

---

Electronic Thesis and Dissertation Repository

---

11-9-2020 1:00 PM

## Deciphering the CK2-dependent phosphoproteome and its integration with regulatory PTM networks

Teresa Nunez de Villavicencio Diaz, *The University of Western Ontario*

Supervisor: David W. Litchfield, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Biochemistry

© Teresa Nunez de Villavicencio Diaz 2020

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Biochemistry Commons](#), [Bioinformatics Commons](#), and the [Computational Biology Commons](#)

---

### Recommended Citation

Nunez de Villavicencio Diaz, Teresa, "Deciphering the CK2-dependent phosphoproteome and its integration with regulatory PTM networks" (2020). *Electronic Thesis and Dissertation Repository*. 7531. <https://ir.lib.uwo.ca/etd/7531>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## Abstract

Protein functions are regulated by the post-translational addition of covalent modifications on certain amino acids. Depending on their distance within the 3-dimensional structure, addition/removal of individual post translational modifications (PTMs) can be impacted by others. This PTM interplay constitutes an essential regulatory mechanism that interconnects the molecular networks in the cell. Protein CK2, a clinically relevant acidophilic Ser/Thr kinase, may be responsible for 10-20% of the human phosphoproteome. Such estimates agree with the number of known substrates, which continues to expand. Furthermore, the demonstration that CK2 participates in hierarchical phosphorylation and has similar sequence determinants to caspases suggest extensive PTM interplay in CK2-dependent signaling.

In this thesis, we explore the role of lysine acetylation in the vicinity of the phosphorylatable residue(s) as a modulator of phosphorylation by CK2. To explore this association a biochemistry approach was followed to decipher the impact of lysine acetylation in CK2 specificity and a proteomics approach was employed for profiling the affected sites in cells. In solution peptides conforming to the CK2 consensus containing lysine or acetyllysine at positions +1, +2, or +3 downstream the phosphorylatable site, were found only to be phosphorylated by CK2 when either determinant was present at position +2. Linear patterns were generated to reflect this specificity and the PTM databases and the human proteome were searched for substrate hits. The boundaries and cellular conditions of the hits were assessed including conservation, mutations, regulatory role, and enzyme-substrate relationships. Several hits matching the patterns were observed in the cells to be both acetylated and phosphorylated. Since chromatin organization proteins were hits the regulatory role of CK2 in this process was summarized. The data processing and analysis steps followed were incorporated in a data-driven R Shiny web application visualRepo.

Collectively, our work shows how CK2-dependent phosphorylation and lysine acetylation network integration could contribute to the complexity of cellular processes. It also provides a computational and analytical framework for further studies exploring the PTM interplay occurring at a kinase consensus sequence. Finally, given the druggability of CK2, our results

offer new molecular insights for exploring the combination of CK2 and lysine deacetylase inhibitors.

## Keywords

protein kinase CK2, post-translational modifications, modification interplay, lysine acetylation, phosphorylation, proteomics, bioinformatics.

## Summary for Lay Audience

Proteins are main components of the cell and are subjected to many chemical changes or modifications, affecting their function. Some chemical modifications in proteins have been described to interplay with one another resulting in intricate regulatory networks. If we were to look at a protein as a sequence of characters, where two proximal characters were to be substituted, i.e., modified, an interplay between these can occur preventing or facilitating the addition of the second modification once the first is added. The protein CK2 is one of many proteins that modify other proteins in the cell, referred to as substrates. CK2 prefers modifying characters (amino acids) contained in a negatively charged sequence. In this thesis, we decipher how modification by CK2 is enabled by a pre-existing modification (lysine acetylation) that removes a positive charge two amino acids away. Biological databases were searched for protein sequences conforming to this view and several candidates were identified. The sequence features and cellular functions of the candidates were also mined from the databases and the literatures. Finally, for the first time a list of hits containing both modifications occurring in the cell were identified. All the steps in the processing and analysis were incorporated into a web application, visualRepo, for facilitating the reuse and reanalysis of the data. Overall, our work can be extended to identify other modifying proteins and the interplay that can occur. Given CK2 is a therapeutic target, our results provide us with information that can be used to devise combination therapies involving CK2 in the future.

## Co-Authorship Statement

Chapter 1: Published manuscript (*Pharmaceuticals* 2017, 10(1), 27) authored by Teresa Nuñez de Villavicencio-Diaz, Adam J. Rabalski and David W. Litchfield. Teresa generated all the data, figures and tables that are presented. All authors contributed to the ideas for this manuscript as well as editing and writing. © 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Chapter 2: Teresa Nunez de Villavicencio Diaz performed all the experiments, code writing, data processing and analysis, and generated all figures (except MALDI images) and tables that are presented. Kristina Jurcic ran all the mass spectrometry samples and Wen Quin synthesized the in solution peptides. The phosphoproteomic data sets of CX-4945 inhibition generated by Adam Rabalski and Eduard Cruise, previous graduate students in our group, were reused in this thesis' *in silico* analysis for finding potential linear pattern hits.

Chapter 3-4: Teresa Nunez de Villavicencio Diaz performed all the experiments, code writing, data processing and analysis, and generated all figures, tables, scripts, and software that are presented. The phosphoproteomic data set generated by Eduard Cruise, was reused in this thesis' *in silico* analysis for finding functional links between CK2 and chromatin organization.

## Acknowledgments

Our work on protein kinase CK2 and kinase regulatory networks has been supported by an Operating Grant from the Canadian Institutes of Health Research and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada. Teresa Nuñez de Villavicencio Diaz has also been supported by a Dean's Doctoral Scholarship from the Schulich School of Medicine & Dentistry.

I am grateful to my supervisor David Litchfield, my advisory committee Dr. Gloor and Dr. Li, Kristina Jurcic, Adam Rabalski, Edward Cruise, and past and present members of the Litchfield lab for the scientific discussions related to CK2, proteomics, bioinformatics, and kinase regulatory networks. I am also grateful to Barb Green for her support as the Graduate Program Assistant.

Finally, I want to give many thanks to my husband Samy for helping and supporting me during this stage of my career.

# Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iv
Co-Authorship Statement.....	v
Acknowledgments.....	vi
Table of Contents.....	vii
List of Tables.....	xii
List of Figures.....	xv
List of Supplemental Tables.....	xix
List of Supplemental Figures.....	xxiii
List of Appendices.....	xxiv
Chapter 1 - Introduction.....	1
1 Protein Kinase CK2: Intricate Relationships within Regulatory Cellular Networks.....	1
1.1 Functional Networks Involving CK2.....	3
1.2 Protein-Protein Interaction Networks Involving CK2.....	5
1.3 CK2 Networks Derived from Text Mining of the Published Literature.....	7
1.4 Extension of CK2 Networks to Include Other Constituents of Regulatory Networks.....	11
1.5 Rationale for the Study.....	14
1.6 Scope of Thesis.....	15
1.7 References.....	16
1.8 Supplemental Materials.....	25
Chapter 2.....	27
2 Protein Kinase CK2: Interplay Between Cellular Acetylation and Phosphorylation Networks.....	27
2.1 CK2-dependent signaling and lysine acetylation.....	32

2.2	Tentative interplay between CK2-dependent phosphorylation and other PTM types .....	33
2.3	Phosphorylation of in solution peptides matching the CK2 motif with lysine and acetyllysine as determinants .....	34
2.4	Analysis of high-throughput datasets resulting from CK2 manipulation .....	38
2.4.1	<i>In silico</i> analysis of <i>in vitro</i> high-throughput CK2 kinase assay .....	38
2.4.2	<i>In silico</i> analysis of CK2 inhibition experiments by CX-4945 treatment. ....	40
2.4.3	Asp-N digestion experiment .....	42
2.4.4	Acetylated proteins functionally linked to CK2 .....	43
2.5	K motif representation in the proteome and modification networks .....	45
2.5.1	Acetylation motif discovery.....	45
2.5.2	K motif proteome-wide acetylation .....	46
2.5.3	K motif proteome-wide phosphorylation.....	48
2.5.4	K motif representation in the human proteome .....	50
2.5.5	K motif boundaries and cellular conditions .....	53
2.5.6	Functional link to CK2.....	65
2.5.7	Candidate K motif sites for CK2-dependent phosphorylation.....	66
2.6	K motif as a ground for PTM interplay .....	68
2.6.1	Modulation of CK2-dependent signaling by lysine deacetylase inhibitors .....	68
2.6.2	Co-occurrence of lysine acetylation and phosphorylation in neighboring residues .....	74
2.6.3	<i>In vivo</i> phospho-acetyl K motif hits.....	82
2.7	Methods and Materials.....	86
2.8	Discussion.....	102
2.8.1	PTM interplay at CK2 target sites .....	103
2.8.2	CK2 specificity against peptides containing lysine or acetyllysine as determinants .....	103



2.8.3	Identification of K motif candidates .....	105
2.8.4	CK2 and lysine acetylation networks .....	106
2.8.5	Boundary and cellular conditions of the K motif hits.....	108
2.8.6	Integration of phosphorylation and lysine acetylation networks .....	109
2.9	Concluding remarks and implications .....	112
2.10	References.....	113
2.11	Supplemental Materials .....	122
Chapter 3	.....	132
3	Protein kinase CK2: Regulatory Role in Cellular Signaling and its Impact on Chromatin Organization.....	132
3.1	Exploring the link between CK2 and chromatin organization.....	134
3.2	CK2-dependent signaling and histones.....	137
3.3	CK2-dependent signaling and chromatin-modifying enzymes.....	140
3.3.1	CK2-dependent phosphorylation of HDAC lysine deacetylases .....	143
3.4	CK2-dependent signaling and regulation of gene expression.....	145
3.4.1	CK2 regulatory role in PRC-mediated gene repression.....	148
3.4.2	CK2 regulatory role in DNA methylation .....	149
3.4.3	CK2 regulatory role in ribosomal RNA expression.....	150
3.4.4	CK2 regulation of transcription factors .....	153
3.5	Chromatin dynamics intersect with the CK2 protein-protein interaction network .....	155
3.6	CK2-dependent phosphoproteome and chromatin dynamics .....	157
3.6.1	Asp-N phosphoproteome .....	160
3.7	Methods and Materials.....	161
3.8	Concluding remarks .....	162
3.9	References.....	163
3.10	Supplemental Materials .....	176

Chapter 4.....	180
4 Strategies for the functional analysis of lists of modification sites.....	180
4.1 Functional analysis strategies .....	181
4.1.1 Pattern scan and discovery .....	181
4.1.2 Post-translational modification data retrieval .....	184
4.1.3 Kinase enrichment and activity inference analyses* .....	187
4.1.4 Kinase-substrate relationship prediction* .....	191
4.1.5 Functional association to writer enzymes .....	193
4.1.6 Molecular interaction .....	196
4.1.7 Functional relevance of modification .....	197
4.1.8 Frequently used tools .....	198
4.2 visualRepo implementation .....	201
4.2.1 Development environment.....	201
4.2.2 Report objects .....	202
4.2.3 User Interface.....	204
4.2.4 Reports .....	207
4.2.5 Functional analysis.....	209
4.3 Concluding remarks .....	212
4.4 References.....	213
4.5 Supplemental Material .....	219
Chapter 5.....	224
5 Discussion .....	224
5.1 Summary of findings and research impact.....	225
5.1.1 Intricate relationships of CK2 with lysine acetylation networks .....	225
5.1.2 CK2-dependent regulation and chromatin organization .....	226
5.1.3 Proteome-wide distribution of CK2 sites containing lysine residues .....	227

