**Article:**
Adams, M, Fleming, JR, Riehle, E et al. (4 more authors) (2019) Scalable, Non-denaturing Purification of Phosphoproteins Using $Ga^{3+}$-IMAC: N2A and M1M2 Titin Components as Study case. Protein Journal, 38 (2). pp. 181-189. ISSN 1572-3887

https://doi.org/10.1007/s10930-019-09815-w

# Scalable, non-denaturing purification of phosphoproteins using $Ga^{3+}$-IMAC: N2A and M1M2 titin components as study case

Michael Adams[1], Jennifer R. Fleming[1], Eva Riehle[1], Tiankun Zhou[1], Thomas Zacharchenko[2], Marija Markovic[1], Olga Mayans[1]

[1]Department of Biology, University of Konstanz, 78457 Konstanz, Germany; [2]School of Biology, University of Leeds, Leeds LS2 9JT, UK.

***To whom correspondence should be addressed:**
Prof. Olga Mayans, Tel: +49-7531 882212, olga.mayans@uni-konstanz.de

**Running Title:** $Ga^{3+}$-IMAC purification of phosphoproteins in native conditions

**Abstract**

The purification of phosphorylated proteins in a folded state and in large enough quantity for biochemical or biophysical analysis remains a challenging task. Here, we develop a new implementation of the method of gallium immobilized metal chromatography ($Ga^{3+}$-IMAC) as to permit the selective enrichment of phosphoproteins in the milligram scale and under native conditions using automated FPLC instrumentation. We apply this method to the purification of the UN2A and M1M2 components of the muscle protein titin upon being monophosphorylated in vitro by cAMP-dependent protein kinase (PKA). We found that UN2A is phosphorylated by PKA at its C-terminus in residue S9578 and M1M2 is phosphorylated in its interdomain linker sequence at position T32607. We demonstrate that the $Ga^{3+}$-IMAC method is efficient, economical and suitable for implementation in automated purification pipelines for recombinant proteins. The procedure can be applied both to the selective enrichment and to the removal of phosphoproteins from biochemical samples.

**Keywords:** Phosphorylation / FPLC Protein purification / titin / PKA

Abbreviations: MS, mass spectrometry; IMAC, immobilized metal affinity chromatography; MOAC, metal oxide affinity chromatography; FPLC, fast protein liquid chromatography; IDA, iminodiacetic acid; IPTG, isopropyl β-D-1-thiogalactopyranoside; TCEP, Tris(2-carboxyethyl)phosphine hydrochloride; PKAcα, cAMP-dependent protein kinase.

## 1. Introduction

Phosphorylation of substrate proteins by protein kinases is a fundamental mechanism of cellular regulation [1]. Most phosphorylation studies aim to identify modification sites in target proteins using high-resolution mass spectrometry (MS) and proteomics techniques (reviewed in [2]). In phosphoproteome profiling, phosphoproteins are part of complex mixtures (commonly cell crudes or enriched broths), where incomplete phosphorylation adds to chemical heterogeneity and low protein abundance to trouble site identification. Thus, phosphopeptide enrichment is key to proteomics and considerable efforts have been directed to developing fractionation methodologies for this purpose. Immobilized metal affinity chromatography (IMAC) is the most cost-effective and successfully applied procedure to this effect [3]. However, IMAC leads to a differential recovery of peptides, having an enrichment bias towards multiply phosphorylated species. To increase the sensitivity for monophosphorylated samples and, thereby, improve phosphopeptide coverage in proteomics, novel metal oxide affinity chromatography (MOAC) matrices have been developed (e.g. [4–6]), the tendency being towards the combined usage of IMAC and MOAC materials [5,7]. IMAC systems have been reported that are based on $Ga^{3+}$, $Fe^{3+}$, $Ti^{4+}$, $Zn^{2+}$ and $Al^{3+}$ ions (e.g. [8–13]) and include commercial products such as $Fe^{3+}$-based PHOS-Select™ IMAC (Sigma) and, the now discontinued, $Ga^{+3}$-IMAC PhosphoProfile™ (Sigma). Traditionally, $Fe^{3+}$ has been the preferred immobilized ion for phosphor-group capture, but $Ga^{3+}$ has been shown to be superior in its binding specificity of phosphorylated species [8,9,13,14]. Beyond reaching a certain threshold of purity, phosphoproteome profiling does not place requirements on high yield or the 3D-fold preservation of protein samples. In fact, IMAC protocols (as well as MOAC) are commonly performed in microscale column formats and on proteolized peptide mixtures using strongly denaturing chemical elutants (e.g. acetonitrile, dihydroxybenzoic acid, ammonium hydroxide, trifluoracetic acid, or phosphoric acid at highly acidic pH values; [8,10–12,15–18]) as this yields the best performance.

