}-oxalic acid | [86] |
|                          | 3-{4-(2-[2-(2-Bromo-acetylaminomethyl)-ethyl]disulfanyl]-ethylcarbamoyl}-cyclohexylcarbamoyl]-pyrrole-2-carboxylic acid | [86] |
|                          | Photocaged phosphothreonine | [73] |
|                          | Sulfitothreonine | [87] |
|                          | p-Carboxymethyl-phenylalanine | [88] |
|                          | p-(Phosphonoamino)-phenylalanine | [89] |
|                          | O-phosphothreonine | [90] |
|                          | Benzyl 2-benzoxycarbonylamino-3-[4-bis(dimethylamino)phosphoryl oxyphenyl]propanoate | [91] |
|                          | 4-Phosphomethyl-L-phenylalanine | [77] |
| Phosphotyrosine          | Photocaged phosphotyrosine | [73] |
|                          | O-phospho-c-threonine | [92] |
| Phosphoserine            | Photocaged phosphoserine | [73] |
|                          | Phosphononemethylene alanine | [161] |

It depends on multiple factors, such as expression levels and performance of a given orthogonal aa-RS/tRNA pair, sequence context of repurposed codons within a gene, number and/or diversity of ncAAs present in the same protein, and cellular bioavailability of ncAAs [63]. Nehring et al. [64] examined the incorporation efficiency of several Tyr analogs by cognate aa-RS/tRNA pairs, at a specific position of a target protein in yeast. Using western blot analysis followed by mass spectrometry characterization, the authors reported significant differences in the performance of the aa-RSs tested [64]. Pott et al. [65] studied the influence of nucleotide context in amber codon suppression efficiency. They generated a library of sequences around the TAG triplet and performed expression assays, identifying hit sequence contexts with relative protein yields of 70–110% (compared to the levels of control proteins without ncAAs), three to five fold higher than the average yield obtained for random sequences [65]. Schwark et al. [63] assessed the role of the release factor-1 (RF-1) in the efficiency of amber codon reassignment. The results of an expression screening with ncAA-containing green fluorescent protein (GFP) mutants, both in wild type and RF-1 deleted E. coli cells, suggested that the contribution of RF-1 in translation efficiency is highly variable in single amber codon constructs, but has a consistent and significant negative effect on multisite mutants [63].

These few examples illustrate the convenience, even necessity, of optimizing the orthogonal translation systems. This can be achieved with high-yield protein expression vectors such as pEVL, or selectivity-enhanced aa-RS, tRNA and/or bacterial elongation factor Tu mutants for stronger interactions [58,66–69]. It is also possible to use a knock-out of the tRNA competitor RF-1 for improved binding to amber/quadruplet codons, or engineered ribosomes for higher decoding efficiency [70–72]. Moreover, the ncAA bioavailability can also be optimized through protection of repurposed codons within a gene, number and/or diversity of ncAAs present in the same protein, and cellular bioavailability of ncAAs [63].
which are imported by the ATP-dependent DppA transporter and eventually hydrolyzed, releasing the ncAAs inside the cell [77]. Alternatively, the use of cell-free protein synthesis (CFPS) for ncAA incorporation has gained popularity in recent years. This approach seems to be particularly suitable for solving the intrinsic cellular bioavailability constraints of phosphoresidues, among other issues [78]. Protein expression with this method occurs in an open reaction environment to which crude cell extracts containing an over-expressed tRNA/aa–RS pair are added along with the ncAA, which is perpetually accessible to the translation machinery. There are several reports of increased production yield due to the implementation of CFPS protocols, as corroborated by western blot and mass spectrometry analyses [78,79]. Oza et al. [80] developed a CFPS-based method to enhance the synthesis efficiency of site-specific Ser–phosphorylated constructs. Notably, the application of this system allowed them to increase the relative yield (i.e. percentage of wild type protein expression) of doubly phosphorylated human MEK1 kinase by ca. 80% compared to previously reported data for the same construct using an in vivo approach [81]. Wakamori et al. [82] took advantage of the CFPS platform to boost the efficiency of the simultaneous incorporation of acetyl-lysine into four specific positions of histone H4, achieving virtually 100% expression of the tetra-acetylated species. Yanagisawa et al. [83] also resorted to CFPS to improve the ratio of Lys-methylated Histone H3 synthesized to ncAA consumed by 10–20-fold compared to in vivo expression, which is particularly relevant when working with expensive and/or difficult to synthesize ncAAs [78].

Despite the evident strengths of CFPS, a great variability in reaction performance has also been documented, the causes of which can be multiple (such as ncAA sequence location in the protein construct) and unmanageable [84]. Some authors have developed strategies to obtain pSer, pThr, and pTyr by adding a protective “hood” against the action of phosphatases [73]. This protective “hood” is a photocleavable o-nitrophenylethyl group, which uncaps the phosphorylated residue upon irradiation with 350 nm light.

There is a great variety of ncAAs developed to mimic both the charge and the volume of phosphoresidues (Table 1). pTyr residue is the amino acid for which the greatest variety of ncAAs are available, in an attempt to mimic the volume and properties of its aromatic ring [73,77,85–91]. However, most authors choose the ncAA p-carboxy-methyl-L-phenylalanine (pCMF) because of the relatively low cost of its synthesis—it is even commercially available—and the efficient development of the technique employed to introduce it in a desired position (see below). Recently, further strategies based on the pEVOL technique have been designed to allow the synthesis and incorporation of pTyr and pThr residues in vivo, thus preventing the action of phosphatases [90,92] (Table 1).