5.1.4	visualRepo web application implementation for data processing and analysis.....	227
5.1.5	Limitations .....	228
5.2	Future directions .....	229
5.3	Concluding remarks .....	229
5.4	References.....	230
	Appendices.....	233
	Curriculum Vitae .....	234

## List of Tables

Table 1.1. CK2 subunits interactors extracted from a human protein-protein interaction network built by retrieving the human interactome from the BioGRID database build 3.4.129 using Cytoscape v3.4.0. *Interaction pairs: a pair of interactors including CK2-CK2 substrate and CK2 substrate-CK2 substrate. ....	5
Table 1.2. Post-translational modifications (PTMs) in the vicinity of CK2 <sup>1</sup> target sites retrieved from PhosphoSitePlus (accessed August 2016). ....	13
Table 2.1. PTM interplay annotated for known human CK2 substrates, retrieved from PSP, in the database PTMcode v2. ....	34
Table 2.2. Linear patterns (K motifs) matching the sequences studied in the in solution peptides. ....	36
Table 2.3. K motif hits represented in the Asp-N phosphoproteomic dataset. ....	43
Table 2.4. Significant motifs in an acidic environment discovered on acetylation sites identified in Scholz <i>et al.</i> 2015 [24] using MoMo v5.0.5. ....	45
Table 2.5. Acetylation sites identified in Scholz <i>et al.</i> 2015 [24] matching the K motif variants and total number of acetylation sites in the vicinity of Ser/Thr residues; unique sequences. ....	46
Table 2.6. Acetylation sites annotated in the PSP database (September 2019) conforming to K motif variants and corresponding phosphorylation data, unique sites. ....	47
Table 2.7. K motif representation in PSP phosphorylation data, unique sites. ....	48
Table 2.8. K motif hits annotated phosphorylated in the database iPTMnet (September 2019) and reported acetylated in iPTMnet and/or the Scholz <i>et al.</i> 2015 data [24]; corresponding PTM interplay from PTMcode v2. Unique sites. ....	49
Table 2.9. PTM types targeting the K motif hits annotated in the database iPTMnet (unique sites). ....	49

Table 2.10. Kinase-substrate relationships corresponding to K motif hits annotated phosphorylated in the PSP and iPTMnet databases (September 2019). .....	50
Table 2.11. K motif representation on the human proteome with FIMO v5.0.5 by using the MoMo v5.0.5 motifs discovered in the acetylation sites identified in Scholz <i>et al.</i> 2015 [24] as input; unique sequences. UniProtKB human sequences reviewed and isoforms April 26, 2019.....	51
Table 2.12. K motif representation on the human proteome using ScanProsite and EMBOSS, unique sites.....	52
Table 2.13. K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off $\geq 0.4$ ). .....	53
Table 2.14. Warnings detected on K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off $\geq 0.4$ ). .....	53
Table 2.15. SLiMSearch4 functional features on K motif hits. ....	57
Table 2.16. Mutations targeting the positions in the K motif returned by SLiMSearch4.....	59
Table 2.17. Enzyme-substrate relationships represented in the positions matched in the K motif (iPTMnet). .....	64
Table 2.18. Summary of GO molecular function, level 5, classification corresponding to the candidate CK2 substrates conforming to the K motif variant.....	67
Table 2.19. Known modifications annotated to the identified phospho-acetyl sites retrieved from iPTMnet (December 2019) and PSP (September 2019) databases. ....	78
Table 2.20. GO molecular function classification, level 5, associated with writers, erasers, and readers of containing co-occurring phosphorylation and lysine acetylation sites.....	79
Table 2.21. Protein complexes represented among the proteins containing the identified phospho-acetyl sites. ....	80

Table 2.22. Phosphorylated and acetylated sites that conform to CK2 consensus and K motif hits identified <i>in vivo</i> in peptides with co-occurring phosphorylation and acetylation; includes canonical and isoform sequences.....	83
Table 2.23. Co-occurring phosphorylation and acetylation sites <i>in vivo</i> at K motif variants (the table includes ambiguously localized sites).....	84
Table 2.24. K motif hits boundaries and cellular properties determined by SLiMSearch4....	85
Table 2.25. K motif in solution peptides.....	87
Table 2.26. Antibodies used.....	89
Table 2.27. Isotopically-labelled amino acids. ....	96
Table 3.1. Histone modification functionally linked to CK2.....	138
Table 3.2. CSNK2A1 subunit interactors members of chromatin organization complexes (GO cellular compartment, level 5). ....	155
Table 3.3. Proteins involved in chromatin organization that contains at least one hit for CK2 consensus sequence with changed phosphorylation when CK2 is manipulated.....	158
Table 4.1. Prediction of kinase-substrate relationships for known human CK2 substrates (PhosphoSitePlus).....	193

# List of Figures

Figure captions were shortened for brevity.

Figure 1.1. CK2-centered functional association network represented with the database STRING v10.0. ....	4
Figure 1.2. Protein-protein interaction network of human CK2 subunits generated using Cytoscape v3.4.0 and the BisoGenet v3.0.0 plugin. ....	6
Figure 1.3. CK2 Functional association networks represented in Cytoscape v3.4.0.....	8
Figure 1.4. Protein-protein interaction network of CK2 subunits expanded to include kinase information of the interactors.....	12
Figure 2.1. <i>In vitro</i> phosphorylation of peptides containing K and AcK (acetyllysine) by CK2.....	35
Figure 2.2. Known CK2 target sites in PhosphoSitePlus (PSP) containing Lys in the vicinity of the phosphorylatable site. ....	38
Figure 2.3. Sites phosphorylated <i>in vitro</i> by CK2 identified by Bian <i>et al.</i> 2013 [49] that contained Lys in the vicinity of the phosphorylatable residue. ....	39
Figure 2.4. Phosphosites identified and quantified in three different phosphoproteomic studies of CX-4945 inhibition that match the K motif variants.....	41
Figure 2.5. Acetyl-proteins identified by Scholtz <i>et al.</i> 2015 functionally linked to CK2.....	44
Figure 2.6. SLiMSearch4 results for the K motif variant +2 (K motif extended (ext_SLiMSearch) and all hits (SLiMSearch)) showing the distribution of the disordered, conservation, and Anchor scores for the hits found.....	56
Figure 2.7. SLiMSearch4 hits for K motif +2 variant conserved counter (Outliers coefficient = 1.5) and conservation, motif v flank scores (Outliers coefficient (Tukey's method) = 5). .	57

Figure 2.8. Curated mutations annotated in COSMIC (February 2020) on sites matching the K motif. A) Number of mutations per K motif variants. ....	62
Figure 2.9. Literature relationships between CK2 and lysine deacetylases and acetyltransferases extracted using Chilibot.....	65
Figure 2.10. K motif hits in the human proteome functionally linked to CK2.....	66
Figure 2.11. Immunoblots of lysine deacetylase inhibitor (KDACi) treatment of U2OS and HeLa cells at different time points monitoring total protein acetylation status and the modulation of CK2 subunits and CK2 substrates. ....	70
Figure 2.12. SILAC workflow of lysine deacetylase (KDAC) inhibitor treatment of U2OS cells. ....	72
Figure 2.13. Known and tentative CK2 target sites and substrates quantified in U2OS cells treated with the lysine deacetylase inhibitors .....	74
Figure 2.14. Label-free workflow to study the acetylated fraction of the phosphoproteome. ....	75
Figure 2.15. Functional classification of phosphoproteins containing acetylation sites in HeLa and USOS cells, biological processes observed in common in Scholz <i>et al.</i> study and our study.....	76
Figure 2.16. Histone code observed. Red: histone modifications annotated in iPTMnet and PSP databases and Scholz <i>et al.</i> 2015 study [24].....	77
Figure 2.17. Established framework. Round boxes: main results; orange border: main experiments. ....	113
Figure 3.1. Bioinformatics workflow implemented for finding chromatin organization proteins functionally linked to CK2. Left: examples of what to expect from literature and ranking analysis. ....	135
Figure 3.2. Chromatin organization proteins functionally linked to CK2. A) Annotations included and score calculation B) AmiGO chromatin remodeling.....	136



Figure 3.3. Merged kinase-substrate relationships found in PhosphoSitePlus (PSP) and iPTMnet databases. A) Summary of available information.....	141
Figure 3.4. Epigenetic regulation of gene expression in <i>Homo sapiens</i> (Reactome identifier: R-HSA-212165). * Examples of CK2 substrates in the pathway. ....	151
Figure 4.1. Motif discovery and pattern matching tools and strategies. ....	183
Figure 4.2. PTM profile of CK2 substrates.....	186
Figure 4.3. Kinase set enrichment analysis.....	190
Figure 4.4. Functional association of a list of modulated phosphoproteins to CK2. ....	195
Figure 4.5. Scholarly literature that co-mention (pairs) the tools and databases described here. ....	200
Figure 4.6. Development environment for visualRepo application implementation.....	202
Figure 4.7. Diagram of classes of reports in the visualRepo application. The fields show the information collected for each category and the methods the steps required to store the information.....	203
Figure 4.8. Diagram of dataAbout and linkTo classes example of data applied to each class. Training set refers to the information collected to CK2 whereas the topic of interest is the list of proteins or genes of interest to functionally associate to CK2. ....	204
Figure 4.9. visualRepo user interface. ....	205
Figure 4.10. visualRepo functional analysis example. ....	206
Figure 4.11. Visualization of report data with visualRepo. ....	208
Figure 4.12. Visualization examples of mutation, motif hit, and site data with visualRepo for K motif hits position +2. ....	209
Figure 4.13. visualRepo site quantification hypothesis testing.....	212



## List of Supplemental Tables

The supplemental materials are provided within this thesis or as separate files whenever necessary due to their size.