In contrast to proteome research, the fractionation of intact, non-denatured phosphoproteins has seen surprisingly little development. Such fractionation is key to biochemical, biophysical and structural research that aims to study the conformational and functional consequences of phosphorylation in the target proteins in analytical terms. These studies require natively folded proteins, in high state of purity and in high quantity (often mg scale). To achieve this goal, proteins are commonly produced recombinantly, isolated, and

subjected to phosphorylation in vitro using a known kinase. Yet, phosphotransfer reactions are often incomplete and capturing the phosphorylated species in reaction mixtures remains challenging. Enrichment protocols as employed in proteomics are here unsuitable and often cost-prohibitive. Thus, the production of phosphomimics via site-directed mutagenesis or via the introduction of non-natural phosphor-amino acids is pursued instead to achieve homogeneous sample preparations [19]. These procedures, however, suffer from their own limitations, with phosphomimics often exhibiting behaviours different to those of natively phosphorylated proteins [2, 20].

For a phosphoprotein purification technique to be applicable to downstream applications in the fields of biochemistry and biophysics, it must be highly selective, scalable up to the milligram range, non-destructive, and affordable. The most promising method to this effect is IMAC as these resins are already introduced in biochemical laboratories, are available commercially in large columns formats, compatible with automated Fast Protein Liquid Chromatography (FPLC) instrumentation, rechargeable and reusable. Yet, only few studies have investigated the efficacy of IMAC in native phosphoprotein enrichment. Machida and co-workers [9] described the microscale purification of phosphoproteins from cell lysates using $Ga^{3+}$-IMAC on iminodiacetic acid (IDA) resin and using phosphate buffer for elution, indicating that $Ga^{3+}$-IMAC has promise for intact phosphoprotein purification. More recently, $Ga^{3+}$-IMAC on sepharose beads has been attempted in the purification under native conditions of the protein OPN [20], but selectivity for the phosphorylated protein was not achieved in that study, likely due to the unusually acidic character of that protein.

Here, we establish a procedure for the enrichment of natively folded phosphoproteins using $Ga^{3+}$-IMAC resin, pre-packed columns and automated FPLC instrumentation. We also develop a protocol for the regeneration of the columns to allow reuse. We apply this method to two recombinant proteins as study cases -specifically the UN2A and M1M2 domain components of the myoprotein titin- and demonstrate that the procedure is effective on monophosphorylated species. Our results show the suitability of the method for the high-yield production of intact phosphoproteins for the study of protein structure and function.

## 2. Materials and Methods

### 2.1 Protein production

The expression clone of UN2A from titin (residues 9472-9581; UniProtKB Q8WZ42) has been previously described [21]. The sequence coding for domains M1M2 (residues 32490-32713; UniProtKB Q8WZ42) was cloned into the pOPINB vector (Oxford Protein Production Facility, UK), which adds a $His_6$-tag and a protease C3 cleavage site N-terminally to the target sequence.

Recombinant overproduction of UN2A and M1M2 titin samples used E. coli Rosetta (DE3) cells (Merck Millipore). Cells were cultivated at 37ºC in Terrific Broth supplemented with 25μg/ml kanamycin and 34μg/ml chloramphenicol. Protein expression was induced at $OD_{600}$=0.9 with 0.5mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cultures further grown overnight at 18ºC. Cells were harvested by centrifugation and lysed by sonication in ice in 25mM HEPES pH 7.5, 300mM NaCl, 1mM dithiothreitol, in the presence of an EDTA-free protease inhibitor cocktail (Roche Applied Science). Lysates were clarified by centrifugation and proteins purified from the supernatant using $Ni^{2+}$-NTA affinity chromatography on a HisTrap™ HP column (GE Healthcare), followed by protease mediated tag removal (Tobacco Etch Virus for N2A and C3 protease for M1M2), subtractive affinity chromatography and size-exclusion chromatography (HiLoad Superdex 75 column, GE Healthcare). Buffers used for N2A purification were as described [21], with the final sample buffer being 25mM HEPES pH 7.5, 100mM NaCl, 1mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). M1M2 purification used 300mM NaCl, 50mM Tris-HCl pH 7.5, except for the final size-exclusion chromatography that used 50mM NaCl, 50mM Tris-HCl pH 7.5. Purified samples were stored at 4°C until further use.

