Computational and Structural Biotechnology Journal 18 (2020) 1852–1863

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

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

ARTICLE INFO

Article history:
Received 30 March 2020
Received in revised form 29 June 2020
Accepted 30 June 2020
Available online 7 July 2020

Keywords:
Molecular biology
Molecular dynamics
Non-canonical amino acid
Phosphomimetic
Protein phosphorylation
Protein structure

ABSTRACT

Post-translational modifications of proteins expand their functional diversity, regulating the response of cells to a variety of stimuli. Among these modifications, phosphorylation is the most ubiquitous and plays a prominent role in cell signaling. The addition of a phosphate often affects the function of a protein by altering its structure and dynamics. However, these alterations are often difficult to study and the functional and structural implications remain unresolved. New approaches are emerging to overcome common obstacles related to the production and manipulation of these samples. Here, we summarize the available methods for phosphoprotein purification and phosphomimetic engineering, highlighting the advantages and disadvantages of each. We propose a general workflow for protein phosphorylation analysis combining computational and biochemical approaches, building on recent advances that enable user-friendly and easy-to-access Molecular Dynamics simulations. We hope this innovative workflow will inform the best experimental approach to explore such post-translational modifications. We have applied this workflow to two different human protein models: the hemeprotein cytochrome *c* and the RNA binding protein HuR. Our results illustrate the usefulness of Molecular Dynamics as a decision-making tool to design the most appropriate phosphomimetic variant.

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 from 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 [6,8,9]. 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 *disorder to order* and *order to disorder* transitions. An example of *order to disor-*

* Corresponding author.
E-mail address: idiazmoreno@us.es (I. Díaz-Moreno).

https://doi.org/10.1016/j.csbj.2020.06.043
2001-0370/

der transition is nucleophosmin (NPM1), sequential phosphorylation of solvent exposed Ser and Thr regulates the thermodynamic stability of the NPM1 oligomerization domain. This phosphorylation event promotes the exposure of other sites buried within the structure for subsequent phosphorylation and drives the switch from structured pentamer to disordered monomer [14]. K-homology splicing regulator protein (KSRP) contains four K-homology domains (KH1-4), in which KH1 phosphorylation acts as a conformational switch between folded (non-phosphorylated protein) and essentially unfolded (phosphorylated protein) states [15]. Unfolding of KH1 domain creates a 14-3-3ζ binding site which prevents the recruitment of the exosome [16]. On the other hand, phosphorylation of IDR eukaryotic translation initiation factor 4E (eIF4E)-binding protein 2 (4E-BP2) protein constitutes an example of *disorder to order* transition by inducing folding into a four-stranded β-domain, which blocks the accessibility of eIF4E [17].

1.1. Overview of phosphoprotein detection and purification methods

Phosphoproteomic investigations usually require purification and often enrichment of endogenous proteins from cells or tissues. 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 (p*I*), 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 p*I* 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 p*I* 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 phosphoresidues and protein folding [20,21].

Many phosphoprotein 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 Fe^{3+}, 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 (Zn^{2+} or Mn^{2+}) 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 phosphoprotein 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 TiO_2, 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 phosphoresidues 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, TiO_2-MOAC and/or IAC, either sequentially or as complementary techniques, to optimize protein enrichment into phosphotyrosine [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 phosphotyrosine proteins [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 (GSH), or maltose-binding protein—to facilitate purification by affinity chromatography (Ni^{2+}-IMAC, GSH-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 co-translationally 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.

Table 1
Non-canonical amino acids mimicking phosphoresidues.

Phosphorylated residues	Unnatural amino acid	Citation
Phosphotyrosine	4′-Carboxymethyloxy-3′-phosphonophenylalanine	[85]
	N-[4-(2-{2-[3-(2-Bromo-acetylamino)-propionylamino]-3-hydroxy-propionylamino}-ethyl)-phenyl]-oxalamic acid	[86]
	3-(4-{2-[2-(2-Bromo-acetylamino)-ethyldisulfanyl]-ethylcarbamoyl}-cyclohexylcarbamoyl)-pyrazine-2-carboxylic acid	[86]
	Photocaged phosphotyrosine	[73]
	Sulfotyrosine	[87]
	p-Carboxymethyl-L-phenylalanine	[88]
	p-(Phosphonoamino)-phenylalanine	[89]
	O-phosphotyrosine	[90]
	Benzyl 2-benzyloxycarbonylamino-3-[4-bis (dimethylamino)phosphoryloxyphenyl] propanoate	[91]
	4-Phosphomethyl-L-phenylalanine	[77]
Phosphothreonine	Photocaged phosphothreonine	[73]
	O-phospho-L-threonine	[92]
Phosphoserine	Photocaged phosphoserine	[73]
	Phosphonomethylene 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 pEVOL, 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 against phosphatases or gene deletion of these enzymes for elevated intracellular levels of phosphoresidues [73,74].

A common issue encountered when introducing phosphorylated residues, is the lack of permeability of cell membranes to highly-charged ncAAs [62,75]. Two main strategies have been developed to increase cellular uptake of phosphoresidues and ncAAs: 1) site-directed mutagenesis of periplasmic transporters for ncAA-specific internalization [76], and 2) addition of ncAAs to the culture medium as dipeptides with any natural amino acid,

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 overexpressed 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- to 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 (*p*CMF) 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* (C*c*) 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.

C*c* is an essential metalloprotein for mitochondrial metabolism and homeostasis, playing a dual-purpose role in cell life and death. Under homeostasis, C*c* acts as an electron carrier within the respiratory chain in the mitochondria. However, stress stimuli induce the release of C*c* to the cytosol and nucleus where it interacts with several targets [96–102]. In the cytosol, C*c* triggers programmed cell death (PCD) by activating the caspase cascade [103–105]. In the nucleus, C*c* is involved in DNA damage response through its interaction with histone chaperones [99,100]. All of these C*c* 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 C*c*, 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 ($HuR_{1\text{-}101}$) (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 RRMs [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], Predikin [132]), support vector machines (PHOSIDA [133], KinasePhos 2.0 [134]), logistic regression (DISPHOS [13]) or structural analysis (pKaPS [135]) (Table 2).