1.4. Case study

Many researchers tend to picture how proteins or biomolecules function by using static atomic-level structures while neglecting the dynamics of the molecule. Nonetheless, Molecular Dynamics (MD) simulations have been used increasingly over the past few years to study macromolecular systems due to the constant progress in computing power. Recently, Hollingsworth and Dror (2018) published an extensive review on MD simulations highlighting their potential application in molecular biology and drug discovery [93]. For example, MD simulations can be used to figure out how a biological system will respond to certain perturbations, such as amino acid mutation [94] or post-translational modification [95].

We herein provide a methodological article to offer a guide for readers who want to study phosphorylated proteins by combining computational and biochemical approaches. To illustrate this, two human proteins—for which phosphorylation in several residues has been studied—have been chosen as models: cytochrome c (Cc) and human antigen R (HuR) (Fig. 1). These two proteins have been selected as their number of phosphorylation sites is known and their relatively small size reduces the computational time required for simulations.

Cc is an essential metalloprotein for mitochondrial metabolism and homeostasis, playing a dual-purpose role in cell life and death. Under homeostasis, Cc acts as an electron carrier within the respiratory chain in the mitochondria. However, stress stimuli induce the release of Cc to the cytosol and nucleus where it interacts with several targets [96–102]. In the cytosol, Cc triggers programmed cell death (PCD) by activating the caspase cascade [103–105]. In the nucleus, Cc is involved in DNA damage response through its interaction with histone chaperones [99,100]. All of these Cc functions are modulated by post-translational modifications [106–108]. In mammals, phosphorylation of residues 28, 47, 48, 58 and 97 has been reported [4,109–113] (Fig. 1A). Tyr48 is the folding nucleus of Cc, and its phosphorylation alters the dynamics of the hemeprotein, thus regulating oxidative phosphorylation and PCD-related functions [114].

HuR is an RNA-binding protein that modulates the processing of mRNAs involved in cell survival, proliferation and apoptosis. HuR is composed of three RNA recognition motifs (RRMs) connected by flexible regions. The modular structure of HuR means the function of the RRM domains can be studied independently and/or in tandem [115–118]. For this reason, computational and biochemical analyses of single RRMs can provide information of molecular changes caused by PTMs. Its function is highly regulated by the phosphorylation catalyzed by several kinases [119]. For example, Janus kinase 3 (JAK3) can phosphorylate Tyr63, Tyr68 and Tyr200 [120]. Tyr63 and Tyr68 are located at the first motif (RRM1), which comprises residues 1–101 (HuR1–101) (Fig. 1B). On the other hand, Tyr200 is located at the linker region between RRM2 and RRM3. Functional data suggests that Tyr200 phosphorylation impairs HuR binding to target transcripts by either shifting its localization or altering the conformation of RRM2 [120]. However, the effect of phosphorylation at RRM1 remains unveiled. Understanding conformational or dynamic changes induced by this PTM may provide a rationale for changes in functionality.

2. Methodology

2.1. In silico approach to study protein phosphorylation

2.1.1. Prediction of phosphorylation sites

Tyr-phosphorylation is the most frequent PTM found in cell signaling, due to the highly conserved pTyr binding domains and the low basal phosphorylation levels of kinase protein targets, essential for a signaling mechanism [3,121]. However, there may be several residues in a protein that can be phosphorylated. Hence, as a first approach, it is useful to check the information on phosphorylation sites provided by database servers, such as PhosphoSite ([122]), Phospho.ELM ([123]) or NetworKIN ([124]) (Table 2). Several machine learning methods have been used to generate prediction tools for phosphorylation sites. These prediction tools are based on artificial neural networks (NetPhos [125], NetPhosBac [126], NetPhosYeast [127]), convolutional neural networks (NetPhosPan[128]), position specific scoring matrices (GPS [129], ScanSite [130]), hidden Markov models (KinasePhos 1.0 [131], Predkin [132]), support vector machines (PHOSIDA [133], KinasePhos 2.0 [134]), logistic regression (DISPHOS [135]) or structural analysis (pKaPS [136]) (Table 2).
packages usually work "out of the box" with detailed manuals. NAMD and OpenMM [136–140]. Latest versions of MD software simulation. Some examples include GROMACS, AMBER, CHARMM. There are several software analysis frameworks available such as MDTRAJ [147], CPPTRAJ [148] or GROMACS [138]. We opted to use the CPPTRAJ module of AMBER for trajectory analysis [148]. For each simulation, root mean square deviation of both atomic and tools to enable first-time use of the software. All MD simulations performed in this work were carried out on a small workstation equipped with a 17-6700 Intel CPU processor and NVIDIA 1050-Ti GPU. With this setup, we achieved a simulation speed of 50 ns/day for a 30,000 atoms system, which could be hugely increased with a high-end workstation and high-performance computing clusters. In our case, we performed MD simulations using the OpenMM [140] software in an OpenCL platform with AMBER 14SB force field [141] and particle mesh Ewald (PME) electrostatics with an Ewald summation cutoff of 10 Å.

The initial structures were based on the nuclear magnetic resonance (NMR) structures of reduced human Cc (PDB ID: 2N9I; [142]), reduced human Y48pCMF Cc mutant (PDB ID: 2N3Y; [114]) and HuR -101 (PDB ID: 5SZW; [118]). All mutants and phosphorylated species structural models were built using the WT protein as a template. The force field parameters for phosphorylated tyrosine [143], unnatural amino acid pCMF [144] and heme group [145] were used to generate the topology and coordinate files required for the simulation, with the TLEAP module of AMBER [139].