### 2.2 Production of PKA kinase

Murine PKA catalytic subunit α (PKAcα) was cloned into the pET-30 Ek/LIC vector (Novagen), which yielded an N-terminally $His_6$-tagged PKA sample. Expression was in E. coli Rosetta (DE3) cells (Merck Millipore), which were grown and harvested as described above for titin samples. Cells were lysed in 25mM HEPES pH 7.5, 30mM NaCl and PKAcα purified

from the supernatant by metal affinity chromatography in a HisTrap™ HP column (GE Healthcare). Binding to the HisTrap™ HP column was in lysis buffer, but elution used 25mM HEPES pH 7.5, 300mM NaCl, 100mM Imidazole. Eluted fractions were dialysed against 25mM HEPES pH 7.5, 100mM NaCl and subjected to size exclusion chromatography (Superdex S200 16/600 column, GE Healthcare). Finally, 50% [v/v] glycerol was added to the purified sample, which was then aliquoted at a concentration of 1mg/mL (estimated by $A_{280}$) and flash frozen in $LN_2$ in gel filtration buffer. The protein was stored at -80°C until use.

### 2.3 Bulk scale phosphorylation assay

Purified UN2A and M1M2 proteins were assayed for phosphorylation by PKAcα. Phosphorylation assays used ca. 30mg of sample protein (in a 15mL volume corresponding to a protein concentration of 2mg/mL) in their final buffers. To this, 1.5mL 10x magnesium buffer (100mM $MgCl_2$, 500mM Tris pH 7.5, 500mM NaCl) was added. Reactions contained 6.6 μg/mL PKAcα and were initiated by addition of ATP (at a final concentration of 1.33 mM and prepared in 50mM HEPES pH 7.5). The mixture was incubated at 37°C and 150rpm for 1h, then gently mixed overnight at room temperature. Phosphorylation was confirmed via Phos-Tag™ (Wako Pure Chemical Industries; [22]) applied to 12% SDS-PAGE.

The SDS-PAGE gels were imaged using a Gel Doc™ EZ Gel imager and the supplied software Image Lab™ 5.2.1 (Bio-Rad). Gel densitometry was performed using Image Lab™ 5.2.1 by automatically detecting the gel bands at the default detection threshold. Relative abundance was then gathered from the lane profile generated by the software.

### 2.4 Gallium-based phosphoprotein purification

Reactions mixtures from phosphorylation assays were subjected to $Ga^{3+}$-IMAC purification. For this, a column was prepared by stripping a commercial 5mL HisTrap™ HP column (GE Healthcare) from $Ni^{2+}$ ions using 20mM sodium phosphate pH 7.4, 0.5M NaCl, 50mM ethylenediaminetetraacetic acid (EDTA). The column was then charged with $Ga^{3+}$ ions by flowing 3 column volumes (CV) of 200mM gallium chloride (Sigma Aldrich). The column was washed in water to remove excess ions and stored in 20% ethanol until use.

For purification, protein mixtures from phosphorylation assays were loaded onto the column at a flow rate of 1 mL/min using an Äkta Start FPLC system (GE Healthcare). To minimize unspecific binding, the column containing the bound samples was washed with a high salt buffer (UN2A: 50 mM MES pH 5.5, 500 mM NaCl; M1M2: 50 mM Tris-HCl pH 6.0, 500 mM NaCl). Upon column washing (5CV), bound protein fractions were eluted using a gradient (5%-to-100%) of 500mM sodium phosphate pH 6.0 (total 5CV) at a flow rate of 5 mL/min. Individual elution fractions were evaluated using Phos-Tag$^{TM}$ (Wako Pure Chemical Industries [22]) in 12% SDS-PAGE.

In all cases, upon each usage, the $Ga^{3+}$-IMAC column was cleaned with 50mM citric acid pH 4.0 that eluted bound phosphate groups. The column was then used in subsequent purification rounds. After ca. 10 usage cycles, the column was stripped of metal ions using EDTA and newly charged with $Ga^{3+}$ ions (as described above) for continued use.