Supp. Table 1.1. STRING network, function, and annotations of the top 50 CK2-related genes. ....	25
Supp. Table 1.2. CK2-interacting proteins in the human interactome.....	25
Supp. Table 1.3. CK2-interacting hub proteins. ....	25
Supp. Table 1.4. CK2-interacting proteins of each subunit.....	25
Supp. Table 1.5. CK2-related GO annotations extracted from the literature and protein complex data. ....	26
Supp. Table 1.6. PTM sites in the vicinity of CK2 phosphorylated sites. ....	26
Supp. Table 1.7. High-resolution images.....	26
Supp. Table 2.1. Chilibot and PubMed query: “ck2” AND “acetylation” with network representation and GO biological process enrichment. ....	122
Supp. Table 2.2. Known CK2 target sites annotated in PhosphoSitePlus (PSP) and iPTMnet databases matching the K motif variants and their function (UniProt). ....	122
Supp. Table 2.3. K motif hits from <i>in silico</i> analysis of <i>in vitro</i> high-throughput CK2 kinase assay.....	122
Supp. Table 2.4. Differentially inhibited phosphosites identified in three different phosphoproteomic studies of CX-4945 inhibition that match the K motif variants. ....	123
Supp. Table 2.5. Motif discovery on acetylation sites identified in Scholz <i>et al.</i> 2015 using MoMo v5.0.5. ....	123

Supp. Table 2.6. K motif hits annotated phosphorylated and acetylated in PSP database (September 2019); does not imply co-occurrence of phosphorylation and acetylation (September 2019).....	123
Supp. Table 2.7. K motif hits annotated phosphorylated in PSP database (September 2019). .....	123
Supp. Table 2.8. K motif hits annotated phosphorylated in the iPTMnet database (September 2019) reported acetylated in iPTMnet and/or the Scholz <i>et al.</i> 2015 data and corresponding PTM interplay from PTMcode v2 (October 2019). ....	123
Supp. Table 2.9. Kinase-substrate relationships corresponding to K motif hits annotated phosphorylated in the PSP and iPTMnet databases (September 2019). ....	123
Supp. Table 2.10. K motif representation on the human proteome (September 2019) with FIMO v5.0.5 by using the MoMo v5.0.5 motifs discovered in the acetylation sites identified in Scholz <i>et al.</i> 2015 as input; unique sequences. UniProtKB human sequences reviewed and isoforms April 26, 2019. ....	124
Supp. Table 2.11. K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off $\geq 0.4$ ). ....	124
Supp. Table 2.12. UniProtKB domain and COSMIC mutations (February 2020) functional features of K motif hits. ....	124
Supp. Table 2.13. Functional features of sequences conforming to K motif hits: Regulatory Sites, CGS, KDAC-KAT, and Sites in Pathways. ....	124
Supp. Table 2.14. Summary of CK2 candidate sites matching the K motif. ....	124
Supp. Table 2.15. CK2 candidate sites modulated by lysine deacetylase treatment in U2OS cells. ....	125
Supp. Table 2.16. Peptides with acetylation and phosphorylation sites co-occurrences. ....	125
Supp. Table 2.17. GO level 5 classification of proteins with co-occurring phosphorylation and acetylation sites. ....	125

Supp. Table 2.18. Known sites (iPTMnet and PSP), known kinase-substrate relationships (iPTMnet), and known regulatory sites (PSP) corresponding to the identified sites in acetylated and phosphorylated peptides (September and December 2019).....	125
Supp. Table 2.19. CK2 consensus and K motif hits found <i>in vivo</i> and their corresponding protein function and site modification information. ....	125
Supp. Table 3.1. Literature text mining for co-mentions of CK2 and chromatin organization and histone proteins. ....	176
Supp. Table 3.2. Chromatin organization data sets. ....	176
Supp. Table 3.3. Functional ranking of chromatin organization proteins to CK2.....	176
Supp. Table 3.4. Subunits of complexes linked to CK2 and chromatin organization. ....	177
Supp. Table 3.5. CK2 interactors by subunits involved in chromatin organization. ....	177
Supp. Table 3.6. GO subcellular compartment classification of CK2 interactors involved in chromatin organization. ....	177
Supp. Table 3.7. Phosphosites modulated when CK2 is manipulated using Asp-N as protein digestion enzyme involved in chromatin organization and transcription. ....	177
Supp. Table 4.1. CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019.....	219
Supp. Table 4.2. CK2 target sequences motif discovery using MoMo v.5.0.5. ....	219
Supp. Table 4.3. FIMO v.5.0.5 proteome scan for motifs discovered with MoMo v.5.0.5..	219
Supp. Table 4.4. SDSE hits that are annotated as phosphorylated in PhosphoSitePlus (PSP) database.....	220
Supp. Table 4.5. PTM types annotated in the vicinity (+/-7 amino acids) of phosphorylated SDSE hits.....	220

Supp. Table 4.6. SDSE hits known kinase-substrate relationships and phosphorylation <i>in vitro</i> by CK2. ....	220
Supp. Table 4.7. Kinase set enrichment analysis and activity inference. ....	220
Supp. Table 4.8. CoPhosK, iGPS, and NetworKIN prediction using as input CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019. ....	220
Supp. Table 4.9. Co-regulated proteins of CK2 subunits retrieved from ProteomeHD. ....	220
Supp. Table 4.10. Co-regulated proteins of CK2 subunits that are known CK2 interactors, June 2020. ....	220
Supp. Table 4.11. Regulatory CK2 target sites and domains represented, September 2019.	221
Supp. Table 4.12. Summary of strategies and tools discussed in Chapter 4. ....	221
Supp. Table 4.13. Network representation of tools and databases included in the strategies that are co-mentioned in the Scholarly literature. ....	221

## List of Supplemental Figures

The supplemental materials are provided within this thesis or as separate files whenever necessary due to their size. Figure captions were shortened for brevity.

Supp. Figure 2.1. Functional association STRING network corresponding to the PubMed query: “ck2” AND “acetylation” represented with Cytoscape v3.8.0; confidence cut-off score: 0.7, and a maximum number of proteins: 100.....	126
Supp. Figure 2.2. MALDI quality analysis of <i>in-solution</i> peptides, > 1000 m/z with > 3% intensity. 1 Dalton difference amide C-terminal. P/p: position, ac: acetylation. ....	127
Supp. Figure 2.3. CTTN modification status in U2OS cells. A) Antibody optimization using U2OS lysate. B) Quantification of antibody optimization. C) PEAKS coverage and CTTN modifications after in-gel digestion and LC-MS/MS analysis. ....	128
Supp. Figure 2.4. SLiMSearch4 scores for K motif variants A) +1 and B) +3; K motif extended (ext_SLiMSearch) and all hits (SLiMSearch).....	129
Supp. Figure 2.5. K motif variants A) +1 and B) +3; K motif extended (ext_SLiMSearch) and all hits (SLiMSearch).....	131
Supp. Figure 2.6. PDB secondary structure information for K motif hits found in the human proteome scan. ....	131
Supp. Figure 3.1. Summary of enzyme-substrate relationships and site information retrieved from iPTMnet and PhosphoSitePlus (PSP) databases. ....	178
Supp. Figure 3.2. Chromatin-remodeling complexes and CK2 highlighting core subunits that are CK2 substrates and/or are modulated upon CK2 inhibition. ....	179
Supp. Figure 4.1. Kinase-substrate prediction for CK2 sites.....	222
Supp. Figure 4.2. Motif and domain information for the CK2 targets sites in HDAC1, Ser421, and Ser423. ....	223

## List of Appendices

Appendix A. Supplemental Materials.....	233
Appendix B. LC-MS runs and protein/peptide identification files.....	233
Appendix C. Scripts and computational resources. ....	233



## List of Abbreviations

BP	Biological process(es)
Bx	Bufexamac
C/Cys	Cysteine
CC	Cellular component(s)
DMSO	Dimethyl sulfoxide
GO	Gene Ontology
HDAC	Histone deacetylase
K/Lys	Lysine
KAT	Lysine acetyltransferase
KDAC	Lysine deacetylase
KDACi	Lysine deacetylase inhibitor
LC	Liquid Chromatography
LC-MS/MS	Liquid chromatography–mass spectrometry
M/Met	Methionine
MF	Molecular function(s)
MS	Mass Spectrometry
ms <sup>2</sup>	Tandem mass spectrometry
m/z	mass to charge ratio
NAM	Nicotinamide
PBS	Phosphate-buffered Saline
PSP	PhosphoSitePlus
R/Arg	Arginine
S/Ser	Serine
SIRT	Sirtuin
SILAC	Stable Isotope Labeling by/with Amino acids in Cell culture
T/Thr	Threonine
TSA	Trichostatin A

## Chapter 1 - Introduction

# 1 Protein Kinase CK2: Intricate Relationships within Regulatory Cellular Networks

Since its original discovery more than 50 years ago, protein kinase CK2 has been implicated in a continually expanding array of biological processes [1]. In this respect, CK2 has been shown to be a participant in the regulation of cellular processes such as transcription [2,3] and translation [4,5,6,7], control of protein stability [8,9,10] and degradation [11,12], cell cycle progression [13], cell survival [14,15,16], and circadian rhythms [17]. CK2 has also been linked to various aspects of tumor progression and suppression and has been shown to be elevated in many forms of cancer [18,19] as well as in virally infected cells [20,21,22]. Consequently, CK2 has recently emerged as a potential therapeutic target with two CK2 inhibitors, namely CX-4945 [23,24] and CIGB-300 [25,26], currently in clinical trials [27,28,29] for cancer treatment.

In humans, CK2 is typically considered to be a tetrameric enzyme comprised of two catalytic subunits (CK2 $\alpha$  and/or CK2 $\alpha'$  subunits that are encoded by the CSNK2A1 and CSNK2A2 genes, respectively) and two regulatory CK2 $\beta$  subunits [1,30] (encoded by the CSNK2B gene). Although typically classified as a protein serine/threonine kinase based on sequence relationships to other members of the protein kinase superfamily, CK2 has also been shown to exhibit protein tyrosine kinase activity [31,32,33,34]. Biochemical characterization of its enzymatic activity has demonstrated that CK2 is an acidophilic kinase with a consensus recognition motif that features aspartic acid and glutamic acid residues as well as some phosphorylated residues as its dominant specificity determinants [35,36,37]. Characterization of its specificity determinants has contributed to the identification of many CK2 substrates that have been shown to be directly phosphorylated by CK2 [38,39]. Phosphoproteomic profiling has also revealed many putative substrates that have been shown to be phosphorylated in cells at sites that match the consensus for phosphorylation by CK2 [40,41]. In fact, analysis of phosphoproteomic datasets typically suggests that CK2 could be responsible for more than 10% of the phosphoproteome [42].

While there is ample evidence that CK2 is an important constituent in the regulation of many fundamental biological processes, there are unresolved issues regarding its regulation in cells. In this respect, questions regarding its regulation arise because the catalytic subunits of CK2 are enzymatically active in the presence or absence of the regulatory CK2 $\beta$  subunit. The role of the regulatory subunit is to provide a negatively charged docking site required for the phosphorylation of a certain group of CK2 substrates that contain a positive stretch of amino acids downstream the CK2 $\alpha$ /  $\alpha'$  target site [39]. Furthermore, the activity of CK2 is generally unaffected by second messengers, and unlike many kinases that are regulated by phosphorylation within an activation loop, the activation loop of CK2 is devoid of regulatory phosphorylation sites [43,44]. The fact that the catalytic subunits of CK2 are fully active when expressed as recombinant proteins in bacteria further suggests that the enzyme may be constitutively active [45,46,47]. Consequently, it is unclear how a constitutively active enzyme can be a key regulatory participant in tightly regulated cellular processes. To investigate this apparent paradox, we have examined the relationship between CK2 and other constituents of the regulatory networks within cells by performing text mining of the published literature and mining of proteomic databases. Furthermore, motivated by our demonstration that CK2 phosphorylation sites overlap with other post-translational modifications to enable CK2 to modulate caspase cleavage [48,49,50] and to participate in hierarchical phosphorylation relationships, we have analyzed proteomic databases to identify post-translational modifications that may regulate, or be regulated by, CK2 phosphorylation. A similar analysis was performed with phosphoproteomic data generated by our group to corroborate and expand the previous findings.