For each simulation, the system was neutralized with sodium and chloride counter-ions according to the total charge of the protein, and solvated with optimal 3-charge, 4-point rigid water model (OPC) water molecules [146]. The whole system was subjected to 2500 steps of energy minimization at 298 K. Langevin thermostat was used to control the temperature with a friction coefficient of 1 ps$^{-1}$ and a step size of 0.002 ps. Each system was subjected to 500 ns MD simulation. To study the stability of the phosphomimetic variants, similar MD simulations were carried out at different temperatures (400 K and 500 K) and the system was subjected to 10 ns simulation.

There are several software analysis frameworks available such as MDTRAJ [147], CPPTRAJ [148] or GROMACS [138]. We opted to use the CPPTRAJ module of AMBER for trajectory analysis [148]. For each simulation, root mean square deviation of both atomic positions (RMSD, rms command on CPPTRAJ), atomic fluctuations (RMSF, atomicfluct command on CPPTRAJ), structure content (struct command on CPPTRAJ), hydrogen bonds (hbond command on CPPTRAJ), surface accessible to solvent (surf command on CPPTRAJ), secondary structure elements ($\alpha$-helix and $\beta$-sheet) are depicted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).
CPPTRAJ and native contacts (nativecontact command on CPPTRAJ) were analyzed to evaluate the perturbation induced by the addition of a phosphate group or a phosphomimetic variant at a specific position of the protein. Statistics of simulated proteins throughout the simulations are listed in Table 3.

2.2. Overexpression of phosphomimetic variants

2.2.1. Site-directed mutagenesis

Incorporation of the unnatural amino acid pCMF was accomplished by co-expressing the unnatural tRNA/amiinoacyl-tRNA synthetase pair specific for pCMF (encoded by pEVOL-pCMF plasmid) along with the protein of interest, the DNA sequence of which was mutated to a TAG amber codon at the desired incorporation site. The DNA encoding HuR$_{1-101}$ was cloned into the pGEX expression vector with an N-terminal 6xHis-tag and was used as a template for the mutagenic polymerase chain reaction (PCR). AMBER mutation was performed on a pBTR1 plasmid comprising the CYCS gene encoding human Cc, along with the CYC3 gene of the yeast Cc heme lyase. CYC3 is required for proper cytoplasmic maturation of human Cc, as it covalently attaches the heme group to the Cc apoprotein. Primers used for mutagenic PCR are detailed in Table S1.

2.2.2. Protein expression

The expression of WT Cc and its phosphomimetic variants (Y48E and Y48pCMF) were performed as previously described [144,149] in E. coli BL21 (DE3) strain. Briefly, ncAA-containing protein was expressed as follows: pre-culture was incubated at 37°C at 150 rpm in minimal medium M9 supplemented with ampicillin.

Table 3

| Protein | Average RMSD (Å)$^a$ | Average RG (Å)$^a$ | Average SAS (Å$^2$)$^b$ | Average RMSF (Å)$^c$ |
|---------|----------------------|-------------------|--------------------------|----------------------|
| Cc WT   | 0.85 ± 0.10          | 12.96 ± 0.04      | 39.20 ± 10.66            | 0.43 ± 0.07          |
| Cc Y48E | 0.71 ± 0.07          | 12.87 ± 0.04      | 61.33 ± 13.42            | 0.45 ± 0.07          |
| Cc Y48pCMF | 1.93 ± 0.08     | 13.24 ± 0.05      | 128.10 ± 13.81           | 0.65 ± 0.14          |
| Cc pY48 | 2.13 ± 0.11          | 13.20 ± 0.05      | 123.61 ± 22.67           | 0.66 ± 0.12          |
| HuR$_{1-101}$ WT | 1.59 ± 0.18 | 12.14 ± 0.07      | −17.25 ± 3.34            | 0.47 ± 0.05          |
| HuR$_{1-101}$ Y68E | 1.73 ± 0.19 | 12.21 ± 0.07      | 34.96 ± 12.93            | 0.83 ± 0.13          |
| HuR$_{1-101}$ Y68pCMF | 2.23 ± 0.25 | 12.33 ± 0.07      | −4.71 ± 8.46             | 1.26 ± 0.43          |
| HuR$_{1-101}$ pY68 | 2.35 ± 0.29 | 12.40 ± 0.08      | 102.76 ± 28.16           | 1.57 ± 0.53          |
| HuR$_{1-101}$ WT | 1.59 ± 0.18 | 12.14 ± 0.07      | 18.84 ± 10.54            | 0.87 ± 0.52          |
| HuR$_{1-101}$ Y63E | 1.56 ± 0.10 | 12.20 ± 0.07      | 66.36 ± 10.54            | 1.10 ± 0.70          |
| HuR$_{1-101}$ Y63pCMF | 1.66 ± 0.12 | 12.18 ± 0.06      | 140.11 ± 10.29           | 0.67 ± 0.32          |
| HuR$_{1-101}$ pY63 | 1.92 ± 0.10 | 12.23 ± 0.06      | 102.27 ± 21.22           | 0.65 ± 0.30          |

$^a$ Average RMSD/RG values were calculated for the plateau (last 100 ns). For HuR$_{1-101}$ trajectories, values were calculated for the structured domain (residues 22–97). For all simulations, RMSD drift was smaller than 0.001 Å/ns.

$^b$ Average SAS values were calculated for the plateau (last 100 ns). Values were calculated for the heme group (Cc trajectories), position 68 of HuR$_{1-101}$ (Tyr68 mutants’ trajectories) and position 63 of HuR$_{1-101}$ (Tyr63 mutants’ trajectories).