**2.5 Mass spectrometry**

UN2A and M1M2 samples were analyzed by reversed phase liquid chromatography (C18 resin; Acclaim PepMap100, Thermo Scientific) nanospray tandem mass spectrometry (LC-MS/MS) using an LTQ-Orbitrap mass spectrometer (Thermo Fisher) and an Eksigent nano-HPLC. After sample injection, the column was washed for 5 min with 95% phase A (0.1% formic acid) and 5% phase B (0.1% formic acid in acetonitrile), and peptides were eluted using a linear gradient of 5%-to-40% phase B, then 80% phase B at 300 nL/min. The LTQ-Orbitrap mass spectrometer was operated in a data dependent mode in which each full MS scan (30 000 resolving power) was followed by five MS/MS scans where the five most abundant molecular ions were dynamically selected and fragmented by collision-induced dissociation (CID) using a normalized collision energy of 35% in the LTQ ion trap. Dynamic exclusion was allowed. Tandem mass spectra were searched against the Swissprot human protein database using Mascot (Matrix Science) with “Trypsin/P” enzyme cleavage, static cysteine alkylation by iodoacetamide and variable methionine oxidation. Mass deconvolution was performed using mMass 5.5.0. The spectra were smoothed using two cycles of the Savitzky-Golay method and a window size of 0.3 Da/z. Peaks were picked with a signal/noise threshold of 3.0, a picking height of 75 and no deisotoping. Charges for the peaks were assigned manually. Monoisotopic deconvolution was performed with a peak grouping window of 0.01Da/z.

### 2.6 Circular Dichroism

Circular dichroism (CD) spectra were recorded on a J-815 CD spectrometer (JASCO) using a quartz cuvette (Hellma) with a light path of 1mm. Samples were studied in 50 mM $NaH_2PO_4$ pH 7.5 at a concentration of 0.44 mg/mL for M1M2; 1.06 mg/mL for pM1M2; 0.16 mg/mL for UN2A; and 0.26 mg/mL for pUN2A. The spectral range was $\lambda$= 190–300 nm. Spectra were acquired five times at 18°C and averaged. The CDSSTR program [23] in the Dichroweb server [24] was used to calculate secondary structure content from spectra. The fit between experimental and reconstructed spectra was evaluated by their normalized root-mean-square deviation (NRMSD) (NRMSD>0.05 signifies a poor fit).

## 3. Results

### 3.1 UN2A and M1M2 components of titin are phosphorylation targets of PKA

Titin is an intrasarcomeric, multi-domain protein of giant dimensions (>3MDa) that orchestrates the response of muscle to mechanical and metabolic stress [25]. Titin is phosphorylated by a range of kinases that modulate, among other functions, its elastic properties. As study cases for the $Ga^{3+}$-IMAC methodology in this work, we choose to test the phosphorylation of the UN2A and M1M2 components of titin by cAMP-dependent protein kinase (PKA). We selected PKA among those kinases known to be relevant to titin physiology [25] as predictions (NetPhos 3.1 Server [26]) suggested that it might phosphorylate both samples.

UN2A is a 100-residue unique sequence (12 kDa), part of the essential N2A element of titin in the sarcomeric I-band [27] (**Fig 1a**). UN2A is the main receptor site of the anti-apoptotic MARP/CARP proteins in titin [21,28]. The recent biophysical characterization of a recombinant UN2A sample revealed that this domain adopts a loose, helical fold with high internal flexibility [21,29]. Phosphorylation assays of recombinant UN2A (**Fig 1b**) in our current study show that UN2A is phosphorylated by PKAcα. UN2A samples incubated with PKAcα were subjected to Phos-Tag™ SDS-PAGE [22]. The Phos-Tag™ gel contains a

phosphate-binding acrylamide tag that causes phosphorylated proteins to migrate more slowly, resulting in additional SDS-PAGE bands that correspond to the level of phosphorylation of the migrating species. The Phos-Tag™ gels of PKAcα-incubated UN2A exhibited one extra band indicative of the presence of one phosphorylation site (**Fig 1d**). This result was confirmed by full-mass determination mass spectrometry (MS) (**Fig 1d**). Subsequently, the phosphorylation site was identified by tryptic digest MS as being located in residue S9578, which lies close to the C-terminus of UN2A in the linker sequence to the subsequent domain, Ig81. The UN2A-Ig81 domain interface has been recently shown to form the complete binding site of CARP in titin [21]. Future studies will be required to establish whether UN2A phosphorylation is physiologically relevant and whether it affects the titin/CARP interaction.