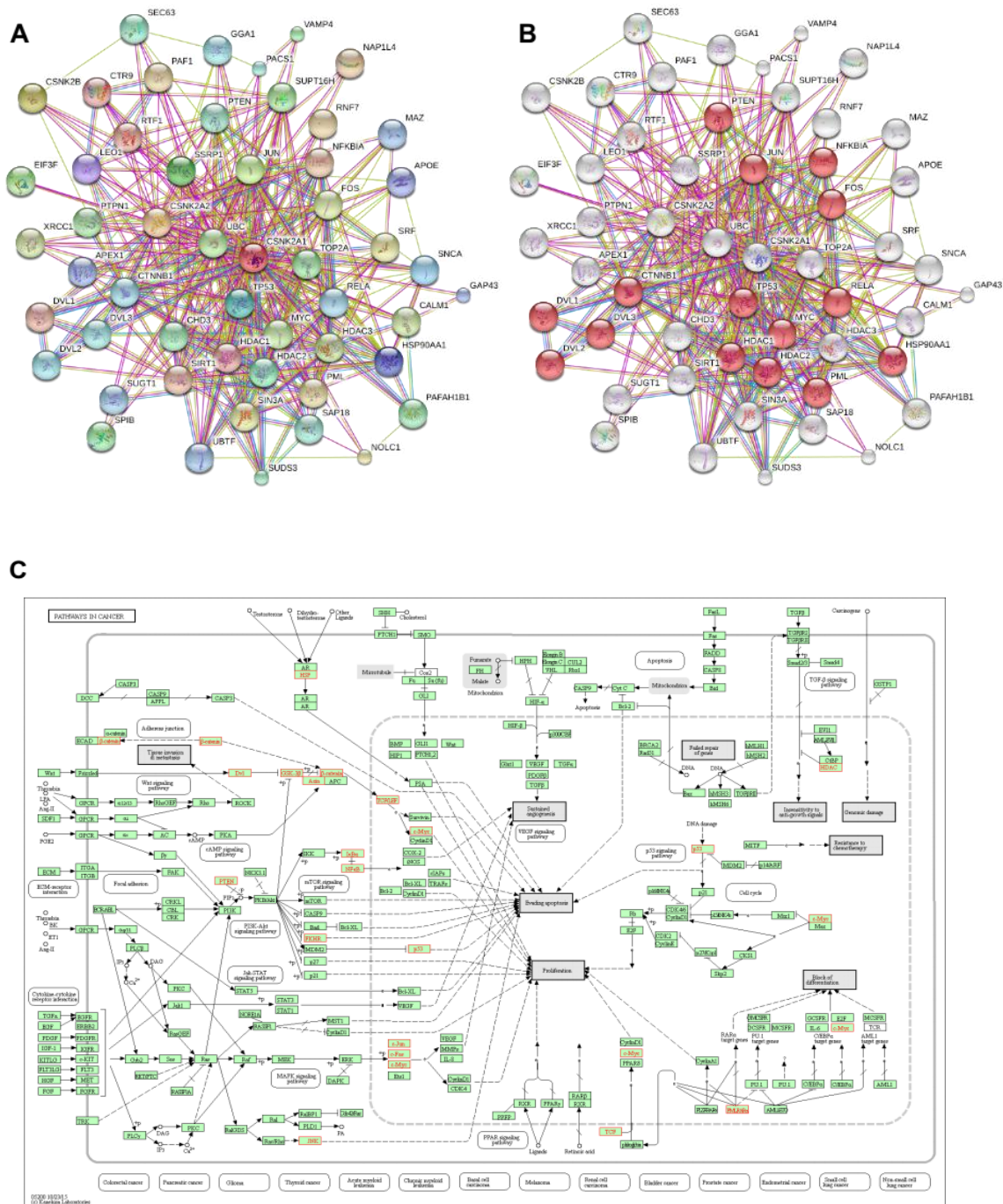
Overall, in this study, we explore database and literature information available and our own high-throughput data in the context of CK2-dependent signaling with the objective of highlighting and discussing the extensive interplay of CK2 with regulatory networks in the cell. The analysis that we have provided is also intended to provide functional perspectives to the data available in the databases since these data are often isolated from the information regarding CK2 that exists within the literature; especially for non-experts in the field and for interpreting data related to CK2 emerging from existing or new high-throughput studies.

## 1.1 Functional Networks Involving CK2

The impact of CK2 in the cell can be, to some extent, predicted by considering the number of biological processes in which it has been implicated and the number of substrates that have been reported to date [39]. Furthermore, it is anticipated that the published literature will represent an important resource for deciphering and validating information that is emerging from genome- and proteome-wide analyses. Consequently, we performed text mining to identify publications that highlight CK2 and aspects of its function or regulation. To this point, CK2 substrate information remains “sparse” in the published literature that is represented by more than 2600 papers (PubMed [51] search: “Casein Kinase II” [Mesh]) directly describing CK2 function and more than 5000 papers mentioning the kinase (GoPubMed [52] search). Nevertheless, assembling this information to obtain a global but detailed view of CK2-dependent networks is one step towards deciphering genome and proteome scale analyses of CK2.

An initial evaluation of the functional relationship of CK2 with other cellular proteins can be obtained by querying STRING v10.0 [53], a database with known and predicted functional associations between proteins (see Fig. 1.1A for a representation of the top 50 proteins functionally associated to CK2, Supp. Table 1.1). According to this analysis, CK2 regulates the activity of at least 15 cancer-related proteins such as the tumor suppressor TP53, the histone deacetylases HDAC1 and HDAC2, and the NFkB subunit RELA (Fig. 1.1B,C, Supp. Table 1.1). CK2-dependent phosphorylation of these proteins has been reported either *in vitro* or *in vivo* [54,55,56,57] with the majority of the target sites identified conforming to the minimum CK2 consensus sequence: [ST]xx[DEpS]. The CK2 functional relationship to these substrates places the kinase in a central position in human protein-protein interaction networks since such proteins are considered “information hubs” [58]. For instance, the proteins TP53, HDAC1, HDAC2, and RELA bind to at least 997, 554, 323, and 271 unique interactors based on the BioGRID protein-protein interaction repository [59] (accessed 30 November 2016). In fact, in a human protein-protein interaction network (built from the BioGRID database in Cytoscape [60,61] v3.4.0, self-loops and duplicated edges removed) CK2 can “influence” approximately 23% of the established interactions (63,988 interaction pairs out of

270,000) if we assume a ‘guilt by association’ approach considering CK2 direct interactions (meaning step 1 interactors: 629 proteins, Supp. Table 1.2) and that of its indirect interactors (meaning step 2 interactors: 11,869 proteins, Supp. Table 1.2).



**Figure 1.1.** CK2-centered functional association network represented with the database STRING v10.0. (A) The network shows the top 50 human genes related to CK2 subunits

CSNK2A1, CSNK2A2, and CSNK2B (number of nodes and edges: 53 and 369, respectively) based on neighborhood, experiments, text mining, and database sources. Briefly, the human CK2 subunits were searched in STRING by gene name using the multiple protein search functionality. The functional association network was then retrieved by selecting 50 as the maximum number of interactors to show from the first shell (step 1) and checking the mentioned interaction sources (these options are found within the “data settings” drop-down menu); (B) A clone of network A with the red nodes representing proteins connected to deregulated pathways (KEGG pathway: hsa05200); (C) A network representation of the map: Pathways in cancer (KEGG pathway: hsa05200). The map was downloaded from the KEGG PATHWAY [63] database (accessed 30 November 2016) and imported to Cytoscape with the CyKEGGParser v1.2.7 plugin [64]. A high-resolution image of this figure is also available in the Appendix A.

A summary of the number of CK2 interactors for each human CK2 subunit is presented in Table 1.1. Based on this analysis and a search using the “find a gene” functionality of the Enrichr tool [62], further hub interactors of CK2 that can be identified include CDK1, XRCC6, CREB1, HNRNPA1, LYN, YWHAQ, FOS, and MAPK1 (Supp. Table 1.3).

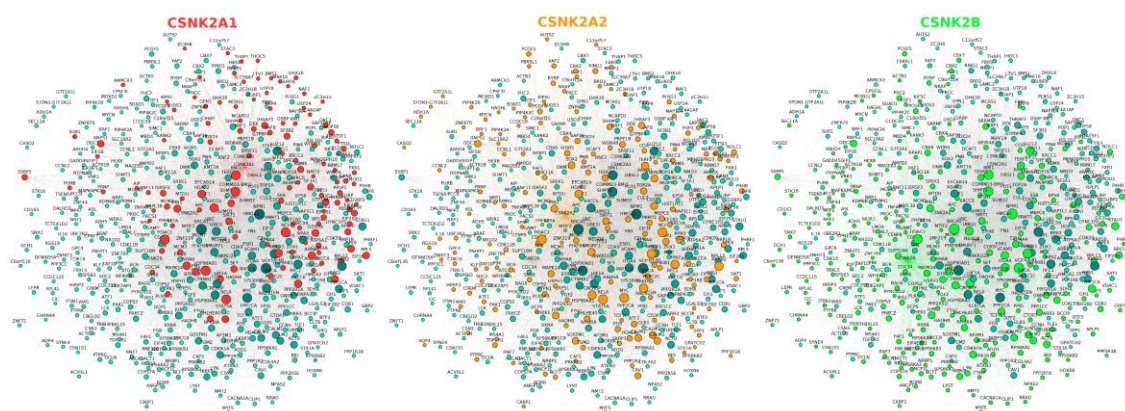
**Table 1.1.** CK2 subunits interactors extracted from a human protein-protein interaction network built by retrieving the human interactome from the BioGRID database build 3.4.129 using Cytoscape v3.4.0. \*Interaction pairs: a pair of interactors including CK2-CK2 substrate and CK2 substrate-CK2 substrate.

Direct interactors of	Number of interactors
CSNK2A1 (1457)	398 unique direct interactors; 435 interaction pairs*
CSNK2A2 (1459)	155 unique direct interactors; 171 interaction pairs.
CSNK2B (1460)	247 unique direct interactors; 270 interaction pairs.
All CK2 subunits	632 unique direct interactors from which 36 are shared by the three subunits and 95 by two.
All CK2 subunits and their direct interactors	12502 unique direct interactors (632 step 1 and 11875 step 2); 63988 interaction pairs.

## 1.2 Protein-Protein Interaction Networks Involving CK2

In a complementary analysis, the direct interactors of CK2 subunits were represented in a protein-protein interaction network (Fig. 1.2, Supp. Table 1.4) using BisoGenet [65] v3.0.0 Cytoscape plugin for the retrieval of physical interaction information. As expected from Table 1.1, differences in the number and identity of the panel of interactors can be observed for each CK2 subunit which suggests a certain degree of functional divergence

as previously highlighted in the literature [66,67,68,69,70], and encourages the development of tools that allow us to differentiate the contributions of the endogenous catalytic subunits to the phosphoproteome. In addition, it points to CSNK2B as a hub itself, suggesting that it may have a role in coordinating interactions with the catalytic subunits of CK2 to modulate phosphorylation of certain substrates. Furthermore, since the interaction network for CSNK2B does not completely overlap that of the catalytic CK2 subunits (Fig. 1.2, Supp. Table 1.4), this analysis reinforces the prospect that CSNK2B has CK2-independent roles within cells as previously described in the literature [67]. In fact, the CSNK2B-dependent interactome has been previously profiled using mouse brain homogenates [71] where CSNK2B is thought to have a crucial role since its mRNA expression levels are 2–3-fold higher compared with other organs, except the testis [72]. In this setting, CSNK2B was found to interact with both cytoplasmic and nuclear-localized proteins involved in protein synthesis, RNA and DNA processing, the cytoskeleton, cell signaling, and transport [71]. Although not included among the references retrieved by BisoGenet, the functional classification of the proteins identified as part of the CSNK2B-dependent interactome is in agreement with the network generated by BisoGenet.