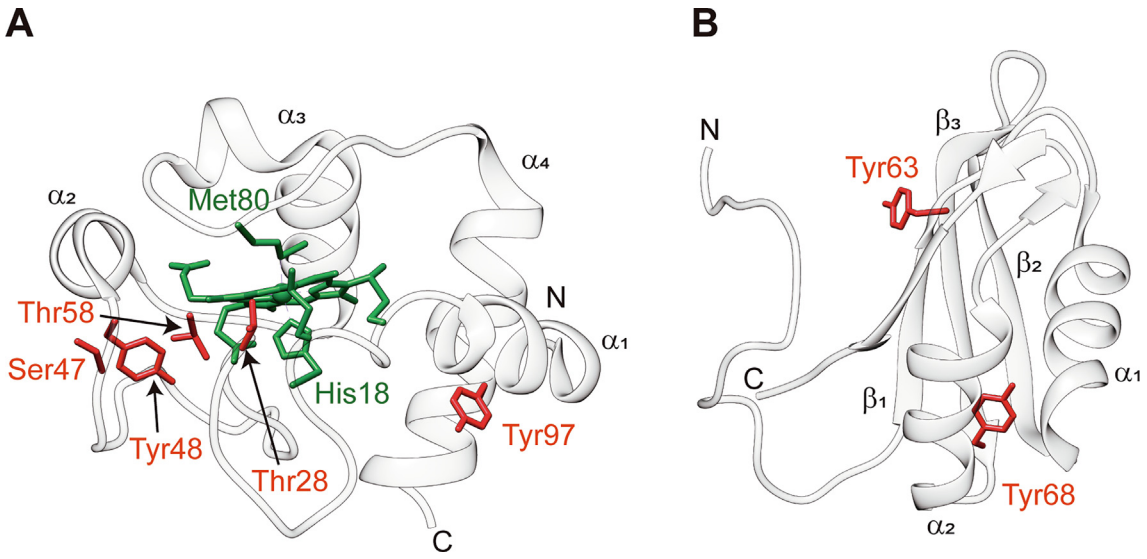

Fig. 1. Phosphorylation sites of human cytochrome *c* and HuR. Richardson representations of human C*c* (A; PDB ID: 2N9I; [142,113]) and HuR_{1-101} (B; PDB ID: 5SZW; [118,114]). Heme group and axial ligands (His18 and Met80) of C*c* are colored in green. Phosphorylation sites described in the literature for mammalian C*c* and HuR are highlighted in red. Thr at position 58 is the most conserved amino acid from bacteria to mammals, but it is not present in human C*c*. N-terminal and C-terminal of the polypeptide chains are marked with N and C, respectively. Secondary structure elements (α-helix and β -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).

Table 2
Databases and prediction tools to study protein phosphorylation.

	Databases	
Phosphosite	https://www.phosphosite.org	[122]
Phospho.ELM	http://phospho.elm.eu.org	[123]
NetworKIN	http://www.networkin.info	[124]
	Prediction tools	
DISPHOS	http://www.dabi.temple.edu/disphos/	[13]
GPS 5.0	http://gps.biocuckoo.cn	[129]
Scansite 4.0	http://scansite.mit.edu	[130]
pKaPS	http://mendel.imp.ac.at/sat/pkaPS	[135]
KinasePhos 2.0	http://KinasePhos2.mbc.nctu.edu.tw	[134]
NetPhos	https://services.healthtech.dtu.dk/service.php?NetPhos-3.1	[125]
NetPhosBac	https://services.healthtech.dtu.dk/service.php?NetPhosBac-1.0	[126]
NetPhosYeast	https://services.healthtech.dtu.dk/service.php?NetPhosYeast-1.0	[127]
NetPhospan	https://services.healthtech.dtu.dk/service.php?NetPhospan-1.0	[128]
PHOSIDA	http://phosida.de	[133]
Predikin	http://predikin.biosci.uq.edu.au	[132]

2.1.2. Analysis of phosphorylation-induced perturbations on protein structure and dynamics

Comparison between MD simulations on wild-type (WT) and phosphorylated protein species enables us to anticipate possible perturbations in the target structure and dynamics induced by phosphorylation at certain positions. MD simulations allows the evaluation of the motions of residues and atoms within a protein structure. This *in silico* information may be useful in assessing the behavior of phosphomimetic variants (Fig. 2). The chosen phosphomimetic variant should present similar dynamic properties and conformational changes as the phosphorylated protein.

Due to the extensive development of MD simulations, there is a broad range of software packages that can be used to perform the simulation. Some examples include GROMACS, AMBER, CHARMM, NAMD and OpenMM [136–140]. Latest versions of MD software packages usually work "out of the box" with detailed manuals 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 i7-6700 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 C*c* (PDB ID: 2N9I; [142]), reduced human Y48*p*CMF C*c* mutant (PDB ID: 2N3Y; [114]) and HuR_{1-101} (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 *p*CMF [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 chlorine 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 (RSMD, *rms* command on CPPTRAJ), atomic fluctuations (RMSF, *atomicfluct* command on CPPTRAJ), structure content (*secstruct* command on CPPTRAJ), hydrogen bonds (*hbond* command on CPPTRAJ), surface accessible to solvent (*surf* command on

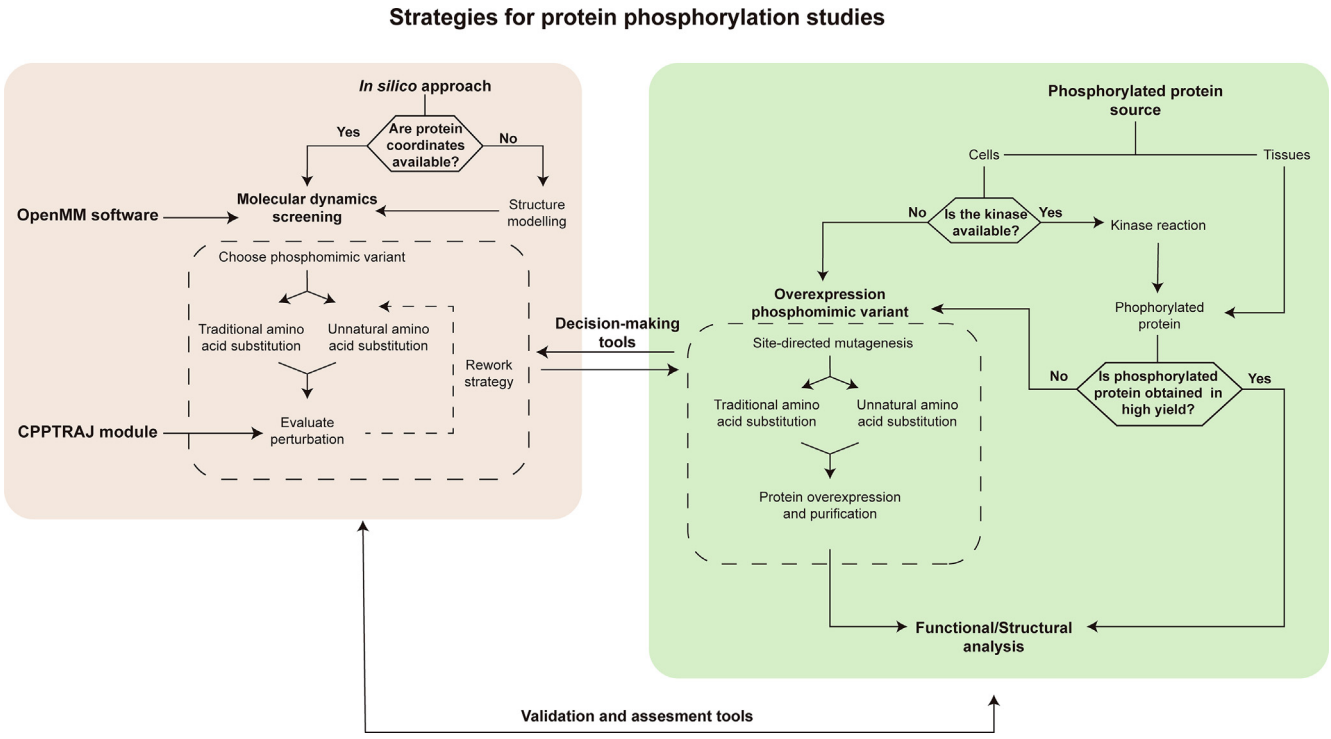

Fig. 2. Proposed experimental approach for protein phosphorylation analysis. The decision-making at each step is the result of assessing *in silico* approaches with wet-lab techniques.