$^c$ Average RMSF values of Cc were calculated for the loop 80–91, for the stretch 42–49 of HuR$_{1-101}$ Tyr68 mutants, and for the stretch 52–69 of HuR$_{1-101}$ Tyr63 mutants.
(100 μg/mL) and chloramphenicol (20 μg/mL). Pre-induction of the pre-culture with 0.02% arabinose is necessary to ensure the presence of the unnatural tRNA/aaRS pair. Protein expression was induced at an optical density (OD600) of 0.6 with 0.02% arabinose and 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG); pCMF and δ-amino-levulinic acid were added after induction at a final concentration of 1 mM each. Cells were incubated for 20 h under agitation (150 rpm) at 30 °C before harvesting by centrifugation.

WT HuR1-101 and its phosphomimetic variants were expressed in the E. coli BL21 (DE3) strain. For WT HuR1-101 and non-containing pCMF mutants, the cells were cultured at 37 °C under agitation (150 rpm) in LB medium supplemented with 100 μg/mL ampicillin. When OD600 reached 0.6, protein expression was induced with 1 mM IPTG for 20 h under agitation (150 rpm) at 30 °C. For the HuR1-101 mutants Y63pCMF and Y68pCMF, cells were cultured at 37 °C at 150 rpm in minimal medium M9, supplemented with 100 μg/mL ampicillin and 20 μg/mL chloramphenicol. Protein expression was induced at OD600 of 0.6 with 0.02% arabinose and 1 mM IPTG, as described previously [144]. pCMF was added after induction at a final concentration of 1 mM. Cells were harvested by centrifugation, resuspended in the lysis buffer 20 mM Tris–HCl (pH 8.0), containing 800 mM NaCl, 10 mM imidazole, 1 mM phenylmethanesulfon fluoride, 0.2 mg/mL lysozyme and 0.02 mg/mL DNase, and ruptured by sonication. Afterwards, cellular debris was separated by centrifugation at 14,000 × g for 30 min at 4 °C.

Protein expression was monitored by SDS-PAGE and Western-Blot. Proteins were separated by SDS-PAGE and transferred onto polyvinylidene difluoride (PVDF) membranes. Membranes were blocked with 5% fat-free dry milk in phosphate buffered saline (PBS)-0.1% Tween-20. Immunoblotting was performed with specific mouse primary antibody against HuR (catalog number 32160702, Sigma) or 6xHis-tag (catalog number 11922416001, Roche) and horseradish peroxidase (HRP)-conjugated secondary antibody against mouse IgG (catalog number A4416, Sigma). The immunoreactive bands were detected using Immobilon Western Chemiluminescent HRP Substrate (Millipore).

3. Results and discussion

3.1. General workflow

A general workflow for protein phosphorylation studies is proposed in Fig. 2. The first decision to be made before carrying out any functional and/or structural analysis is choosing the source of phosphorylated protein. If the phosphorylated protein is obtained with a high yield, either by direct purification from tissues/cells extracts or further purification after a kinase in vitro reaction, the desired physicochemical analysis can be performed.

However, as pointed out in Section 1.1, obtaining large amounts of phosphorylated protein is not always feasible. For this reason, it is necessary to overexpress a phosphomimetic variant. At this stage, an in silico approach can help to indicate which variant is the most suitable for your system.

To perform the MD simulation screening, the 3D atomic coordinates for the desired protein (preferably from the crystallographic or NMR structure) must be available at the PDB. Homology modeling software such as SWISS-MODEL (https://swissmodel.expasy.org; [150]), MODELLER (https://salilab.org/modeller; [151]) or ROBETTA (http://robetta.bakerlab.org; [152]) can be used to generate the protein coordinates if they are not available. If X-ray diffraction data from homologous proteins is available crystallographic water molecules can be included in the model. Once structural models have been generated, a comparison of phosphorylation-induced perturbations should be performed by analyzing the MD simulations of the WT protein versus the phosphorylated species (see Section 2.1.2). Subsequent analysis of multiple MD simulations of candidate phosphomimetic variants helps to identify which molecule best mirrors the conformational changes induced by phosphorylation. Computational-based interaction network analysis may also be helpful to evaluate phosphomimetic constructs [153].

As soon as the selection of the phosphomimetic variant has been completed based on the MD data (decision-making tool), overexpression and purification of the chosen protein may be performed. This stage requires optimization of the expression conditions as mutations may affect protein folding and stability. Mimicking phosphorylation of residues exposed to the solvent – which are the most common phosphorylation sites – is often a straight forward process. However, if the phosphoresidue is located in the protein core (i.e. Tyr68 of HuR), the addition of negative charges may induce conformational changes or impair correct protein folding. MD simulations at varying conditions (temperature, ionic force or pH) can inform and detect possible difficulties in protein expression and folding.

Finally, MD simulations, in a feedback-like loop, can validate and assess the gathered functional/structural information of the phosphorylated/phosphomimetic protein.

3.2. Cytochrome c phosphorylation

A clear example where it was possible to verify the poor choice of glutamate to mimic pTyr is in Cc. The specific Cc kinase is unknown and a broad study on the functional implications of phosphorylation is unachievable with purified phosphorylated Cc because of low yields obtained from cell extracts [106,113,144,154,155]. The proposed methodology (see Section 3.1) was used to analyze different phosphomimetic variants (mutation by glutamate or substitution with pCMF) for the study of Cc Tyr48 phosphorylation (pY48) [144].