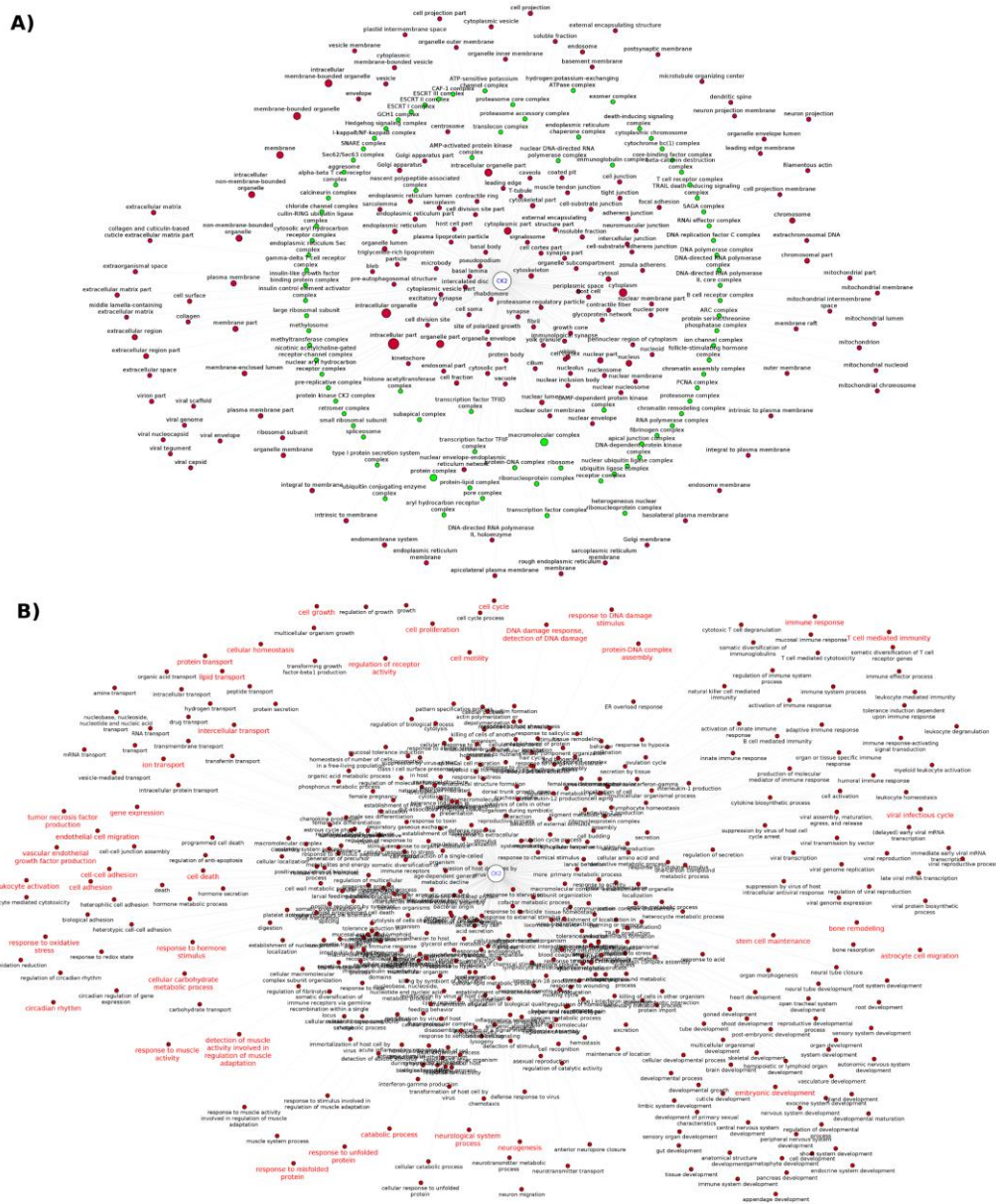
**Figure 1.2.** Protein-protein interaction network of human CK2 subunits generated using Cytoscape v3.4.0 and the BisoGenet v3.0.0 plugin. The three networks are clones, each of which highlights the interactors of CSNK2A1 (red nodes), CSNK2A2 (orange nodes), and CSNK2B (green nodes) subunits, respectively. Briefly, the network was represented by querying SysBiomics (BisoGenet’s interaction database) through the plugin’s interface using the human CK2 subunit gene names and selecting “protein-protein interaction” as the biorelation type and the input nodes and neighbors to step 1 method as

the criteria for building the network. A high-resolution image of this figure is also available in the Appendix A.

### 1.3 CK2 Networks Derived from Text Mining of the Published Literature

A way of summarizing and systematizing CK2 knowledge relies on the use of text mining to access literature information. In this regard, the analysis of GO cellular component annotations using GoPubMed revealed more than 200 subcellular locations and protein complexes studied in the context of CK2 (summarized in Fig. 1.3, Supp. Table 1.5). This analysis also reflects the functional pleiotropy of CK2, which can associate with molecular machineries such as the ribosome, spliceosome, proteasome, and chromatin remodeling complexes, and to other smaller more dynamic complexes such as TRAIL-death inducing complex and the Ikappa-NFkB complex. Directly interrogating the literature followed by data extraction can provide information that otherwise may be missed if we only consider specialized databases such as the mammalian protein complexes database CORUM [73] where CSNK2A1, CSNK2A2, and CSNK2B (Supp. Table 1.5) are listed only as members of the “PDGF treated Ksr1-CK2-MEK-14-3-3 complex”, the “MKP3-CK2alpha complex”, the “Casein kinase II-HMG1 complex”, and the “UV-activated FACT complex” with CSNK2B also listed as a member of the “Fgf2-Ck2 complex” (accessed 30 November 2016). However, protein complex data extracted from the CORUM database has been manually curated whereas the data extracted from the literature needs to be critically analyzed since the extraction process may be ambiguous and biased towards the algorithm used by the tool [74], in this case, GoPubMed. A similar analysis can be made for the GO biological processes. The retrieval of CK2 related GO biological process from literature highlights frequently studied core processes such as cell cycle, cell proliferation, DNA damage, cell death, and viral infectious cycle, as well as other hot topics in the field (Fig. 1.3B, Supp. Table 1.5). The latter includes the involvement of CK2 in embryonic development [75,76,77,78], T cell-mediated immunity [79,80], inflammation [81], glucose homeostasis [82,83,84,85], ion transport [86], bone remodeling [78,87], neurogenesis [88,89], neurological system process [90], response to misfolded and unfolded protein [90], stem cell differentiation and maintenance [91], and response to muscle activity [92,93].





**Figure 1.3.** CK2 Functional association networks represented in Cytoscape v3.4.0 for illustrating the: (A) Cellular components mentioned together with CK2 in the literature identified by GoPubMed search. The green nodes represent protein complexes and the node size highlights terms frequently co-occurring with CK2; (B) Biological processes mentioned together with CK2 in the literature identified by GoPubMed search; the red label-nodes indicate hot topics in CK2 research. Briefly, the term “Casein Kinase II”[Mesh] was queried using the GoPubMed text mining tool and the output was downloaded to generate a network in SIF format [61] for Cytoscape input by specifying CK2 and the GO subcellular components/biological processes as nodes and their co-occurrence in the literature as the interactions (binary type: yes/no). The node size was not set differentially for the biological processes network as this will affect the

visualization and readability. A high-resolution image of this figure is also available in the Appendix A.

Text mining of the CK2-related literature can also provide insights regarding holoenzyme-dependent regulation, which relates to the events where CK2-mediated phosphorylation of a given substrate is positively or negatively regulated by holoenzyme formation and the presence or absence of CSNK2B [45]. In this case, the analysis of protein-protein interaction data for CK2 subunits alone is insufficient for assuming holoenzyme-dependent regulation. A recent *in silico* study relied on text mining for retrieving known holoenzyme-dependent substrates and generated sequence patterns for predicting novel candidates based on structural information of the known substrates [39]. Information on the holoenzyme-dependent substrates can also be obtained from the PhosphoSitePlus database [94,95] by querying “substrates of CK2B”; however the list obtained is not comprehensive (accessed February 2017) when compared to the text mining study [39]. Furthermore, substrates known to be holoenzyme-dependent such as PDX1 [96] and CFTR [97] are cataloged as phosphorylated by the catalytic subunit in this database (query: “substrates of CK2A1”). To avoid such inconsistencies and misleading information, researchers are encouraged to carefully review the evidence provided in databases such as PhosphoSitePlus, which are obtained through automated literature text mining and thus error prone.

In addition to text mining, systematic proteomic studies represent rich resources for potentially uncovering CK2-regulated biological processes and pathways when datasets from these studies are uploaded to repositories [98] such as PRIDE and MassIVE and/or provided as supplementary information. To explore the availability of proteomics data for CK2, we performed a Pubmed search and, as a result, retrieved two proteome studies [99,100], one interactome study [71], and four phosphoproteome studies [40,41,101,102]. However, a PRIDE search only returned one of these studies, which explores the function of CSNK2A1 from *Ostreococcus Tauri*, in a minimal circadian system [101] (ID: PXD000975). Consequently, for the remainder of these studies, information available (e.g., protein and/or phosphorylation site identification and/or relative quantification values) is limited to that provided in the original paper or by direct request to the authors. We also searched PhosphoSitePlus and found that only one phosphoproteome study out

of the four identified in Pubmed was included (accessed February 2017). This phosphoproteomic study explores the short-term response of HEK-293T cells treated with the CK2 inhibitor quinalizarin [102]. However, PhosphoSitePlus only mentions the proteins and phosphosites identified without providing any quantitative results.