M1M2 (25,7 kDa) is a dual-Ig component of titin located immediately C-terminal to the titin kinase domain, in the sarcomeric M-line **(Fig 1a)**. The kinase region of titin is a multi-domain scaffold that recruits the turn-over factors MuRF1, MuRF2, Nbr1 and p62 to titin [25]. It is unknown whether the immediately adjacent M1M2 domains might participate in the scaffolding functions of this region. M1 and M2 are separated by a 32-residue long linker that is unstructured according to our sequence-based predictions: predictions of secondary structure used the Jpred4 server [30] and disorder was analyzed using the meta-server MetaDisorder [31]. Predictions were further confirmed with I-TASSER [32]. The length of the interdomain linker in M1M2 (and in titin's M-band in general) contrasts notably with linker lengths in the rest of titin, where domains are connected by short linkers 0-4 residues in length [33]. The functionality of the extended linkers in M-line titin is unknown. In this work, we expressed recombinantly M1M2 as a soluble protein product. The protein, however, had a complex polydisperse chromatographic and electrophoretic behaviour, with SDS-resistant oligomers being observed in gels (**Fig 1c**). We found that M1M2 is a target of PKAcα, with Phos-Tag™ SDS-PAGE analysis and full-length MS revealing the presence of one phosphorylation site in this segment (**Fig 1e**). Site identification using tryptic digest MS revealed the targeted residue to be T32607 (**Fig 1e**), which is located in the linker sequence joining M1 and M2. Remarkably, in twitchin from Mytillus galloprovincialis (the bivalve homologue of titin), the equivalent M1M2 linker sequence binds to myosin, with the binding being released upon phosphorylation of the linker by PKA [34,35]. This PKA-dependent myosin/twitchin interaction is responsible for tension maintenance in the "catch" state of muscle in that animal group - a state where muscular tension is maintained for long periods of time with little energy consumption thanks to physical linkages between actin, myosin and twitchin. The M1M2 linker of vertebrate titin and bivalve twitchin

do not share sequence conservation (**Fig 2**) and vertebrate striated muscle does not adopt a "catch" state. However, since titin is an important scaffold in the sarcomere, it is tantalizing to envision that PKA phosphorylation at this locus might affect titin interactions with M-line associated proteins.

As negative controls in this study, the multi-domain titin kinase (A170-kinase-M1; **Fig 1a**) and the ankyrin-repeat domain of the protein CARP (binding partner of UN2A-Ig81) were also assayed for PKAcα phosphorylation. Neither protein was found to be a substrate of PKAcα (despite weak positive predictions; NetPhos 3.1 Server [26]), pointing to a possible specificity of the sites identified in UN2A and M1M2.

### 3.2 Purification of monophosphorylated UN2A and M1M2 samples

Reaction mixtures from phosphorylation assays contained both phosphorylated and unphosphorylated species due to incomplete phosphotransfer. Based on Phos-Tag™ SDS-PAGE densitometry we estimated that UN2A and M1M2 reactions mixtures contained approx. 56% pUN2A and 65% pM1M2 phosphorylated species, respectively (**Fig 3d,f**). To enrich the phosphorylated fractions, we subjected the reaction mixtures to $Ga^{3+}$-IMAC. For this, a $Ga^{3+}$-IMAC column that was compatible with automated FPLC equipment was produced by repurposing a pre-packed, commercial, metal affinity 5mL HisTrap™ HP column (GE Healthcare) (as described in Methods). $Ga^{3+}$ affinity chromatography was then performed on UN2A and M1M2 reaction mixtures using an Äkta Start FPLC system (GE Healthcare), with proteins being bound during in-column flow with no previous incubation or prolonged exposure to the resin. The large column format and its compatibility with FPLC instrumentation allowed here up-scaling the purification procedure, so that approx. 30 mg total protein (~15 mL total volume) was applied to the IMAC column for each sample. In both cases, bound fractions were eluted with sodium phosphate buffer at near physiological pH (pH 6.0). Upon usage, the column was cleaned from bound phosphate groups by washing with citrate buffer pH 4.0 that protonated and eluted bound phosphate groups; the column was then ready for further reuse.

The purification of pUN2A and pM1M2 species using this approach yielded a distinct elution peak in $Ga^{3+}$-IMAC chromatograms (**Fig 3a,b**). Analysis of elution peaks using Phos-Tag™ SDS-PAGE densitometry confirmed enrichment of phosphorylated species, with

pUN2A amounting now to 92% of the purified sample and pM1M2 being 93.5% (**Fig 3c-f**). Eluted fractions contained approx. 4-5 mg of total phosphoprotein in each case, suggesting that under the conditions tested the phophoprotein binding efficiency was ca. 0.8 mg protein/mL $Ga^{3+}$-IMAC resin.