Initial MD calculations of WT Cc and modelled mutants Y48E, Y48pCMF and pY48 Cc revealed that the Y48pCMF mutation induces a localized distortion of the structure similar to that computed for the phosphorylated species. Radius of gyration (RG) values of phosphomimetic Y48pCMF and phosphorylated Cc displayed a similar behavior, while Y48E profile resembles that of WT Cc (Fig. S2). Indeed, surface accessible to solvent (SAS) analysis revealed that the solvation of the heme group in the nCα mutant (128.1 ± 18.2 Å²) mirrors more closely the solvation experienced by the phosphorylated variant (123.6 ± 22.7 Å²) as compared with the Y48E mutant (61.3 ± 13.4 Å²) (Fig. 3 and Fig. S2). The loop containing residues 20–30 showed the highest root mean square fluctuations (RMSF) for all mutants, and were similar to values previously reported for Thr28 and Thr58 mutants [113,155,156] (Fig. S2). RMSF values of residues surroundings position 48 from Y48pCMF Cc were slightly higher than those from pY48 Cc, and RMSF values from Y48E Cc were slightly lower in comparison with those from WT Cc. Notably, the RMSF values at the loop between residues 80–91, where one of the axial ligands (Met80) of the heme group is located, were similar for both phosphomimetic Y48pCMF and phosphorylated pY48 Cc species, whereas WT and Y48E mutant displayed analogous values. In summary, the Y48pCMF variant showed a behavior closer to the one displayed by phosphorylated cytochrome c with respect to the overall properties calculated by MD simulations.

This data was used as a starting point to characterize the Y48pCMF Cc mutant [114]. This mutant impairs Cc diffusion between respiratory complexes, boosts hemeprotein peroxidase and reactive oxygen species (ROS) scavenging activities, and hamper caspase-dependent apoptosis [114]. Further experimental NMR relaxation measurements and hydrogen–deuterium
Comparison of biophysical and functional properties of two phosphomimetic cytochrome c mutants with respect to the native phosphorylated protein.

|                  | in vivo pY48 Cc | Y48E Cc | Y48pCMF Cc |
|------------------|-----------------|---------|------------|
| Detection by anti-pTyr | Yes             | No      | No         |
| $E^\circ_{pH \, 7.0}$ (mV) | –               | 192.0 ± 5.0 | 208.5 ± 0.3 |
| Alkaline transition ($pK_a$) | 7.0             | [162,130] | [144,116] |
| Oxygen consumption, $V_{max}$ (50%) | 30%             | [154,126] | [114,100] |
| CcG activity, $K_m$ values (µM) | 3.7             | –       | –          |
| Caspase activity | –               | Unable to activate caspase-3 [154,126] | Decrease by 47.7% compared to WT [114,100] |
| Bound to cardiolipin | –               | 30% lower affinity [154,126] | 27% more affinity [114,100] |
| Peroxidase activity | –               | Cc peroxidase activity only inducible at high cardiolipin concentration [154,126] | In the absence of CL, the mutant shows 3 times more activity than WT [114,100] |

= : Not available.

3.3. HuR phosphorylation

We also applied the workflow detailed above to study HuR$_{1-101}$ phosphorylation. Perturbation analysis of the MD simulation of phosphorylated Tyr68 (pY68) HuR$_{1-101}$ revealed that phosphorylation at this position induces conformational changes at the $\alpha_1$ helix (residues 33–47), thereby increasing Tyr68 solvent accessibility (Fig. 4A, Fig. S3-4). SAS of position 68 in phosphorylated HuR$_{1-101}$ (102.7 ± 28.2 Å$^2$) was higher than WT HuR$_{1-101}$ (−17.3 ± 3.3 Å$^2$). Analogous MD simulations were performed to evaluate the behavior of phosphomimetic variants (Fig. S3-4). Substitution of Tyr68 by pCMF slightly destabilizes the $\alpha_1$ helix, as does Tyr68 phosphorylation (Fig. 4B, Fig. S3). RMSF values for this stretch are comparable between pY68 and Y68pCMF HuR$_{1-101}$ (Fig. S4). However, the solvent accessibility does not increase (−4.7 ± 8.5 Å$^2$), displaying similar values as the WT protein (Fig. S4). On the other hand, Tyr68-by-Glu68 substitution does increase the solvent accessibility of the Glu residue (35.0 ± 13.0 Å$^2$) without affecting the secondary structure of HuR$_{1-101}$ nor inducing significant conformational changes (Fig. S4). Interestingly, the Glu side-chain flips with respect to its initial conformation, popping out of the hydrophobic pocket (Fig. 4C). MD simulations of the Y68F mutant, in which Tyr68 is replaced with Phe, was carried out as a control to check that this substitution does not alter HuR$_{1-101}$ structure (Fig. 4D, Fig. S3).

Based on the findings of this data, we decided to study Tyr68 phosphorylation using ncAA substitution, as it better mirrors the generated perturbation. Nonetheless, HuR$_{1-101}$ Y68pCMF, like Y68A and Y68E mutants, could not be expressed in vitro. In fact, the only mutant at position 68 that was successfully expressed was Y68F HuR$_{1-101}$ recombinant variant (Fig. 4E). Evolutionary conservation analysis performed with ConSurf web server [159] revealed that the most conserved residue at position 68 is Tyr (88%) followed by Phe (12%). Intriguingly, ncAA HuR$_{1-101}$ mutant at position 63 (Y63pCMF) and the variant Y63F were all well expressed (Fig. 4E). A plausible explanation is that Tyr68 is located at the hydrophobic core of HuR$_{1-101}$, and as an aromatic amino acid also showed a higher peroxidase activity in the absence of cardiolipin with heightened dynamics in the Met80-containing loop [158].