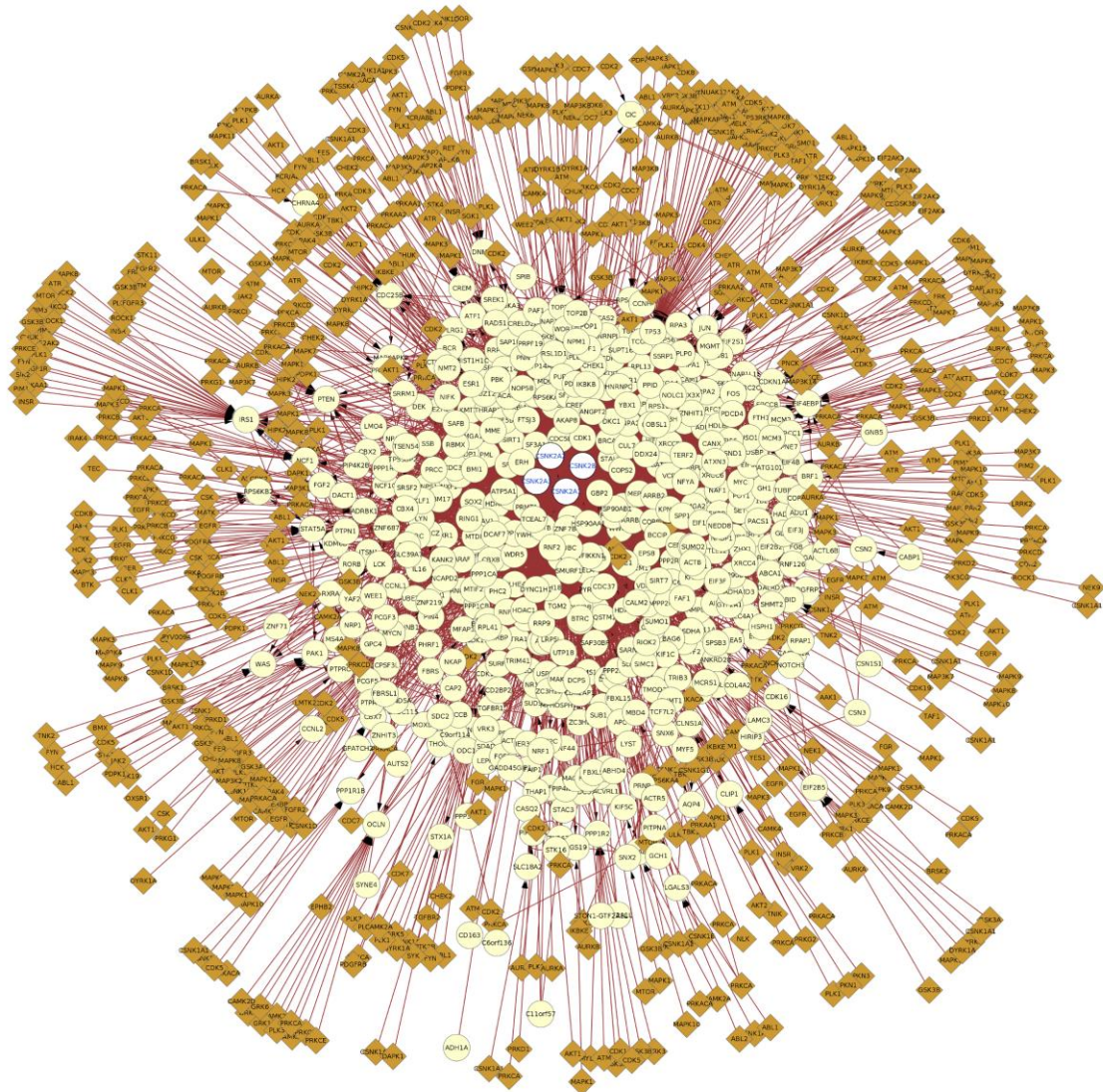
The quinalizarin phosphoproteomic study [102] identified 28 downregulated putative CK2 phosphosites with several of the target proteins displaying a role in cell death and/or survival including TPD52 (isoform 2), STX12, BCLAF1, AKAP12, RAD50, and PDCD4 (isoform 2). Overall, most of the phosphorylation-modulated substrates were classified as nuclear and were found to be involved in biological processes classified as transcription, mRNA and rRNA processing, gene expression, and DNA replication. Comparable functional annotations were obtained in two of the other CK2 phosphoproteomic studies where the phosphosites identified to belong to proteins mostly localized to the nucleus as components of the spliceosome [40,41].

Intriguingly, the quinalizarin phosphoproteomic study also revealed that several “CK2 attributable” phosphosites increased upon treatment with the inhibitor. As this result seems paradoxical, the authors proposed both technical and biological explanations for this observation including off-target effects of the inhibitor and an increase in the protein abundance of those substrates in particular [102]. Since it is evident that CK2 is connected to a plethora of regulatory hubs through protein-protein interaction and hierarchical phosphorylation, we performed a kinase-motif matching analysis using the PhosphoMotif Finder functionality of the HPRD database [103] to determine if other kinases could, in theory, be responsible for the upregulation of phosphosites that had been putatively identified as CK2-dependent phosphosites. As a result of this analysis, we identified at least two instances where the modulated phosphosite matched motifs for other kinases besides CK2. For example, the vicinity of the residue S1068 of TP53BP1 matches the minimal CK2 consensus sequence pSXX[E/D] as well as the pSQ and XpSQ substrate motifs of the ATM kinase and the DNA-dependent protein kinase, respectively. Interestingly, TP53BP1 does interact with ATM, which phosphorylates several residues in the protein upon DNA damage to promote its tumor suppressor functions [104]. Another example is AKAP12 where the vicinity of S627 matches substrate recognition

motifs pSXX[E/D], RXRXX[pS/pT], [R/K]XRXXpS, RVRRPpSESDK, and RRPpS conforming to CK2, AKT, MAPKAPK1, AMP-activated protein kinase 2, and PKA/PKC motifs, respectively. AKAP12 has been shown to interact at least with PKC and PKA [105], with PKA phosphorylating S627 and three other residues [106] of the protein. Moreover, in proteomic databases, arginine and lysine residues proximal to S627 of AKAP12, are reported to be methylated [98]. Altogether, these observations illustrate the complexity of CK2-signaling networks and the challenges associated with interpreting changes in the phosphoproteome arising from the modulation of CK2. Accordingly, all possible sources of information for other kinases and modifying enzymes that may act upon CK2 target sequences (e.g., arginine-methyltransferases and lysine-acetyltransferases) need to be taken into consideration.

## 1.4 Extension of CK2 Networks to Include Other Constituents of Regulatory Networks

As a logical extension of examining direct interactions with CK2, understanding how CK2 integrates with other signaling networks in the cell requires consideration of how CK2 substrates may be acted upon by other constituents of signaling networks. For example, we have considered phosphorylation information regarding CK2 substrates with other kinases that also modify the CK2 substrates and/or interactors (Fig. 1.4). In this respect, Fig. 1.4 shows that at least 171 other kinases (Supp. Table 1.6) are capable of phosphorylating sites in CK2 interactors, which suggests a likelihood of functional interplay among phosphorylation sites. This is reflected, for example, in the fact that CK2-dependent phosphorylation often participates in hierarchical phosphorylation with other kinases, which generates a CK2 target sequence with a phosphorylated serine that functions as the dominant specificity determinant [36].



**Figure 1.4.** Protein-protein interaction network of CK2 subunits expanded to include kinase information of the interactors retrieved from PhosphositePlus, the network was represented with Cytoscape v3.4.0 using the BisoGenet v3.0.0 plugin and the PhosphositePlus Web Service Client Module. The yellow nodes represent CK2 interactors and the brown nodes represent kinases that phosphorylate (edges with arrow) the interactors. The red edges represent phosphorylation events. The network was represented as explained in Fig. 1.2. Briefly, all the proteins were selected, and the kinase data was added by importing the information from PhosphoSitePlus using the plugin and selecting the gene name as the matching key. A high-resolution image of this figure is also available in the Appendix A.

As noted above, there is evidence that many CK2 substrates are also phosphorylated by other kinases. When considering the prospect for hierarchical phosphorylation involving CK2, analysis of the phosphorylation events in the vicinity of the CK2 phosphosites may

further contribute to identifying such relationships. Consequently, we searched PhosphoSitePlus database serine, threonine, and tyrosine phosphorylation and CK2 substrate data, mostly reflecting data generated through a proteomics approach [95], for those phosphosites occurring in the primary structure at different distance windows up- or downstream of the CK2 target site (Table 1.2, Supp. Table 1.6). This analysis was also extended to other post-translational modifications reported in the PhosphoSitePlus database, such as ubiquitination and acetylation (Table 1.2, Supp. Table 1.6). Although this analysis is restricted to CK2 substrate data available in PhosphoSitePlus, it is useful to illustrate the importance of considering the modification status of target sequences when studying CK2-dependent phosphorylation, a concept that also applies to other kinases. For example, the link between AKT-mediated phosphorylation and arginine methylation has been previously reported [107,108,109]. The post-translational modification analysis indicated that the vicinity of CK2 sites may constitute ‘hot spots’ for phosphorylation. Consideration of these sites brings together at least 76 different kinases including CDK7, STK1, PRKCA, PLK1, GSK3, CDC7, SRC, CDK2, MAPK3, GRK2, PRKCZ, CDK9, MAPK1, and IKBKE.

**Table 1.2.** Post-translational modifications (PTMs) in the vicinity of CK2<sup>1</sup> target sites retrieved from PhosphoSitePlus (accessed August 2016).

PTM type	# sites at -4/+4	# sites at -7/+7	# sites at -36/+36
Acetylation	15	25	146
O-N-acetylgalactosamine	2; overlap: 1	5; overlap: 1	19; overlap: 1
O-N-acetylglucosamine	1; overlap: 2	1; overlap: 2	5; overlap: 2
Methylation (m1,m2,m3,me)	4, 3, 1, -	8, 7, 1, -	29, 22, 2, 3
Phosphorylation	262; overlap: 482	395; overlap: 482	1177; overlap: 482
Sumoylation	1	8	50
Ubiquitination	11	21	171

\* CK2 sites were retrieved for bovine CSNK2A1 (P68399), human CSNK2A1 (P68400), human CSNK2A2 (P19784), human CSNK2B (P67870), mouse CSNK2A1 (Q60737), rat CSNK2A1 (P19139). Overlap: Modification occurring at the CK2 target site.

## 1.5 Rationale for the Study

Taken together, systematic analysis of the literature and databases strongly reinforces the view that CK2 is involved in a broad spectrum of biological processes. At the same time, it is important to recognize that there are significant limitations with information that is available both within the peer-reviewed literature and in databases. In addition to clearly highlighting the involvement of CK2 in a broad array of biological processes, database mining and network analysis clearly reveals the extensive interplay between CK2 and other constituents of signaling networks. This interplay is evident from the identification of ‘hot spots’ where CK2 phosphorylation sites are localized proximal to other phosphorylation sites or other post-translational modifications within the primary sequence of its substrates. The proximity of post-translational modifications to one another raises the very interesting prospect of one post-translational modification being regulated by others. As noted earlier, we have previously demonstrated that CK2 can modulate the cleavage of caspase substrates when CK2 phosphorylates residues adjacent to the cleavage site [48]. This is a clear example of how phosphorylation by CK2 can modulate susceptibility to another post-translational modification. The demonstration that phosphorylated residues can sometimes ‘prime’ a substrate for hierarchical phosphorylation by CK2 demonstrates that phosphorylation by CK2 can also be regulated by other modifications [36]. Considering the prevalence of post-translational modifications that reside within CK2 substrates within close proximity of the CK2 phosphorylation site (Table 1.2), it will be important to consider the relationship between CK2 phosphorylation and these other modifications. From the perspective of its participation in regulatory processes, the interplay between CK2 and other pathways could yield very intricate and precise control of processes. Considering the emergence of CK2 as a potential therapeutic target, the intricate relationships of CK2 within the regulatory networks also have important implications for the application of CK2 inhibitors. From this perspective, the modulation of CK2 could impact other pathways since phosphorylation by CK2 could have regulatory consequences for other pathways. Similarly, modulation of other pathways could also affect CK2 or its functions as the modification status of the target sequence may ultimately regulate its recognition by CK2.