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 *p*CMF was accomplished by co-expressing the unnatural tRNA/aminoacyl-tRNA synthetase pair specific for *p*CMF (encoded by pEVOL-*p*CMF 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 $HuR_{1\text{-}101}$ 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 C*c*, along with the CYC3 gene of the yeast C*c* heme lyase. CYC3 is required for proper cytoplasmic maturation of human C*c*, as it covalently attaches the heme group to the C*c* apoprotein. Primers used for mutagenic PCR are detailed in Table S1.

2.2.2. Protein expression

The expression of WT C*c* and its phosphomimetic variants (Y48E and Y48*p*CMF) 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
Statistics of simulated proteins along MD trajectories.

	Average RMSD (Å)[a]	Average RG (Å)[a]	Average SAS (Å^2)[b]	Average RMSF (Å)[c]
C*c* WT	0.85 ± 0.10	12.96 ± 0.04	39.20 ± 10.66	0.43 ± 0.07
C*c* Y48E	0.73 ± 0.07	12.87 ± 0.04	61.33 ± 13.42	0.45 ± 0.07
C*c* Y48*p*CMF	1.93 ± 0.08	13.24 ± 0.05	128.10 ± 18.21	0.65 ± 0.14
C*c* pY48	2.13 ± 0.11	13.20 ± 0.05	123.61 ± 22.67	0.66 ± 0.12
$HuR_{1\text{-}101}$ WT	1.59 ± 0.18	12.14 ± 0.07	−17.25 ± 3.34	0.47 ± 0.05
$HuR_{1\text{-}101}$ Y68E	1.73 ± 0.19	12.21 ± 0.07	34.96 ± 12.93	0.83 ± 0.13
$HuR_{1\text{-}101}$ Y68*p*CMF	2.23 ± 0.25	12.33 ± 0.07	−4.71 ± 8.46	1.26 ± 0.43
$HuR_{1\text{-}101}$ pY68	2.35 ± 0.29	12.40 ± 0.08	102.76 ± 28.16	1.57 ± 0.53
$HuR_{1\text{-}101}$ WT	1.59 ± 0.18	12.14 ± 0.07	18.84 ± 10.54	0.87 ± 0.52
$HuR_{1\text{-}101}$ Y63E	1.56 ± 0.10	12.20 ± 0.07	66.36 ± 10.54	1.10 ± 0.70
$HuR_{1\text{-}101}$ Y63*p*CMF	1.66 ± 0.12	12.18 ± 0.06	140.11 ± 10.29	0.67 ± 0.32
$HuR_{1\text{-}101}$ 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 $HuR_{1\text{-}101}$ 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 (C*c* trajectories), position 68 of $HuR_{1\text{-}101}$ (Tyr68 mutants' trajectories) and position 63 of $HuR_{1\text{-}101}$ (Tyr63 mutants' trajectories).
[c] Average RMSF values of C*c* were calculated for the loop 80–91, for the stretch 42–49 of $HuR_{1\text{-}101}$ Tyr68 mutants, and for the stretch 52–69 of $HuR_{1\text{-}101}$ 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 (OD_{600}) of 0.6 with 0.02% arabinose and 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG); *p*CMF and δ-aminolevulinic 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 HuR_{1-101} and its phosphomimetic variants were expressed in the *E. coli* BL21 (DE3) strain. For WT HuR_{1-101} and non-containing *p*CMF mutants, the cells were cultured at 37 °C under agitation (150 rpm) in LB medium supplemented with 100 μg/mL ampicillin. When OD_{600} reached 0.6, protein expression was induced with 1 mM IPTG for 20 h under agitation (150 rpm) at 30 °C. For the HuR_{1-101} mutants Y63*p*CMF and Y68*p*CMF, 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 OD_{600} of 0.6 with 0.02% arabinose and 1 mM IPTG, as described previously [144]. *p*CMF 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 phenylmethanesulfonyl 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 modelling 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 a 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 C*c*. The specific C*c* kinase is unknown and a broad study on the functional implications of phosphorylation is unachievable with purified phosphorylated C*c* 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 *p*CMF) for the study of C*c* Tyr48 phosphorylation (pY48) [144].

Initial MD calculations of WT C*c* and modelled mutants Y48E, Y48*p*CMF and pY48 C*c* revealed that the Y48*p*CMF mutation induces a localized distortion of the structure similar to that computed for the phosphorylated species. Radius of gyration (RG) values of phosphomimetic Y48*p*CMF and phosphorylated C*c* displayed a similar behavior, while Y48E profile resembles that of WT C*c* (Fig. S2). Indeed, surface accessible to solvent (SAS) analysis revealed that the solvation of the heme group in the ncAA mutant (128.1 ± 18.2 $Å^2$) mirrors more closely the solvation experienced by the phosphorylated variant (123.6 ± 22.7 $Å^2$) as compared with the Y48E mutant (61.3 ± 13.4 $Å^2$) (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 Y48*p*CMF C*c* were slightly higher than those from pY48 C*c*, and RMSF values from Y48E C*c* were slightly lower in comparison with those from WT C*c*. 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 Y48*p*CMF and phosphorylated pY48 C*c* species, whereas WT and Y48E mutant displayed analogous values. In summary, the Y48*p*CMF 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 Y48*p*CMF C*c* mutant [114]. This mutant impairs C*c* diffusion between respiratory complexes, boosts hemeprotein peroxidase and reactive oxygen species (ROS) scavenging activities, and hampers caspase-dependent apoptosis [114]. Further experimental NMR relaxation measurements and hydrogen–deuterium

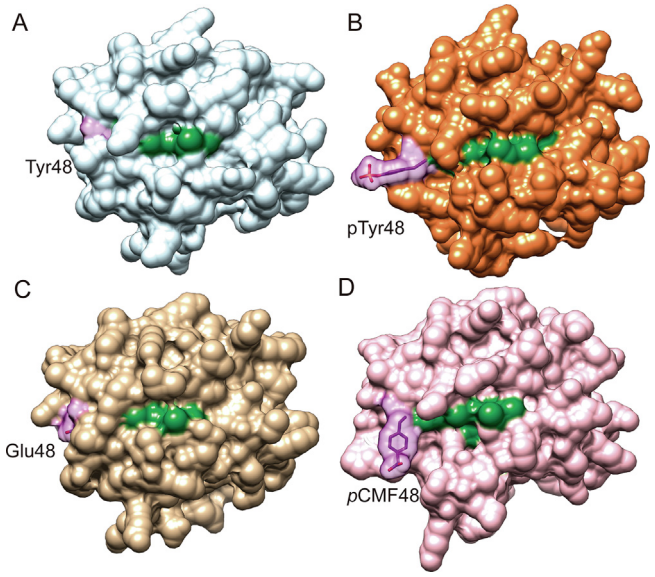

Fig. 3. Cytochrome *c* conformational changes induced by Tyr48 phosphorylation. Surface representation of human C*c* (A, WT in white; B, phosphorylated at Tyr48 in orange; C, Tyr-to-Glu mutant in gold; D, Tyr-to-*p*CMF mutant in pale rose) after MD simulation. Heme group is colored in green and position 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).

exchange experiments on the Y48*p*CMF variant revealed that the structure of the ncAA mutant exhibits enhanced dynamics around the *p*CMF-containing loop (also observed in MD simulations), which could explain the changes observed in its functionality (Table 4). Upon cardiolipin binding, C*c* undergoes conformational changes which allows H_2O_2 to access the heme crevice [157]. In the absence of cardiolipin, Y48*p*CMF C*c* displayed higher peroxidase activity compared to the Y48E phosphomimetic variant. This disparity in activity may be explained by the difference in the heme group SAS, which is two-times higher for the ncAA variant. Another explanation could be the increased dynamics of the Met80-containing loop for the Y48*p*CMF variant. The Y48H C*c* variant (naturally occurring mutation related to thrombocytopenia) also showed a higher peroxidase activity in the absence of cardiolipin with heightened dynamics in the Met80-containing loop [158].