Fig. 3. Cytochrome c conformational changes induced by Tyr48 phosphorylation. Surface representation of human Cc (A, WT in white; B, phosphorylated at Tyr48 in orange; C, Tyr-to-Glu mutant in gold; D, Tyr-to-pCMF mutant in pale rose) after MD simulation. Heme group is colored in green and position 48 in purple. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).
key in protein folding may further stabilize the protein. To test this hypothesis, we performed several MD simulations of WT, Y68F and Y68A HuR1-101 variants at two different temperatures (400 and 500 K) to analyze their respective stability (Fig. S5). Trajectory analyses indicated that the most stable mutant at position 68 is Y68F—in fact, the only one can be successfully expressed in vitro—and that Y68F correlation between native contacts and radius of gyration throughout the whole simulation is similar to that of WT HuR1-101. Altogether, these findings suggest a crucial role for the aromatic ring (provided by either Tyr or Phe) at position 68 for correct HuR1-101 folding.

A recent NMR analysis of HuR1-101 binding to an RNA oligonucleotide derived from adenine and uracil-rich element (ARE) 1 of the cyclooxygenase 2 mRNA revealed that the RNA binding site comprises Tyr63 among other residues in the β1 and β3 strands [114]. Moreover, Tyr63 has been found phosphorylated in Hodgkin’s lymphoma cell lines [160]. Further analysis performed by Yoon et al. [120] detected Y63 phosphorylation after in vitro kinase assays with recombinant purified JAK3 kinase and HuR. As the Y63pCMF mutant is recombinantly expressed, we wanted to explore how effectively it mirrors phosphorylation at this position. For Tyr63, similar simulations on phosphorylated (pY63) and phosphomimetic variants (Y63E and Y63pCMF) of HuR1-101 were carried out (Figs. S6–7). Tyr63 is located at the surface of the protein, where it contacts the unstructured N-terminal tail. Perturbation analysis of the Y63E HuR1-101 mutant revealed that substitution by Glu undermines the contacts between the N-terminal tail and the protein core. This mutant displayed higher values of RG, which is consistent with partial unfolding of the N-terminal (Fig. S6). This perturbation is not induced by phosphorylation nor substitution by ncAA pCMF (Fig. 5). Moreover, RMSF values of the area surrounding position 63 are higher than those for WT, pY63 and Y63pCMF HuR1-101. Analysis of the Y63E mutant could thus lead to an overestimation of the effects of phosphorylation. For this reason, the functional implications of Tyr63 phosphorylation will be assessed using the Y63pCMF HuR1-101 mutant in future studies.

4. Concluding remarks

The structural and functional analysis of phosphorylated proteins can be achieved by several approaches. The purpose of the herein presented workflow is to help researchers choose the most suitable alternative when an extensive study of the functional effects of phosphorylation is unfeasible with phosphorylated spe-
cies purified from cells and/or tissues. MD simulations have become more accessible due to the development of user-friendly software packages and the increase in computing power at a modest cost. However, computational approaches are often ignored by molecular biologists. For this reason, we believe that a combination of computational and biochemical approaches will help to better understand how phosphorylation events affect protein structure and function.

CRediT authorship contribution statement

Gonzalo Pérez-Meijas: Conceptualization, Methodology, Software, Visualization, Investigation, Writing - original draft, Writing - review & editing. Alejandro Velázquez-Cruz: Writing - original draft, Writing - review & editing. Alejandra Guerra-Castellano: Investigation, Writing - original draft, Writing - review & editing. Blanca Baños-Jaime: Resources. Antonio Díaz-Quintana: Software, Writing - review & editing. Katiuska González-Arzola: Writing - review & editing. Miguel Ángel De Rosa: Supervision, Writing - review & editing, Project administration, Funding acquisition. Irene Díaz-Moreno: Conceptualization, Supervision, Writing - review & editing, Project administration, Funding acquisition.

Acknowledgements

This work was supported by the Ministry of Science and Innovation (PGC2018-096049-B-100), European Regional Development Fund (FEDER), Andalusian Government (BIO-198, US-1254317 and US-1257019) and TA Instruments. G.-M. and A.V.-C. were awarded a PhD fellowship from the Spanish Ministry of Education, Culture and Sport (FP17/04604 and FP16/01513). Molecular graphics and analyses were performed with the UCSF Chimera software, developed by the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

Conflict of interest

All authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2020.06.043.