## 1.6 Scope of Thesis

Given the intricacies of CK2 signaling and its integration with other modification networks in the cell beyond phosphorylation, we set out to decipher CK2's functional relationships with such networks. For this the following hypothesis was tested: PTMs, in particular acetyllysine, occurring in the vicinity of the phosphorylatable residue modulates phosphorylation of substrates by protein kinase CK2. Correspondingly, in Chapter 2 *in vitro*, *in silico*, and cell experiments are presented that determine the CK2 specificity against peptide substrates containing lysine or acetyllysine residues, the extent of modification and acetylation interplay at candidate CK2 target sites containing lysine or acetyllysine, the cellular and functional boundaries of the candidate hits, the CK2-dependent phosphoproteome upon modulation of the acetylome, and the acetylome and phosphoproteome overlap as pairs of co-occurring sites. Interestingly, lysine and acetyllysine were found to be tolerated by CK2 when present in position +2 from the phosphorylatable residue and at positions +1 and +3, in a highly negative sequence environment. A list of candidate sites was generated including known CK2 target sites, previously annotated sites in PTM databases and the literature, and known regulatory sites which can be validated in further studies. But equally important, experimental evidence was gathered of the co-occurrence of lysine acetylation and phosphorylation in the cell at many modified peptide instances including tentative CK2 sites and at sites that are frequently observed by mass spectrometry(MS)-based proteomics and annotated in databases. The latter includes sites in writers, erasers, and/or readers of acetylation, phosphorylation, and methylation and components of cellular complexes involved in transcription repression and chromatin remodeling among other functions. Given that CK2 is known to phosphorylate and as a result regulate the activity of several writers of acetylation in the cell and that a number of these are part of the transcriptional repression and chromatin remodeling machinery the link between CK2 and chromatin organization was summarized and expanded upon by integrating literature mining with phosphoproteomic and molecular interaction data in Chapter 3.

Finally, the experiments performed required the processing and analysis of large data sets; the individual scripts used for this were integrated into an analytical framework and



used to develop a data-driven R Shiny web application called visualRepo. The development process and the features of visualRepo are covered in Chapter 4 along with strategies for the functional analysis of lists of modification sites applied in Chapter 2 and coded into visualRepo. Importantly, visualRepo can be applied and extended if necessary, for exploring the association between modification sites and other modification writers and erasers of interest.

## 1.7 References

1. Litchfield, D.W. Protein kinase CK2: Structure, regulation and role in cellular decisions of life and death. *Biochem. J.* **2003**, 369, 1–15. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
2. Cabrejos, M.E.; Allende, C.C.; Maldonado, E. Effects of phosphorylation by protein kinase CK2 on the human basal components of the RNA polymerase II transcription machinery. *J. Cell. Biochem.* **2004**, 93, 2–10. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
3. Lüscher, B.; Christenson, E.; Litchfield, D.W.; Krebs, E.G.; Eisenman, R.N. Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation. *Nature* **1990**, 344, 517–522. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
4. Szebeni, A.; Hingorani, K.; Negi, S.; Olson, M.O.J. Role of protein kinase CK2 phosphorylation in the molecular chaperone activity of nucleolar protein b23. *J. Biol. Chem.* **2003**, 278, 9107–9115. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
5. Borgo, C.; Franchin, C.; Salizzato, V.; Cesaro, L.; Arrigoni, G.; Matricardi, L.; Pinna, L.A.; Donella-Deana, A. Protein kinase CK2 potentiates translation efficiency by phosphorylating eIF3j at Ser127. *Biochim. Biophys. Acta - Mol. Cell Res.* **2015**, 1853, 1693–1701. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
6. Riera, M.; Roher, N.; Miró, F.; Gil, C.; Trujillo, R.; Aguilera, J.; Plana, M.; Itarte, E. Association of protein kinase CK2 with eukaryotic translation initiation factor eIF-2 and with grp94/endoplasmic reticulum chaperone. *Mol. Cell. Biochem.* **1999**, 191, 97–104. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
7. Gandin, V.; Masvidal, L.; Cargnello, M.; Gyenis, L.; McLaughlan, S.; Cai, Y.; Tenkerian, C.; Morita, M.; Balanathan, P.; Jean-Jean, O.; et al. mTORC1 and CK2 coordinate ternary and eIF4F complex assembly. *Nat. Commun.* **2016**, 7, 11127. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)][[Green Version](#)]
8. Niechi, I.; Silva, E.; Cabello, P.; Huerta, H.; Carrasco, V.; Villar, P.; Cataldo, L.R.; Marcelain, K.; Armisen, R.; Varas-Godoy, M.; Fernandez, C.; et al. Colon cancer cell invasion is promoted by protein kinase CK2 through increase of endothelin-converting enzyme-1c protein stability. *Oncotarget* **2015**, 6, 42749–42760. [[Google Scholar](#)] [[PubMed](#)]
9. Patsoukis, N.; Li, L.; Sari, D.; Petkova, V.; Boussiotis, V.A. PD-1 Increases PTEN Phosphatase Activity While Decreasing PTEN Protein Stability by Inhibiting Casein Kinase 2. *Mol. Cell. Biol.* **2013**, 33, 3091–3098. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]

10. Zhang, C.; Vilks, G.; Canton, D.A.; Litchfield, D.W. Phosphorylation regulates the stability of the regulatory CK2beta subunit. *Oncogene* **2002**, *21*, 3754–3764. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
11. Shen, J.; Channavajhala, P.; Seldin, D.C.; Sonenshein, G.E. Phosphorylation by the protein kinase CK2 promotes calpain-mediated degradation of IkappaBalpha. *J. Immunol.* **2001**, *167*, 4919–4925. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
12. Scaglioni, P.P.; Yung, T.M.; Choi, S.C.; Baldini, C.; Konstantinidou, G.; Pandolfi, P.P.; Pandolfi, P.P. CK2 mediates phosphorylation and ubiquitin-mediated degradation of the PML tumor suppressor. *Mol. Cell. Biochem.* **2008**, *316*, 149–154. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
13. Homma, M.K.; Homma, Y. Cell cycle and activation of CK2. *Mol. Cell. Biochem.* **2008**, *316*, 49–55. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
14. Ahmed, K.; Gerber, D.A.; Cochet, C. Joining the cell survival squad: An emerging role for protein kinase CK2. *Trends Cell Biol.* **2002**, *12*, 226–230. [[Google Scholar](#)] [[CrossRef](#)]
15. Piazza, F.A.; Ruzzene, M.; Gurrieri, C.; Montini, B.; Bonanni, L.; Chioetto, G.; Di Maira, G.; Barbon, F.; Cabrelle, A.; Zambello, R.; et al. Multiple myeloma cell survival relies on high activity of protein kinase CK2. *Blood* **2006**, *108*, 1698–1707. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
16. Duncan, J.S.; Turowec, J.P.; Duncan, K.E.; Vilks, G.; Wu, C.; Luscher, B.; Li, S.S.-C.; Gloor, G.B.; Litchfield, D.W. A Peptide-Based Target Screen Implicates the Protein Kinase CK2 in the Global Regulation of Caspase Signaling. *Sci. Signal.* **2011**, *4*, ra30. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
17. Tsuchiya, Y.; Akashi, M.; Matsuda, M.; Goto, K.; Miyata, Y.; Node, K.; Nishida, E. Involvement of the Protein Kinase CK2 in the Regulation of Mammalian Circadian Rhythms. *Sci. Signal.* **2009**, *2*, ra26. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
18. Trembley, J.H.; Wang, G.; Unger, G.; Slaton, J.; Ahmed, K. Protein Kinase CK2 in Health and Disease: CK2: A key player in cancer biology. *Cell. Mol. Life Sci.* **2009**, *66*, 1858–1867. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
19. Ortega, C.E.; Seidner, Y.; Dominguez, I. Mining CK2 in Cancer. *PLoS ONE* **2014**, *9*, e115609. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
20. Schuck, S.; Ruse, C.; Stenlund, A. CK2 Phosphorylation Inactivates DNA Binding by the Papillomavirus E1 and E2 Proteins. *J. Virol.* **2013**, *87*, 7668–7679. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
21. Marin, O.; Sarno, S.; Boschetti, M.; Pagano, M.A.; Meggio, F.; Ciminale, V.; D’Agostino, D.M.; Pinna, L.A. Unique features of HIV-1 Rev protein phosphorylation by protein kinase CK2 (‘casein kinase-2’). *FEBS Lett.* **2000**, *481*, 63–67. [[Google Scholar](#)] [[CrossRef](#)]
22. Ohtsuki, K.; Maekawa, T.; Harada, S.; Karino, A.; Morikawa, Y.; Ito, M. Biochemical characterization of HIV-1 Rev as a potent activator of casein kinase II in vitro. *FEBS Lett.* **1998**, *428*, 235–240. [[Google Scholar](#)] [[CrossRef](#)]
23. Kim, J.; Kim, S.H. Druggability of the CK2 inhibitor CX-4945 as an anticancer drug and beyond. *Arch. Pharm. Res.* **2012**, *35*, 1293–1296. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]