During this work and under the conditions described, we experienced that the $Ga^{3+}$-IMAC column maintained its maximal binding efficiency for approximately 5 cycles of reuse (binding/elution/cleaning). The column remained usable for up to 10-12 cycles, at which point it had to be regenerated by stripping the gallium ions with EDTA and newly reloading it with $Ga^{3+}$ ions (as described in Methods). Although the number of reusing cycles was consistent for UN2A and M1M2 samples in this study, it should be considered that column reusability will vary depending on buffer conditions and the specific protein assayed.

### 3.3 Structural integrity of $Ga^{3+}$-IMAC purified pUN2A and pM1M2 samples

To confirm the non-destructive nature of $Ga^{3+}$-IMAC, pre- and post- purification samples were analyzed using circular dichroism (CD) (**Fig 3g,h**). CD spectra from unmodified UN2A confirmed that the secondary structure of this domain is primarily α-helical (estimated as ~70%), closely reproducing the spectral properties previously reported [21]. The CD spectrum of the phosphorylated sample, pUN2A, was in excellent agreement with that of the untreated sample, indicating that neither the introduced phosphor-group nor the purification procedure caused significant alterations to the sample. The CD spectrum of M1M2 was noisier and of lower quality, but it confirmed the expectation that this sample has an all-β secondary structure composition, with 42% β-strand content. This content is in good agreement with secondary structure estimates based on the crystal structure of M1 (PDB 2BK8) and other known Ig domains from titin [33], whose β-strand content commonly amounts to 45-50%. The β-strand content estimated from CD data supports the view that the linker sequence is unstructured. The spectrum of M1M2 was approximately reproduced by the pM1M2 sample (estimated 42% β-sheet content), suggesting that also in this case there was no significant structural difference between the treated and untreated samples.

In brief, CD spectra confirmed the gentle nature of $Ga^{3+}$-IMAC purification and its suitability for preparing natively folded phosphoproteins for downstream analysis. As UN2A is

largely α-helical and M1M2 is β-sheet rich, we concluded that protein fold had no influence in the efficacy of the method.

## 4. Discussion and Conclusion

We demonstrate that the utilization of a stripped and $Ga^{3+}$-charged commercially pre-packed, large IMAC column allows the larger-scale purification of phosphoproteins in native conditions. In contrast to other available methods for phosphoprotein enrichment, the elution conditions are gentle and do not perturb the structural fold of proteins. We also show that the method affords scalability and automation. By developing a column regeneration procedure, we also improve affordability. Contrary to conventional studies, where proteins are incubated with the $Ga^{3+}$-IMAC resin in batch for a prolonged time and then run in microscale columns in gravity flow, we use in-column flow binding and automated FPLC equipment and show that the method is efficient under these conditions. Furthermore, we demonstrate that the method is sufficiently sensitive as to capture and resolve monophosphorylated species. We tested the method in two samples of contrasting characteristics. UN2A is an α-helical, monomeric and monodisperse protein. M1M2 is a dual-domain, β-rich protein that contains a long unstructured linker and that forms oligomers in solution of difficult chromatographic resolution and resistant to SDS, which leads to a complex electrophoretic migration pattern. The method successfully segregates phosphorylated and non-phosphorylated formations in both samples. The procedure is a valuable tool for research into the structure and function of phosphoproteins. In combination with ablation of undesirable phosphorylation sites using mutagenesis, it can provide singly phosphorylated species for functional and structural characterization in good yield and purity. Conversely, it can be applied to the clearing of protein samples from recombinant expression in eukaryotic systems, which often are affected by unwanted phosphorylation by kinases present in the host cell system. This could serve as an alternative methodology in cases where the addition of phosphatases is not desirable. Due to its versatility, the method shows promise for a wide variety of applications, as it offers the specificity of other protocols commonly used for phosphoprotein purification, whilst avoiding the drawbacks of denaturing conditions, low yield, or high cost.

**Acknowledgements**

We acknowledge the financial support of DFG SFB969 and the Leducq Foundation (TNE-13CVD04). JRF is supported by an EU Marie Sklodowska-Curie Individual Fellowship (TTNPred, 753054). We thank the Proteomics Unit of the University of Konstanz for the contribution of mass spectrometry to this work.

**Author contributions**

MA, JRF and OM conceived the study; MA and TZa performed experiments and analysed data for M1M2; MA, ER and TZh performed experiments and analysed data for UN2A; MM designed protocols for $Ga^{3+}$-IMAC column reuse and regeneration; OM, MA and JRF wrote the manuscript; all authors made manuscript revisions.