3.3. HuR phosphorylation

We also applied the workflow detailed above to study $HuR_{1\text{-}101}$ phosphorylation. Perturbation analysis of the MD simulation of phosphorylated Tyr68 (pY68) $HuR_{1\text{-}101}$ revealed that phosphorylation at this position induces conformational changes at the α_1 helix (residues 33–47), thereby increasing Tyr68 solvent accessibility (Fig. 4A, Fig. S3-4). SAS of position 68 in phosphorylated $HuR_{1\text{-}101}$ (102,7 ± 28.2 Å^2) was higher than WT $HuR_{1\text{-}101}$ (−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 *p*CMF slightly destabilizes the α_1 helix, as does Tyr68 phosphorylation (Fig. 4B, Fig. S3). RMSF values for this stretch are comparable between pY68 and Y68*p*CMF $HuR_{1\text{-}101}$ (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 $HuR_{1\text{-}101}$ 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 $HuR_{1\text{-}101}$ 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, $HuR_{1\text{-}101}$ Y68*p*CMF, 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 $HuR_{1\text{-}101}$ 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 $HuR_{1\text{-}101}$ mutant at position 63 (Y63*p*CMF) and the variant Y63F were all well expressed (Fig. 4E). A plausible explanation is that Tyr68 is located at the hydrophobic core of $HuR_{1\text{-}101}$, and as an aromatic amino acid

Table 4
Comparison of biophysical and functional properties of two phosphomimetic cytochrome *c* mutants with respect to the native phosphorylated protein.

	in vivo pY48 C*c* [110,96]	Y48E C*c* [154,126]	Y48*p*CMF C*c* [144,116]
Detection by anti-pTyr	Yes	No [154,126]	No [144,116]
$E'^{0}_{,pH\ 7.0}$ (mV)	–	192.0 ± 5.0 [154,126] 187.3 ± 0.8 [162,130]	208.5 ± 0.3 [144,116]
Alkaline transition (pK_a)		7.0 [162,130]	6.3 [144,116]
Oxygen consumption, V_{max} (% of diminished capacity)	50%	30% [154,126]	60% [114,100]
CcO activity, K_m values (μM)	3.0	3.7 [154,126]	–
Caspase activity	–	Unable to activate caspase-3 [154,126] Decrease by 60% (*ca.*) compared to WT [162,130]	Decrease by 47.7% compared to WT [114,100]
Bound to cardiolipin	–	30% lower affinity [154,126]	27% more affinity [114,100]
Peroxidase activity	–	C*c* peroxidase activity only inducible at high cardiolipin concentration [154,126] Larger than WT in the absence of CL [162,130]	In the absence of CL, the mutant shows 3 times more activity than WT [114,100] In the presence of CL, the peroxidase activity increases twice as much as in the absence of CL [114,100]

– : Not available.

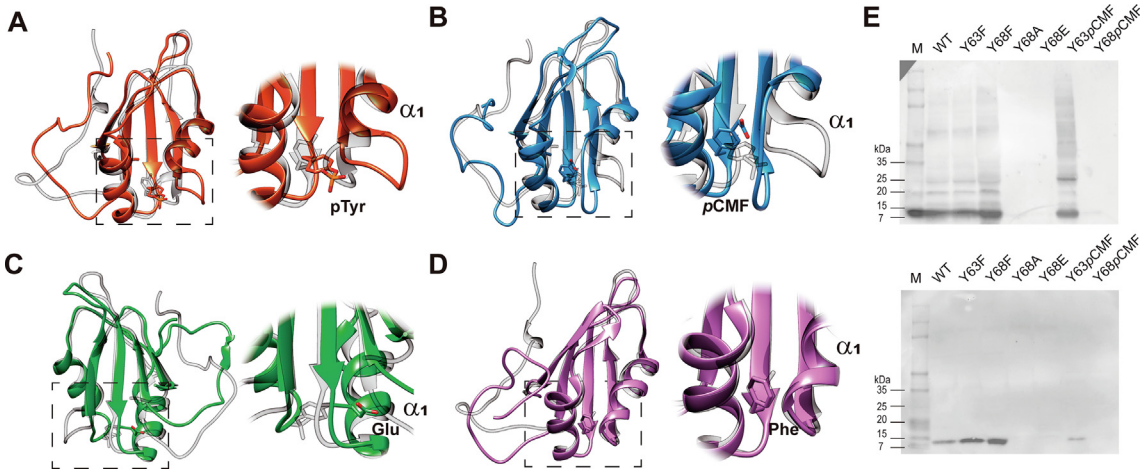

Fig. 4. HuR conformational changes induced by Tyr68 phosphorylation. A, D) Superimposition of ribbon representation of Tyr68 phosphorylated (red) or mutated by Phe (purple) and non-phosphorylated (light gray) HuR_{1-101} structures after 500 ns of MD simulation. B, C) Overlay of the structures mutated in Tyr68 by unnatural amino acid *p*CMF (blue) or by canonical Glu (green) amino acids with respect to the phosphorylated Tyr68 (light gray) HuR_{1-101} after 500 ns of MD simulation. E) Western blots using anti-HuR (*upper* panel) and anti-6xHis-tag (*lower* panel) antibodies from *E. coli* cell extracts previously transformed with plasmids encoding different HuR_{1-101} mutants at positions 63 and 68. Molecular mass markers are also depicted (lane M). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

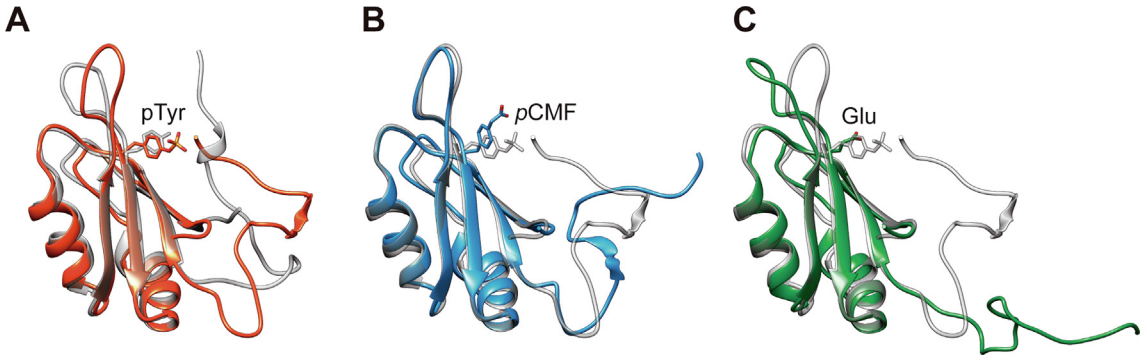

Fig. 5. HuR Tyr63 phosphorylation is mimicked better by *p*CMF than Glu substitution. A) Superimposition of ribbon representation of Tyr63 phosphorylated (red) and non-phosphorylated (light gray) HuR_{1-101} structures after 500 ns of MD simulation. B-C) Overlay of the structures mutated in Tyr63 by the unnatural amino acid *p*CMF (blue) or by canonical amino acid Glu (green) with respect to the phosphorylated Tyr68 (light gray) HuR_{1-101} after 500 ns of MD simulation. (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 HuR_{1-101} 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 HuR_{1-101}. Altogether, these findings suggest a crucial role for the aromatic ring (provided by either Tyr or Phe) at position 68 for correct HuR_{1-101} folding.