References

[1] Duan G, Walther D. The roles of post-translational modifications in the context of protein interaction networks. PLoS Comput Biol 2015;11:e1004049.
[2] Draznin B. Molecular mechanisms of insulin resistance: serine phosphorylation of insulin receptor substrate-1 and increased expression of p85alpha: the two sides of a coin. Diabetes 2006;55:2397–29.
[3] Hunter T. Tyrosine phosphorylation: thirty years and counting. Curr Opin Cell Biol 2005;21:140–6.
[4] Sanderson TH, Mahapatra G, Pecina P, et al. Cytochrome c is tyrosine 97 phosphorylated by neuroprotective insulin treatment. PLoS ONE 2013;8:e78627.
[5] Guo X, Wang X, Wang Z, et al. Site-specific proteasome phosphorylation controls cell proliferation and tumorigenesis. Nat Cell Biol 2016;18:202–12.
[6] Ardito F, Giuliani M, Perrone D, Troiano G, Lo Muzio L. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). Int J Mol Med 2017:40:271–80.
[7] Hunter T. Why nature chose phosphate to modify proteins. Philos Trans R Soc B Biol Sci 2012;367:2513–6.
[8] Fakami Y, Lipmann F. Reversal of Rous sarcoma-specific immunoglobulin phosphorylation on tyrosine (ADP as phosphate acceptor) catalyzed by the src gene kinase. Proc Natl Acad Sci U S A 1983;80:1872–6.
[9] Kole HK, Abdel-Ghany M, Racker E. Specific dephosphorylation of phosphoproteins by protein-serine and -tyrosine kinases. Proc Natl Acad Sci U S A 1988;85:5649–53.
[10] Hornbeck PV, Chabra I, Korshunov JM, Szpyrek E, Zhang B. PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation. Proteomics 2004;4:645–54.
[11] Patrick J, Baker EN, Delbarre LTJ. Histidine phosphorylation in biological systems. Biochim Biophys Acta, Proteins 2008;1784:100–5.
[12] Ciesla J, Fraczyk T, Rode W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. Acta Biochim Pol 2011;58:137–46.
[13] Jakobscheva DM, Radyo P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 2004;32:1037–49.
[14] Mitreà DM, Grace CR, Buljan M, et al. Structural polymorphism in the N-terminal oligomerization domain of NPM1. Proc Natl Acad Sci U S A 2014;111:4466–71.
[15] Díaz-Moreno I, Hollingsworth D, Kelly G, et al. Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. Nucleic Acids Res 2010;38:5193–205.
[16] Díaz-Moreno I, Hollingsworth D, Frenkel T, et al. Phosphorylation-mediated unfolding of a KH domain regulates KSRP localization via 14–3–3 binding. Nat Struct Mol Biol 2009;16:238–46.
[17] Bah A, Vernon RM, Siddiqui Z, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature 2015;519:106–9.
[18] Schmidt SR, Schweikart F, Andersson ME. Current methods for phosphoprotein isolation and enrichment. J Chromatogr, B: Anal Technol Biomed Life Sci 2007;849:154–62.
[19] Díaz-Moreno I, Hollingsworth D, Frenkel T, et al. Phosphorylation-mediated unfolding of a KH domain regulates KSRP localization via 14–3–3 binding. Nat Struct Mol Biol 2009;16:238–46.
[20] Bai J, Vernon RM, Siddiqui Z, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature 2015;519:106–9.
[21] Bjerum EJ, Jensen JH, Tolborg JL. pICalculax: improved prediction of pIC50 values of protein kinase inhibitors. Bioorg Med Chem Lett 2010;20:73–6.
[22] Timberg G, Liu D, Popp C, et al. Enzymatic phosphorylation of histidine as a tool for modifying proteins and peptide mimics. J Am Chem Soc 2012;134:15004–11.
[23] Klimke RM, Klinzing K, Blumel S, et al. Comparison of protein phosphorylation in cell signaling and its use as targeted therapy. Proteomics 2017;17:2573–82.
[24] Zhu K, Zhao J, Lubman DM, Miller FR, Barder TJ. Protein pI shifts due to dephosphorylation of tyrosine. Proteomics 2014;14:3078–87.
[25] Kole HK, Abdel-Ghany M, Racker E. Specific dephosphorylation of phosphoproteins by protein-serine and -tyrosine kinases. Proc Natl Acad Sci U S A 1988;85:5649–53.
[26] Hornbeck PV, Chabra I, Korshunov JM, Szpyrek E, Zhang B. PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation. Proteomics 2004;4:645–54.
[27] Patrick J, Baker EN, Delbarre LTJ. Histidine phosphorylation in biological systems. Biochim Biophys Acta, Proteins 2008;1784:100–5.
[28] Ciesla J, Fraczyk T, Rode W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. Acta Biochim Pol 2011;58:137–46.
[29] Jakobscheva DM, Radyo P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 2004;32:1037–49.
[30] Mitreà DM, Grace CR, Buljan M, et al. Structural polymorphism in the N-terminal oligomerization domain of NPM1. Proc Natl Acad Sci U S A 2014;111:4466–71.
[31] Díaz-Moreno I, Hollingsworth D, Kelly G, et al. Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. Nucleic Acids Res 2010;38:5193–205.
[32] Díaz-Moreno I, Hollingsworth D, Frenkel T, et al. Phosphorylation-mediated unfolding of a KH domain regulates KSRP localization via 14–3–3 binding. Nat Struct Mol Biol 2009;16:238–46.
[33] Bai J, Vernon RM, Siddiqui Z, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature 2015;519:106–9.
[34] Schmidt SR, Schweikart F, Andersson ME. Current methods for phosphoprotein isolation and enrichment. J Chromatogr, B: Anal Technol Biomed Life Sci 2007;849:154–62.
[35] Díaz-Moreno I, Hollingsworth D, Frenkel T, et al. Phosphorylation-mediated unfolding of a KH domain regulates KSRP localization via 14–3–3 binding. Nat Struct Mol Biol 2009;16:238–46.
[36] Bai J, Vernon RM, Siddiqui Z, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature 2015;519:106–9.
[37] Bjerum EJ, Jensen JH, Tolborg JL. pICalculax: improved prediction of pIC50 values of protein kinase inhibitors. J Am Chem Soc 2012;134:15004–11.
[38] Klimke RM, Klinzing K, Blumel S, et al. Comparison of protein phosphorylation in cell signaling and its use as targeted therapy. Proteomics 2017;17:2573–82.
[39] Zhu K, Zhao J, Lubman DM, Miller FR, Barder TJ. Protein pI shifts due to dephosphorylation of tyrosine. Proteomics 2014;14:3078–87.
[40] Kole HK, Abdel-Ghany M, Racker E. Specific dephosphorylation of phosphoproteins by protein-serine and -tyrosine kinases. Proc Natl Acad Sci U S A 1988;85:5649–53.
Jin X, Park O-J, Hong SH. Incorporation of non-standard amino acids into proteins: challenges, recent achievements, and applications. ACS Chem Biol 2019;14:331–40.