**Figure legends**

**Fig 1:** Phosphorylation of UN2A and M1M2 by PKA

**a.** Schematic composition of the N2A element and the M-line multi-domain kinase region of titin. Shadowed are domains under study in this work; **b.** and **c.** gel filtration chromatogram and associated Coomassie-stained SDS-PAGE of purified samples. The molecular mass of the N2A construct used in this study is 13.4 kDa and that of M1M2 is 25.6 kDa. M1M2 migrates in SDS-PAGE with a lower apparent molecular mass. This is consistent with the observation of the dimeric form of the sample in the gels and with our conclusion that the sample is partly resistant to denaturation by SDS. All bands in the gels of UN2A and M1M2 were confirmed to correspond to the proteins under study using MS protein identification analysis; **d.** and **e.** Phos-Tag SDS-PAGE, full-mass MS analysis and sequence segments showing the phosphorylated residue for each sample. Throughout, UN2A results are shown on the left and M1M2 data on the right.

**Fig 2:** Sequence alignment of M1M2 from vertebrate titin and invertebrate homologs

Sequence alignment of the M1M2 linker sequence from vertebrate titin with its corresponding sequences in invertebrate homologs. The phosphorylation site in M1M2 found in this study as well as the residue found previously to be phosphorylated in Mytillus twitchin [34,35] are indicated.

**Fig 3:** Purification and characterization of phosphorylated forms of UN2A and M1M2
**a.** and **b.** $Ga^{3+}$-IMAC chromatograms showing flow-through, wash and elution fractions. Elution used a sodium phosphate gradient; **c.** and **e.** Phos-Tag™ SDS-PAGE corresponding to chromatograms in a. and b. (FT: flow-through; E: Elution); **d.** and **f.** Histograms of densitometry measurements for gels shown in c. and e.; **g.** and **h.** CD spectra of non-phosphorylated samples (solid line) and phosphorylated species purified by $Ga^{3+}$-IMAC (dashed lines). Secondary structure and disorder content was estimated from spectra as follows: UN2A (helix:70% / disorder: 21%; NRMSD: 0.003); pUN2A (helix: 66% / disorder: 22%; NRMSD: 0.002); M1M2 (strand:42% / disorder: 56%, NRMSD: 0.153); pM1M2 (strand 42% / disorder: 55%; NRMSD: 0.164).

## References

[1] Humphrey SJ, James DE, Mann M (2015) Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. Trends Endocrinol Metab 26:676–687

[2] Dephoure N, Gould KL, Gygi SP, Kellogg DR (2013) Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. Mol Biol Cell 24:535–542

[3] Fíla J, Honys D (2012) Enrichment techniques employed in phosphoproteomics. Amino Acids 43:1025–1047

[4] Li Y, Xu X, Qi D, Deng C, Yang P, Zhang X (2008) Novel Fe3O4@TiO2 core-shell microspheres for selective enrichment of phosphopeptides in phosphoproteome analysis. J Proteome Res 7:2526–2538

[5] Lin H, Deng C (2016) Development of immobilized Sn(4+) affinity chromatography material for highly selective enrichment of phosphopeptides. Proteomics. 16:2733–2741

[6] Zou X, Jie J, Yang B. (2017) Single-Step Enrichment of N-Glycopeptides and Phosphopeptides with Novel Multifunctional Ti(4+)-Immobilized Dendritic Polyglycerol Coated Chitosan Nanomaterials. Anal Chem 89:7520–7526

[7] Yang DS, Ding XY, Min HP, Li B, Su MX, Niu MM, Di B, Yan F (2017) Design and synthesis of an immobilized metal affinity chromatography and metal oxide affinity chromatography hybrid material for improved phosphopeptide enrichment. J Chromatogr A 1505:56–62

[8] Posewitz MC, Tempst P (1999) Immobilized gallium(III) affinity chromatography of phosphopeptides. Anal Chem 71:2883–2892

[9] Machida M, Kosako H, Shirakabe K, Kobayashi M, Ushiyama M, Inagawa J, Hirano J, Nakano T, Bando Y, Nishida E, Hattori S (2007) Purification of phosphoproteins by immobilized metal affinity chromatography and its application to phosphoproteome analysis. FEBS J 274:1576–1587

[10] Lai AC, Tsai CF, Hsu CC, Sun YN, Chen YJ (2012) Complementary Fe(3+)- and Ti(4+)-immobilized metal ion affinity chromatography for purification of acidic and basic phosphopeptides. Rapid Commun Mass Spectrom 26:2186–2194

[11] Zhu L, Zhang J, Guo Y (2014) Enhanced detection and desalting free protocol for phosphopeptides eluted from immobilized Fe (III) affinity chromatography in direct MALDI TOF analysis. J Proteomics 96:360–365