A recent NMR analysis of HuR_{1-101} binding to an RNA oligonucleotide derived from adenine and uracil-rich element (ARE) I of the cyclooxygenase 2 mRNA revealed that the RNA binding site comprises Tyr63 among other residues in the β 1 and β 3 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 Y63*p*CMF mutant is recombinantly expressed, we wanted to explore how effectively it mirrors phosphorylation at this position. For Tyr63, similar MD simulations on phosphorylated (pY63) and phosphomimetic variants (Y63E and Y63*p*CMF) of HuR_{1-101} 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 HuR_{1-101} 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 *p*CMF (Fig. 5). Moreover, RMSF values of the area surrounding position 63 are higher than those for WT, pY63 and Y63*p*CMF HuR_{1-101}. Analysis of the Y63E mutant could thus led to an overestimation of the effects of phosphorylation. For this reason, the functional implications of Tyr63 phosphorylation will be assessed using the Y63*p*CMF HuR_{1-101} 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-Mejías: 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-I00), European Regional Development Fund (FEDER), Andalusian Government (BIO-198, US-1254317 and US-1257019) and TA Instruments. G.P.-M. and A.V.-C. were awarded a PhD fellowship from the Spanish Ministry of Education, Culture and Sport (FPU17/04604 and FPU16/01513). 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;55:2392–7.
[3] Hunter T. Tyrosine phosphorylation: thirty years and counting. Curr Opin Cell Biol 2009;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] Fukami 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;85:5849–53.
[10] Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, Zhang B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. Proteomics 2004;4:1551–61.
[11] Puttick J, Baker EN, Delbaere LTJ. Histidine phosphorylation in biological systems. Biochim Biophys Acta, Proteins Proteomics 2008;1784:100–5.
[12] Cieśla J, Frączyk T, Rode W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. Acta Biochim Pol 2011;58:137–48.
[13] Iakoucheva LM, Radivojac P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 2004;32:1037–49.
[14] Mitrea 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, Hollingworth 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, Hollingworth D, Frenkiel TA, 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] Zhu K, Zhao J, Lubman DM, Miller FR, Barder TJ. Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. Anal Chem 2005;77:2745–55.
[20] Halligan BD. ProMoST: a tool for calculating the pI and molecular mass of phosphorylated and modified proteins on two-dimensional gels. Methods Mol Biol 2009;527(283–98):ix.
[21] Bjerrum EJ, Jensen JH, Tolborg JL. pICalculax: improved prediction of isoelectric point for modified peptides. J Chem Inf Model 2017;57:1723–7.
[22] Block H, Maertens B, Spriestersbach A, et al. Immobilized-metal affinity chromatography (IMAC): a review. Methods Enzymol 2009;463:439–73.
[23] Ruprecht B, Koch H, Domasinska P, et al. Optimized enrichment ofpPhosphoproteomes by Fe-IMAC column chromatography. Methods Mol Biol 2017;1550:47–60.
[24] Kinoshita E, Kinoshita-Kikuta E, Takiyama K, Koike T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. Mol Cell Proteomics 2006;5:749–57.
[25] Horinouchi T, Terada K, Higashi T, Miwa S. Using Phos-Tag in Western Blotting analysis to evaluate protein phosphorylation. Methods Mol Biol 2016;1397:267–77.
[26] Kumar G. A simple method for detecting phosphorylation of proteins by using Zn(2+)-Phos-Tag SDS-PAGE at neutral pH. Methods Mol Biol 2018;1853:223–9.
[27] Nagy Z, Comer S, Smolenski A. Analysis of protein phosphorylation using Phos-Tag gels. Curr Protoc Protein Sci 2018;93:e64.
[28] Yang C, Zhong X, Li L. Recent advances in enrichment and separation strategies for mass spectrometry-based phosphoproteomics. Electrophoresis 2014;35:3418–29.
[29] Leitner A. Enrichment strategies in phosphoproteomics. Methods Mol Biol 2016;1355:105–21.
[30] Simon ES, Young M, Chan A, Bao Z-Q, Andrews PC. Improved enrichment strategies for phosphorylated peptides on titanium dioxide using methyl esterification and pH gradient elution. Anal Biochem 2008;377:234–42.
[31] Collins MO, Yu L, Choudhary JS. Analysis of protein phosphorylation on a proteome-scale. Proteomics 2007;7:2751–68.
[32] Paradela A, Albar JP. Advances in the analysis of protein phosphorylation. J Proteome Res 2008;7:1809–18.
[33] Francavilla C, Hekmat O, Blagoev B, Olsen JV. SILAC-based temporal phosphoproteomics. Methods Mol Biol 2014;1188:125–48.
[34] Wang M-C, Lee Y-H, Liao P-C. Optimization of titanium dioxide and immunoaffinity-based enrichment procedures for tyrosine phosphopeptide using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal Bioanal Chem 2015;407:1343–56.
[35] Possemato AP, Paulo JA, Mulhern D, et al. Multiplexed phosphoproteomic profiling using titanium dioxide and immunoaffinity enrichments reveals complementary phosphorylation events. J Proteome Res 2017;16:1506–14.
[36] Buhs S, Gerull H, Nollau P. Identification of tyrosine phosphorylated proteins by SH2 domain affinity purification and mass spectrometry. Methods Mol Biol 2017;1555:407–18.
[37] Ke M, Chu B, Lin L, Tian R. SH2 Domains as affinity reagents for phosphotyrosine protein enrichment and proteomic analysis. Methods Mol Biol 2017;1555:395–406.
[38] Kaneko T, Huang H, Cao X, et al. Superbinder SH2 domains act as antagonists of cell signaling. Sci Signal 2012;5:ra68.
[39] Humphrey SJ, Azimifar SB, Mann M. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. Nat Biotechnol 2015;33:990–5.
[40] Casado P, Hijazi M, Britton D, Cutillas PR. Impact of phosphoproteomics in the translation of kinase-targeted therapies. Proteomics 2017;17.
[41] Yang W, Freeman MR, Kyprianou N. Personalization of prostate cancer therapy through phosphoproteomics. Nat Rev Urol 2018;15:483–97.
[42] Minic Z, Dahms TES, Babu M. Chromatographic separation strategies for precision mass spectrometry to study protein-protein interactions and protein phosphorylation. J Chromatogr, B: Anal Technol Biomed Life Sci 2018;1102–1103:96–108.