Park HS, Hohn MJ, Umehara T, et al. Expanding the genetic code of Escherichia coli with phosphoserine. Science (80-) 2011;333:1151–4.

Zisch AH, Pazzagli C, Freeman AL, et al. Replacing two conserved tyrosines of the histone H4 tail revealed with a human nucleosome core particle with genetically-incorporated H4 tetra-acetylation. Sci Rep 2015;5:17204.

Chemla Y, Ozer E, Shaferman M, et al. Simplified methodology for a modular and genetically expandable protein synthesis in cell-free systems. Synth Syst Biotechnol 2019;4:189–96.

Kawahata N, Yang MG, Luke GP, et al. A novel phosphotyrosine mimetic 4’-carboxymethyloxy-3’-phosphono-phenylalanine (Cpp): exploitation in the design of nonpeptide inhibitors of p5306(Ser20) SH2 domain. Bioorg Med Chem Lett 2001;11:2319–23.

Erlanson DA, McDowell BS, He MM, et al. Discovery of a new phosphotyrosine mimetic for PTPIP using breakaway tethering. J Am Chem Soc 2003;125:6502–3.

Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in Escherichia coli. Nat Biotechnol 2000;18:1364–72.

Fan C, Ip K, Söll D. Expanding the genetic code of Escherichia coli by programmed cell death in plants. Mol Cell 2017;292:185–95.

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. Structural and functional analysis of novel human cytochrome c targets in apoptosis. Mol Cell Proteomics 2014;13. 1439 LP – 1456

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. New Arabidopsis thaliana cytochrome c partners: a look into the elusive role of cytochrome c in programmed cell death in plants. Mol Cell Proteomics 2013;12:3666–76.

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. Improved stability and reduced activity of N-acetylated-mono-methyl-L-lysine into histone proteins by cell-based and cell-free protein synthesis. Nat Chem Biol 2017;13:845–9.

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. Structural and functional analysis of novel human cytochrome c targets in apoptosis. Mol Cell Proteomics 2014;13. 1439 LP – 1456

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. An enhanced system for unnatural amino acids. ACS Chem Biol 2019;14:1560–73.

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. New Arabidopsis thaliana cytochrome c partners: a look into the elusive role of cytochrome c in programmed cell death in plants. Mol Cell Proteomics 2013;12:3666–76.

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. Structural and functional analysis of novel human cytochrome c targets in apoptosis. Mol Cell Proteomics 2014;13. 1439 LP – 1456

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. New Arabidopsis thaliana cytochrome c partners: a look into the elusive role of cytochrome c in programmed cell death in plants. Mol Cell Proteomics 2013;12:3666–76.

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. Structural and functional analysis of novel human cytochrome c targets in apoptosis. Mol Cell Proteomics 2014;13. 1439 LP – 1456

Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, et al. An enhanced system for unnatural amino acids. ACS Chem Biol 2019;14:1560–73.

Zhu P, Galken PR, Mehl RA, Cooley RB. A highly versatile expression system for the production of multiple phosphorylated proteins. ACS Chem Biol 2019;14:1564–72.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Park HS, Hohn MJ, Umehara T, et al. Expanding the genetic code of Escherichia coli with phosphoserine. Science (80-) 2011;333:1151–4.

Wakamori M, Fujii Y, Suka N, et al. Intra- and inter-nucleosomal interactions of the histone H4 tail revealed with a human nucleosome core particle with genetically-incorporated H4 tetra-acetylation. Sci Rep 2015;5:17204.

Chemla Y, Ozer E, Shaferman M, et al. Simplified methodology for a modular and genetically expandable protein synthesis in cell-free systems. Synth Syst Biotechnol 2019;4:189–96.

Kawahata N, Yang MG, Luke GP, et al. A novel phosphotyrosine mimetic 4’-carboxymethyloxy-3’-phosphono-phenylalanine (Cpp): exploitation in the design of nonpeptide inhibitors of p5306(Ser20) SH2 domain. Bioorg Med Chem Lett 2001;11:2319–23.

Erlanson DA, McDowell BS, He MM, et al. Discovery of a new phosphotyrosine mimetic for PTPIP using breakaway tethering. J Am Chem Soc 2003;125:6502–3.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.

Thorsness PE, Koshland DE. Constitutive activation of Mek1 by mutation of serine-199: Preliminary report. Exp Gerontol 1987;22:1042–5.
The page contains a list of authors, followed by a series of references to scientific publications. The references are formatted in a traditional academic style, with the authors' names, publication years, and titles. The titles cover a range of topics from mitochondrial function to phosphorylation sites, and from computational methods to protein structure prediction. The references are cited in a manner that is typical for academic literature, with proper use of punctuation and formatting to indicate publication types (e.g., journals, books, conference proceedings). The overall content indicates a focus on biochemistry, molecular biology, and computational methods in the study of protein function and interactions.