[12] Yao Y, Dong J, Dong M, Liu F, Wang Y, Mao J, Ye M, Zou H (2017) An immobilized titanium (IV) ion affinity chromatography adsorbent for solid phase extraction of phosphopeptides for phosphoproteome analysis. J Chromatogr A. 1498:22–28

[13] Aryal UK, Ross AR (2010) Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. Rapid Commun Mass Spectrom 24:219–231

[14] Sykora C, Hoffmann R, Hoffmann P (2007) Enrichment of multiphosphorylated peptides by immobilized metal affinity chromatography using Ga(III)- and Fe(III)-complexes. Protein Pept Lett 14:489–496

[15] Aryal UK, Olson DJ, Ross AR (2008) Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests. J Biomol Tech 19:296–310

[16] Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou H. (2008) A multidimensional chromatography technology for in-depth phosphoproteome analysis. Mol Cell Proteomics. 7:1389–1396

[17] Yue XS, Hummon AB (2013) Combination of multistep IMAC enrichment with high-pH reverse phase separation for in-depth phosphoproteomic profiling. J Proteome Res 12:4176–4186

[18] Steen H, Stensballe A, Jensen ON (2007) Phosphopeptide Purification by IMAC with Fe(III) and Ga(III). CSH Protoc. https://doi: 10.1101/pdb.prot4607

[19] Liu WR, Wang YS, Wan W (2011) Synthesis of proteins with defined posttranslational modifications using the genetic noncanonical amino acid incorporation approach. Mol Biosyst 7:38–47

[20] Ravi A, Guo S, Rasala B, Tran M, Mayfield S, Nikolov ZL (2018) Separation Options for Phosphorylated Osteopontin from Transgenic Microalgae Chlamydomonas reinhardtii. Int J Mol Sci. https://doi: 10.3390/ijms19020585

[21] Zhou T, Fleming JR, Franke B, Bogomolovas J, Barsukov I, Rigden DJ, Labeit S, Mayans O. (2016) CARP interacts with titin at a unique helical N2A sequence and at the domain Ig81 to form a structured complex. FEBS Lett. 590:3098–3110

[22] Kinoshita E, Yamada A, Takeda H, Kinoshita-Kikuta E, Koike T. (2005) Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins. J Sep Sci. 28:155–162

[23] Johnson WC (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins. 35:307–312

[24] Whitmore L, Wallace BA. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers. 89:392-400

[25] M. Krüger, S. Kötter, Titin, a central mediator for hypertrophic signaling, exercise-induced mechanosignaling and skeletal muscle remodeling, Front. Physiol. 7 (2016) 1–8. doi:10.3389/fphys.2016.00076

[26] Blom N, Gammeltoft S, and Brunak S (1999) Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of Molecular Biology: 294:1351-1362

[27] Bang ML1, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H, Labeit S (2001) The complete gene sequence of titin, expression of an unusual ≈700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. Circ Res 89:1065–1072

[28] Miller MK, Bang ML, Witt CC, Labeit D, Trombitas C, Watanabe K, Granzier H, McElhinny AS, Gregorio CC, Labeit S (2003) The muscle ankyrin repeat proteins: CARP, ankrd2/Arpp and DARP as a family of titin filament-based stress response molecules. J Mol Biol. 333:951–964

[29] Tiffany H, Sonkar K, Gage MJ (2017) The insertion sequence of the N2A region of titin exists in an extended structure with helical characteristics. Biochim Biophys Acta Proteins Proteom. 1865:1–10

[30] Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: A protein secondary structure prediction server. Nucleic Acids Res 43:W389–W394

[31] Kozlowski LP, Bujnicki JM (2012) MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. BMC Bioinformatics. 13:111

[32] Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 9:1–8

[33] Zacharchenko T, von Castelmur E, Rigden DJ, Mayans O. (2015) Structural advances on titin: towards an atomic understanding of multi-domain functions in myofilament mechanics and scaffolding Biochem Soc Trans 43:850–855

[34] Funabara D, Kinoshita S, Watabe S, Siegman MJ, Butler TM, Hartshorne DJ (2001) Phosphorylation of molluscan twitchin by the cAMP-dependent protein kinase. Biochemistry. 40:2087–2095

[35] Funabara D, Hamamoto C, Yamamoto K, Inoue A, Ueda M, Osawa R, Kanoh S, Hartshorne DJ, Suzuki S, Watabe S. (2007) Unphosphorylated twitchin forms a complex with actin and myosin that may contribute to tension maintenance in catch. J Exp Biol. 210:4399–4410