[43] Sommerfeld MR, Metzger S, Stosik M, Tennagels N, Eckel J. In vitro phosphorylation of insulin receptor substrate 1 by protein kinase C-zeta: functional analysis and identification of novel phosphorylation sites. Biochemistry 2004;43:5888–901.
[44] Baudin F, Müller CW. Bacterial expression, purification, and crystallization of tyrosine phosphorylated STAT proteins. Methods Mol Biol 2013;967:301–17.
[45] Shin S-B, Woo S-U, Lee Y-J, Yim H. Comparative analysis of a FRET-based PLK1 kinase assay to identify PLK1 inhibitors for chemotherapy. Anticancer Res 2017;37:1177–83.
[46] Duchemin A-M, Neff NH, Hadjiconstantinou M. Aromatic L-amino acid decarboxylase phosphorylation and activation by PKGIalpha in vitro. J Neurochem 2010;114:542–52.
[47] Babon JJ, Murphy JM. In vitro JAK kinase activity and inhibition assays. Methods Mol Biol 2013;967:39–55.
[48] Loeffler DA, Smith LM, Klaver AC, Martić S. Effects of antibodies to phosphorylated and non-phosphorylated tau on in vitro tau phosphorylation at Serine-199: Preliminary report. Exp Gerontol 2015;67:15–8.
[49] Thorsness PE, Koshland DEJ. Inactivation of isocitrate dehydrogenase by phosphorylation is mediated by the negative charge of the phosphate. J Biol Chem 1987;262:10422–5.
[50] Pearlman SM, Serber Z, Ferrell JEJ. A mechanism for the evolution of phosphorylation sites. Cell 2011;147:934–46.
[51] Huang W, Erikson RL. Constitutive activation of Mek1 by mutation of serine phosphorylation sites. Proc Natl Acad Sci U S A 1994;91:8960–3.
[52] Zheng L, Dominski Z, Yang X-C, et al. Phosphorylation of Stem-Loop Binding Protein (SLBP) on Two threonines triggers degradation of SLBP, the sole cell cycle-regulated factor required for regulation of histone mRNA processing, at the end of S phase. Mol Cell Biol 2003;23:1590–601.
[53] Strickfaden SC, Winters MJ, Ben-Ari G, et al. A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. Cell 2007;128:519–31.
[54] Dissmeyer N, Schnittger A. Use of phospho-site substitutions to analyze the biological relevance of phosphorylation events in regulatory networks. Methods Mol Biol 2011;779:93–138.
[55] Stateva SR, Salas V, Benaim G, et al. Characterization of phospho-(tyrosine)-mimetic calmodulin mutants. PLoS ONE 2015;10:e0120798.
[56] Zisch AH, Pazzagli C, Freeman AL, et al. Replacing two conserved tyrosines of the EphB2 receptor with glutamic acid prevents binding of SH2 domains without abrogating kinase activity and biological responses. Oncogene 2000;19:177–87.
[57] Jonson PH, Petersen SB. A critical view on conservative mutations. Protein Eng 2001;14:397–402.
[58] Fang KY, Lieblich SA, Tirrell DA. Incorporation of non-canonical amino acids into proteins by global reassignment of sense codons. Methods Mol Biol 2018;1798:173–86.
[59] Terasaka N, Iwane Y, Geiermann A-S, Goto Y, Suga H. Recent developments of engineered translational machineries for the incorporation of non-canonical amino acids into polypeptides. Int J Mol Sci 2015;16:6513–31.
[60] Young TS, Ahmad I, Yin JA, Schultz PG. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. J Mol Biol 2010;395:361–74.
[61] Chatterjee A, Sun SB, Furman JL, Xiao H, Schultz PG. A versatile platform for single- and multiple-unnatural amino acid mutagenesis in*Escherichia coli*. Biochemistry 2013;52:1828–37.
[62] Leisle L, Valiyaveetil F, Mehl RA, Ahern CA. Incorporation of Non-Canonical Amino Acids. Adv Exp Med Biol 2015;869:119–51.
[63] Schwark DG, Schmitt MA, Fisk JD. Dissecting the contribution of release factor interactions to amber stop codon reassignment efficiencies of the methanocaldococcus jannaschii orthogonal pair. Genes (Basel) 2018;9:546.
[64] Nehring S, Budisa N, Wiltschi B. Performance analysis of orthogonal pairs designed for an expanded eukaryotic genetic code. PLoS ONE 2012;7:e31992.
[65] Pott M, Schmidt MJ, Summerer D. Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. ACS Chem Biol 2014;9:2815–22.
[66] Guo J, Melançon 3rd CE, Lee HS, Groff D, Schultz PG. Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. Angew Chem Int Ed Engl 2009;48:9148–51.
[67] Ohtsuki T, Yamamoto H, Doi Y, Sisido M. Use of EF-Tu mutants for determining and improving aminoacylation efficiency and for purifying aminoacyl tRNAs with non-natural amino acids. J Biochem 2010;148:239–46.
[68] Wang N, Ju T, Niu W, Guo J. Fine-tuning interaction between aminoacyl-tRNA synthetase and tRNA for efficient synthesis of proteins containing unnatural amino acids. ACS Synth Biol 2015;4:207–12.
[69] Kwok HS, Vargas-Rodriguez O, Melnikov SV, Söll D. Engineered aminoacyl-tRNA synthetases with improved selectivity toward noncanonical amino acids. ACS Chem Biol 2019;14:603–12.
[70] Wang K, Neumann H, Peak-Chew SY, Chin JW. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. Nat Biotechnol 2007;25:770–7.
[71] Johnson DBF, Xu J, Shen Z, et al. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol 2011;7:779–86.
[72] Zhu P, Gafken PR, Mehl RA, Cooley RB. A highly versatile expression system for the production of multiply phosphorylated proteins. ACS Chem Biol 2019;14:1564–72.
[73] Rothman DM, Petersson EJ, Vázquez ME, et al. Caged phosphoproteins. J Am Chem Soc 2005;127:846–7.
[74] Steinfeld JB, Aerni HR, Rogulina S, Liu Y, Rinehart J. Expanded cellular amino acid pools containing phosphoserine, phosphothreonine, and phosphotyrosine. ACS Chem Biol 2014;9:1104–12.
[75] Smolskaya S, Andreev YA. Site-specific incorporation of unnatural amino acids into *Escherichia coli* recombinant protein: methodology development and recent achievement. Biomolecules 2019:9.
[76] Ko W, Kumar R, Kim S, Lee HS. Construction of bacterial cells with an active transport system for unnatural amino acids. ACS Synth Biol 2019;8:1195–203.
[77] Luo X, Fu G, Wang RE, et al. Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. Nat Chem Biol 2017;13:845–9.
[78] Venkat S, Chen H, Gan Q, Fan C. The application of cell-free protein synthesis in genetic code expansion for post-translational modifications. Front Pharmacol 2019;10:248.
[79] Jin X, Park O-J, Hong SH. Incorporation of non-standard amino acids into proteins: challenges, recent achievements, and emerging applications. Appl Microbiol Biotechnol 2019;103:2947–58.
[80] Oza JP, Aerni HR, Pirman NL, et al. Robust production of recombinant phosphoproteins using cell-free protein synthesis. Nat Commun 2015;6:8168.
[81] Park HS, Hohn MJ, Umehara T, et al. Expanding the genetic code of Escherichia coli with phosphoserine. Science (80-) 2011;333:1151–4.
[82] 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.
[83] Yanagisawa T, Takahashi M, Mukai T, et al. Multiple site-specific installations of Nε-monomethyl-L- lysine into histone proteins by cell-based and cell-free protein synthesis. ChemBioChem 2014;15:1830–8.
[84] Chemla Y, Ozer E, Shaferman M, et al. Simplified methodology for a modular and genetically expanded protein synthesis in cell-free systems. Synth Syst Biotechnol 2019;4:189–96.
[85] Kawahata N, Yang MG, Luke GP, et al. A novel phosphotyrosine mimetic 4'-carboxymethyloxy-3'-phosphonophenylalanine (Cpp): exploitation in the design of nonpeptide inhibitors of pp60(Src) SH2 domain. Bioorg Med Chem Lett 2001;11:2319–23.
[86] Erlanson DA, McDowell RS, He MM, et al. Discovery of a new phosphotyrosine mimetic for PTP1B using breakaway tethering. J Am Chem Soc 2003;125:5602–3.
[87] Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in *Escherichia coli*. Nat Biotechnol 2006;24:1436–40.
[88] Xie J, Supekova L, Schultz PG. A genetically encoded metabolically stable analogue of phosphotyrosine in *Escherichia coli*. ACS Chem Biol 2007;2:474–8.
[89] Serwa R, Wilkening I, Del Signore G, et al. Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. Angew Chem Int Ed Engl 2009;48:8234–9.
[90] Fan C, Ip K, Söll D. Expanding the genetic code of *Escherichia coli* with phosphotyrosine. FEBS Lett 2016;590:3040–7.
[91] Hoppmann C, Wong A, Yang B, et al. Site-specific incorporation of phosphotyrosine using an expanded genetic code. Nat Chem Biol 2017;13:842–4.
[92] Zhang MS, Brunner SF, Huguenin-Dezot N, et al. Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. Nat Methods 2017;14:729–36.
[93] Hollingsworth SA, Dror RO. Molecular Dynamics simulation for all. Neuron 2018;99:1129–43.
[94] Cordero-Morales JF, Jogini V, Lewis A, et al. Molecular driving forces determining potassium channel slow inactivation. Nat Struct Mol Biol 2007;14:1062–9.
[95] Fields JB, Németh-Cahalan KL, Freites JA, et al. Calmodulin gates Aquaporin 0 permeability through a positively charged cytoplasmic loop. J Biol Chem 2017;292:185–95.
[96] 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.
[97] 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.
[98] Martínez-Fábregas J, Díaz-Moreno I, González-Arzola K, Díaz-Quintana A, De la Rosa MA. A common signalosome for programmed cell death in humans and plants. Cell Death Dis 2014;5:e1314.
[99] González-Arzola K, Díaz-Moreno I, Cano-González A, et al. Structural basis for inhibition of the histone chaperone activity of SET/TAF-Iβ by cytochrome *c*. Proc Natl Acad Sci U S A 2015;112:9908–13.
[100] González-Arzola K, Díaz-Quintana A, Rivero-Rodríguez F, et al. Histone chaperone activity of *Arabidopsis thaliana* NRP1 is blocked by cytochrome *c*. Nucleic Acids Res 2017;45:2150–65.
[101] González-Arzola K, Velázquez-Cruz A, Guerra-Castellano A, et al. New moonlighting functions of mitochondrial cytochrome *c* in the cytoplasm and nucleus. FEBS Lett 2019;593:3101–19.
[102] Díaz-Moreno I, Velázquez-Cruz A, Curran-French S, Díaz-Quintana A, De la Rosa MA. Nuclear cytochrome *c* - a mitochondrial visitor regulating damaged chromatin dynamics. FEBS Lett 2018;592:172–8.
[103] Desagher S, Martinou JC. Mitochondria as the central control point of apoptosis. Trends Cell Biol 2000;10:369–77.

[104] Jiang X, Wang X. Cytochrome C-mediated apoptosis. Annu Rev Biochem 2004;73:87–106.
[105] Elena-Real CA, Díaz-Quintana A, González-Arzola K, et al. Cytochrome *c* speeds up caspase cascade activation by blocking 14-3-3ε-dependent Apaf-1 inhibition. Cell Death Dis 2018;9:365.
[106] Kalpage HA, Bazylianska V, Recanati MA, et al. Tissue-specific regulation of cytochrome *c* by post-translational modifications: respiration, the mitochondrial membrane potential, ROS, and apoptosis. FASEB J 2019;33:1540–53.
[107] Díaz-Moreno I, García-Heredia JM, Díaz-Quintana A, Teixeira M, De La Rosa MA. Nitration of tyrosines 46 and 48 induces the specific degradation of cytochrome *c* upon change of the heme iron state to high-spin. Biochim Biophys Acta - Bioenerg 2011;1807:1616–23.
[108] Ly HK, Utesch T, Díaz-Moreno I, et al. Perturbation of the redox site structure of cytochrome *c* variants upon tyrosine nitration. J Phys Chem B 2012;116:5694–702.
[109] Lee I, Salomon AR, Yu K, et al. New prospects for an old enzyme: mammalian cytochrome *c* is tyrosine-phosphorylated in vivo. Biochemistry 2006;45:9121–8.
[110] Yu H, Lee I, Salomon AR, Yu K, Hüttemann M. Mammalian liver cytochrome *c* is tyrosine-48 phosphorylated in vivo, inhibiting mitochondrial respiration. Biochim Biophys Acta - Bioenerg 2008;1777:1066–71.
[111] Abriata LA, Cassina A, Tórtora V, et al. Nitration of solvent-exposed tyrosine 74 on cytochrome *c* triggers heme iron-methionine 80 bond disruption nuclear magnetic resonance and optical spectroscopy studies. J Biol Chem 2009;284:17–26.
[112] Zhao X, León IR, Bak S, et al. Phosphoproteome analysis of functional mitochondria isolated from resting human muscle reveals extensive phosphorylation of inner membrane protein complexes and enzymes. Mol Cell Proteomics 2011;10(M110):000299.
[113] Wan J, Kalpage HA, Vaishnav A, et al. Regulation of respiration and apoptosis by cytochrome *c* Threonine 58 phosphorylation. Sci Rep 2019;9:15815.
[114] Moreno-Beltrán B, Guerra-Castellano A, Díaz-Quintana A, et al. Structural basis of mitochondrial dysfunction in response to cytochrome *c* phosphorylation at tyrosine 48. Proc Natl Acad Sci U S A 2017;114:E3041–50.
[115] Pieper D, Schirmer S, Prechtel AT, et al. Functional characterization of the HuR:CD83 mRNA interaction. PLoS ONE 2011;6:e23290.
[116] Pabis M, Popowicz GM, Stehle R, et al. HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRMs. Nucleic Acids Res 2019;47:1011–29.
[117] Díaz-Quintana A, García-Mauriño SM, Díaz-Moreno I. Dimerization model of the C-terminal RNA recognition motif of HuR. FEBS Lett 2015;589:1059–66.
[118] Lixa C, Mujo A, de Magalhães MTQ, et al. Oligomeric transition and dynamics of RNA binding by the HuR RRM1 domain in solution. J Biomol NMR 2018;72:179–92.
[119] Grammatikakis I, Abdelmohsen K, Gorospe M. Posttranslational control of HuR function. Wiley Interdiscip Rev RNA 2017;8:e1372.
[120] Yoon J-H, Abdelmohsen K, Srikantan S, et al. Tyrosine phosphorylation of HuR by JAK3 triggers dissociation and degradation of HuR target mRNAs. Nucleic Acids Res 2014;42:1196–208.
[121] Liu BA, Nash PD. Evolution of SH2 domains and phosphotyrosine signalling networks. Philos Trans R Soc B Biol Sci 2012;367:2556–73.
[122] Hornbeck PV, Zhang B, Murray B, et al. PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. Nucleic Acids Res 2015;43:D512–20.
[123] Dinkel H, Chica C, Via A, et al. Phospho.ELM: A database of phosphorylation sites-update 2011. Nucleic Acids Res 2011;39:D261–7.
[124] Linding R, Jensen LJ, Pasculescu A, et al. NetworKIN: A resource for exploring cellular phosphorylation networks. Nucleic Acids Res 2008;36:D695–9.
[125] Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 1999;294:1351–62.
[126] Miller ML, Soufi B, Jers C, et al. NetPhosBac - A predictor for Ser/Thr phosphorylation sites in bacterial proteins. Proteomics 2009;9:116–25.
[127] Ingrell CR, Miller ML, Jensen ON, Blom N. NetPhosYeast: Prediction of protein phosphorylation sites in yeast. Bioinformatics 2007;23:895–7.
[128] Fenoy E, Izarzugaza JMG, Jurtz V, Brunak S, Nielsen M. A generic deep convolutional neural network framework for prediction of receptor-ligand interactions-NetPhosPan: application to kinase phosphorylation prediction. Bioinformatics 2019;35:1098–107.
[129] Xue Y, Ren J, Gao X, et al. GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. Mol Cell Proteomics 2008;7:1598–608.
[130] Obenauer JC, Cantley LC, Yaffe MB. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res 2003;31:3635–41.
[131] Huang HDA, Lee TYI, Tzeng SW, et al. Incorporating hidden Markov models for identifying protein kinase-specific phosphorylation sites. J Comput Chem 2005;26:1032–41.
[132] Saunders NFW, Brinkworth RI, Huber T, Kemp BE, Kobe B. Predikin and PredikinDB: A computational framework for the prediction of protein kinase peptide specificity and an associated database of phosphorylation sites. BMC Bioinf 2008;9:245.
[133] Gnad F, Ren S, Cox J, et al. PHOSIDA (phosphorylation site database): Management, structural and evolutionary investigation, and prediction of phosphosites. Genome Biol 2007;8:R250.
[134] Wong YH, Lee TY, Liang HK, et al. KinasePhos 2.0: A web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. Nucleic Acids Res 2007;35:W588–94.
[135] Neuberger G, Schneider G, Eisenhaber F. pkaPS: Prediction of protein kinase A phosphorylation sites with the simplified kinase-substrate binding model. Biol Direct 2007;2:1.
[136] Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. J Comput Chem 2005;26:1781–802.
[137] Brooks BR, Brooks 3rd CL, Mackerell ADJ, et al. CHARMM: the biomolecular simulation program. J Comput Chem 2009;30:1545–614.
[138] Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1–2:19–25.
[139] Case DA, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham III TE, et al. DMY and PAK. AMBER 2020 n.d.
[140] Eastman P, Swails J, Chodera JD, et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS Comput Biol 2017;13:e1005659.
[141] Maier JA, Martinez C, Kasavajhala K, et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11:3696–713.
[142] Imai M, Saio T, Kumeta H, et al. Investigation of the redox-dependent modulation of structure and dynamics in human cytochrome *c*. Biochem Biophys Res Commun 2016;469:978–84.
[143] Homeyer N, Horn AHC, Lanig H, Sticht H. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. J Mol Model 2006;12:281–9.
[144] Guerra-Castellano A, Díaz-Quintana A, Moreno-Beltrán B, et al. Mimicking tyrosine phosphorylation in human cytochrome *c* by the evolved tRNA synthetase technique. Chem - A Eur J 2015;21:15004–12.
[145] Autenrieth F, Tajkhorshid E, Baudry J, Luthey-Schulten Z. Classical force field parameters for the heme prosthetic group of cytochrome *c*. J Comput Chem 2004;25:1613–22.
[146] Izadi S, Anandakrishnan R, Onufriev AV. Building water models: A different approach. J Phys Chem Lett 2014;5:3863–71.
[147] McGibbon RT, Beauchamp KA, Harrigan MP, et al. MDTraj: A modern open library for the analysis of molecular dynamics trajectories. Biophys J 2015;109:1528–32.
[148] Roe DR, Cheatham 3rd TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 2013;9:3084–95.
[149] Guerra-Castellano A, Díaz-Quintana A, Pérez-Mejías G, et al. Oxidative stress is tightly regulated by cytochrome *c* phosphorylation and respirasome factors in mitochondria. Proc Natl Acad Sci U S A 2018;115:7955–60.
[150] Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46:W296–303.
[151] Webb B, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol 2017;1654:39–54.
[152] Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 2004;32:W526–31.
[153] O'Rourke KF, Gorman SD, Boehr DD. Biophysical and computational methods to analyze amino acid interaction networks in proteins. Comput Struct Biotechnol J 2016;14:245–51.
[154] Pecina P, Borisenko GG, Belikova NA, et al. Phosphomimetic substitution of cytochrome *c* tyrosine 48 decreases respiration and binding to cardiolipin and abolishes ability to trigger downstream caspase activation. Biochemistry 2010;49:6705–14.
[155] Mahapatra G, Varughese A, Ji Q, et al. Phosphorylation of cytochrome *c* threonine 28 regulates electron transport chain activity in kidney: Implications for amp kinase. J Biol Chem 2017;292:64–79.
[156] Guerra-Castellano A, Díaz-Moreno I, Velázquez-Campoy A, et al. Structural and functional characterization of phosphomimetic mutants of cytochrome *c* at threonine 28 and serine 47. Biochim Biophys Acta - Bioenerg 2016;1857:387–95.
[157] Díaz-Quintana A, Pérez-Mejías G, Guerra-Castellano A, De la Rosa MA, Díaz-Moreno I. Wheel and deal in the mitochondrial inner membranes: the tale of cytochrome *c* and cardiolipin. Oxid Med Cell Longev 2020;2020:6813405.
[158] Deacon OM, Karsisiotis AI, Moreno-Chicano T, et al. Heightened dynamics of the oxidized Y48H variant of human cytochrome *c* increases its peroxidatic activity. Biochemistry 2017;56:6111–24.
[159] Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 2016;44:W344–50.
[160] Gu T-L, Cherry J, Tucker M, et al. Identification of activated Tnk1 kinase in Hodgkin's lymphoma. Leukemia 2010;24:861–5.
[161] Rogerson DT, Sachdeva A, Wang K, et al. Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. Nat Chem Biol 2015;11:496–503.
[162] García-Heredia JM, Díaz-Moreno I, Díaz-Quintana A, et al. Specific nitration of tyrosines 46 and 48 makes cytochrome *c* assemble a non-functional apoptosome. FEBS Lett 2012;586:154–8.