24. Chon, H.J.; Bae, K.J.; Lee, Y.; Kim, J. The casein kinase 2 inhibitor, CX-4945, as an anti-cancer drug in treatment of human hematological malignancies. *Front. Pharmacol.* **2015**, *6*, 70. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
25. Perea, S.E.; Reyes, O.; Baladron, I.; Perera, Y.; Farina, H.; Gil, J.; Rodriguez, A.; Bacardi, D.; Marcelo, J.L.; Cosme, K.; et al. CIGB-300, a novel proapoptotic peptide that impairs the CK2 phosphorylation and exhibits anticancer properties both in vitro and in vivo. *Mol. Cell. Biochem.* **2008**, *316*, 163–167. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
26. Benavent Acero, F.; Capobianco, C.S.; Garona, J.; Cirigliano, S.M.; Perera, Y.; Urtreger, A.J.; Perea, S.E.; Alonso, D.F.; Farina, H.G. CIGB-300, an anti-CK2 peptide, inhibits angiogenesis, tumor cell invasion and metastasis in lung cancer models. *Lung Cancer* **2016**. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
27. Martins, L.R.; Lúcio, P.; Melão, A.; Antunes, I.; Cardoso, B.A.; Stansfield, R.; Bertilaccio, M.T.S.; Ghia, P.; Drygin, D.; Silva, M.G.; et al. Activity of the clinical-stage CK2-specific inhibitor CX-4945 against chronic lymphocytic leukemia. *Leukemia* **2014**, *28*, 179–182. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
28. Martins, L.R.; Perera, Y.; Lúcio, P.; Silva, M.G.; Perea, S.E.; Barata, J.T. Targeting chronic lymphocytic leukemia using CIGB-300, a clinical-stage CK2-specific cell-permeable peptide inhibitor. *Oncotarget* **2014**, *5*, 258–263. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
29. Perea, S.E.; Baladron, I.; Garcia, Y.; Perera, Y.; Lopez, A.; Soriano, J.L.; Batista, N.; Palau, A.; Hernández, I.; Farina, H.; et al. CIGB-300, a synthetic peptide-based drug that targets the CK2 phosphoacceptor domain. Translational and clinical research. *Mol. Cell. Biochem.* **2011**, *356*, 45–50. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
30. Niefind, K.; Raaf, J.; Issinger, O.-G. Protein Kinase CK2 in Health and Disease: Protein Kinase CK2: from structures to insights. *Cell. Mol. Life Sci.* **2009**, *66*, 1800–1816. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
31. Wilson, L.K.; Dhillon, N.; Thorner, J.; Martin, G.S. Casein kinase II catalyzes tyrosine phosphorylation of the yeast nucleolar immunophilin Fpr3. *J. Biol. Chem.* **1997**, *272*, 12961–12967. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
32. Donella-Deana, A.; Cesaro, L.; Sarno, S.; Brunati, A.M.; Ruzzene, M.; Pinna, L.A. Autocatalytic tyrosine-phosphorylation of protein kinase CK2 alpha and alpha' subunits: implication of Tyr182. *Biochem. J.* **2001**, *357*, 563–567. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
33. Basnet, H.; Su, X.B.; Tan, Y.; Meisenhelder, J.; Merkurjev, D.; Ohgi, K.A.; Hunter, T.; Pillus, L.; Rosenfeld, M.G. Tyrosine phosphorylation of histone H2A by CK2 regulates transcriptional elongation. *Nature* **2014**, *516*, 267–271. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
34. Vilck, G.; Weber, J.E.; Turowec, J.P.; Duncan, J.S.; Wu, C.; Derksen, D.R.; Zien, P.; Sarno, S.; Donella-Deana, A.; Lajoie, G.; et al. Protein kinase CK2 catalyzes tyrosine phosphorylation in mammalian cells. *Cell. Signal.* **2008**, *20*, 1942–1951. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
35. Marin, O.; Meggio, F.; Draetta, G.; Pinna, L.A. The consensus sequences for cdc2 kinase and for casein kinase-2 are mutually incompatible. A study with peptides

- derived from the beta-subunit of casein kinase-2. *FEBS Lett.* **1992**, 301, 111–114. [[Google Scholar](#)] [[CrossRef](#)]
36. St-Denis, N.; Gabriel, M.; Turowec, J.P.; Gloor, G.B.; Li, S.S.-C.; Gingras, A.-C.; Litchfield, D.W. Systematic investigation of hierarchical phosphorylation by protein kinase CK2. *J. Proteomics* **2015**, 118, 49–62. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  37. Litchfield, D.W.; Arendt, A.; Lozeman, F.J.; Krebs, E.G.; Hargrave, P.A.; Palczewski, K. Synthetic phosphopeptides are substrates for casein kinase II. *FEBS Lett.* **1990**, 261, 117–120. [[Google Scholar](#)] [[CrossRef](#)]
  38. Meggio, F.; Pinna, L.A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **2003**, 17, 349–368. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  39. Nuñez de Villavicencio-Díaz, T.; Mazola, Y.; Yasser, P.; Cruz, Y.; Guirola-Cruz, O.; Perea, S.E. Predicting CK2 beta-dependent substrates using linear patterns. *Rep. Biochem. Biophys.* **2015**, 25, 20–27. [[Google Scholar](#)] [[CrossRef](#)]
  40. Wang, C.; Ye, M.; Bian, Y.; Liu, F.; Cheng, K.; Dong, M.; Dong, J.; Zou, H. Determination of CK2 Specificity and Substrates by Proteome-Derived Peptide Libraries. *J. Proteome Res.* **2013**, 12, 3813–3821. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  41. Bian, Y.; Ye, M.; Wang, C.; Cheng, K.; Song, C.; Dong, M.; Pan, Y.; Qin, H.; Zou, H. Global screening of CK2 kinase substrates by an integrated phosphoproteomics workflow. *Sci. Rep.* **2013**, 3, 3460. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  42. Salvi, M.; Sarno, S.; Cesaro, L.; Nakamura, H.; Pinna, L.A. Extraordinary pleiotropy of protein kinase CK2 revealed by weblogo phosphoproteome analysis. *Biochim. Biophys. Acta* **2009**, 1793, 847–859. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  43. Sarno, S.; Ghisellini, P.; Pinna, L.A. Unique activation mechanism of protein kinase CK2. The N-terminal segment is essential for constitutive activity of the catalytic subunit but not of the holoenzyme. *J. Biol. Chem.* **2002**, 277, 22509–22514. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  44. Olsen, B.B.; Guerra, B.; Niefind, K.; Issinger, O.-G. Structural Basis of the Constitutive Activity of Protein Kinase CK2. *Methods Enzymol.* **2010**, 484, 515–529. [[Google Scholar](#)]
  45. Pinna, L.A. Protein kinase CK2: A challenge to canons. *J. Cell Sci.* **2002**, 115, 3873–3878. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  46. Olsten, M.E.K.; Weber, J.E.; Litchfield, D.W. CK2 interacting proteins: Emerging paradigms for CK2 regulation? *Mol. Cell. Biochem.* **2005**, 274, 115–124. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  47. Turowec, J.P.; Duncan, J.S.; French, A.C.; Gyenis, L.; St Denis, N.A.; Vilks, G.; Litchfield, D.W. Protein kinase CK2 is a constitutively active enzyme that promotes cell survival: Strategies to identify CK2 substrates and manipulate its activity in mammalian cells. *Methods Enzymol.* **2010**, 484, 471–493. [[Google Scholar](#)] [[PubMed](#)]
  48. Turowec, J.P.; Duncan, J.S.; Gloor, G.B.; Litchfield, D.W. Regulation of caspase pathways by protein kinase CK2: identification of proteins with overlapping CK2

- and caspase consensus motifs. *Mol. Cell. Biochem.* **2011**, 356, 159–167. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
49. Turowec, J.P.; Vilks, G.; Gabriel, M.; Litchfield, D.W. Characterizing the convergence of protein kinase CK2 and caspase-3 reveals isoform-specific phosphorylation of caspase-3 by CK2 $\alpha'$ : Implications for pathological roles of CK2 in promoting cancer cell survival. *Oncotarget* **2013**, 4, 560–571. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  50. Duncan, J.S.; Turowec, J.P.; Vilks, G.; Li, S.S.C.; Gloor, G.B.; Litchfield, D.W. Regulation of cell proliferation and survival: Convergence of protein kinases and caspases. *Biochim. Biophys. Acta - Proteins Proteomics* **2010**, 1804, 505–510. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  51. Lu, Z. PubMed and beyond: A survey of web tools for searching biomedical literature. *Database* **2011**, 2011, baq036. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  52. Doms, A.; Schroeder, M. GoPubMed: Exploring PubMed with the Gene Ontology. *Nucleic Acids Res.* **2005**, 33, W783–W786. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  53. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.; Roth, A.; Bork, P.; et al. The STRING database in 2017: Quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **2017**, 45, D362–D368. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  54. McKendrick, L.; Milne, D.; Meek, D. Protein kinase CK2-dependent regulation of p53 function: Evidence that the phosphorylation status of the serine 386 (CK2) site of p53 is constitutive and stable. *Mol. Cell. Biochem.* **1999**, 191, 187–199. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  55. Khan, D.H.; He, S.; Yu, J.; Winter, S.; Cao, W.; Seiser, C.; Davie, J.R. Protein Kinase CK2 Regulates the Dimerization of Histone Deacetylase 1 (HDAC1) and HDAC2 during Mitosis. *J. Biol. Chem.* **2013**, 288, 16518–16528. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  56. Tsai, S.-C.; Seto, E. Regulation of histone deacetylase 2 by protein kinase CK2. *J. Biol. Chem.* **2002**, 277, 31826–31833. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  57. Dominguez, I.; Sonenshein, G.E.; Seldin, D.C. Protein kinase CK2 in health and disease: CK2 and its role in Wnt and NF-kappaB signaling: Linking development and cancer. *Cell. Mol. Life Sci.* **2009**, 66, 1850–1857. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  58. Futreal, P.A.; Coin, L.; Marshall, M.; Down, T.; Hubbard, T.; Wooster, R.; Rahman, N.; Stratton, M.R. A census of human cancer genes. *Nat. Rev. Cancer* **2004**, 4, 177–183. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  59. Chatr-aryamontri, A.; Oughtred, R.; Boucher, L.; Rust, J.; Chang, C.; Kolas, N.K.; O'Donnell, L.; Oster, S.; Theesfeld, C.; Sellam, A.; et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **2016**, gkw1102. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
  60. Killcoyne, S.; Carter, G.W.; Smith, J.; Boyle, J. Cytoscape: A Community-Based Framework for Network Modeling. *Methods Mol. Biol.* **2009**, 563, 219–239. [[Google Scholar](#)]

61. Cline, M.S.; Smoot, M.; Cerami, E.; Kuchinsky, A.; Landys, N.; Workman, C.; Christmas, R.; Avila-Campilo, I.; Creech, M.; Gross, B.; et al. Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* **2007**, *2*, 2366–2382. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
62. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
63. Kanehisa, M.; Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
64. Nersisyan, L.; Samsonyan, R.; Arakelyan, A. CyKEGGParser: Tailoring KEGG pathways to fit into systems biology analysis workflows. *F1000Research* **2014**, *3*, 145. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
65. Martin, A.; Ochagavia, M.E.; Rabasa, L.C.; Miranda, J.; Fernandez-de-Cossio, J.; Bringas, R. BisoGenet: A new tool for gene network building, visualization and analysis. *BMC Bioinform.* **2010**, *11*, 91. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
66. Filhol, O.; Giacosa, S.; Wallez, Y.; Cochet, C. Protein kinase CK2 in breast cancer: The CK2 $\beta$  regulatory subunit takes center stage in epithelial plasticity. *Cell. Mol. Life Sci.* **2015**, *72*, 3305–3322. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
67. Bibby, A.C.; Litchfield, D.W. The multiple personalities of the regulatory subunit of protein kinase CK2: CK2 dependent and CK2 independent roles reveal a secret identity for CK2beta. *Int. J. Biol. Sci.* **2005**, *1*, 67–79. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
68. Vilk, G.; Saulnier, R.B.; St Pierre, R.; Litchfield, D.W. Inducible expression of protein kinase CK2 in mammalian cells. Evidence for functional specialization of CK2 isoforms. *J. Biol. Chem.* **1999**, *274*, 14406–14414. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
69. Messenger, M.M.; Saulnier, R.B.; Gilchrist, A.D.; Diamond, P.; Gorbsky, G.J.; Litchfield, D.W. Interactions between protein kinase CK2 and Pin1. Evidence for phosphorylation-dependent interactions. *J. Biol. Chem.* **2002**, *277*, 23054–23064. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
70. Bosc, D.G.; Graham, K.C.; Saulnier, R.B.; Zhang, C.; Prober, D.; Gietz, R.D.; Litchfield, D.W. Identification and characterization of CKIP-1, a novel pleckstrin homology domain-containing protein that interacts with protein kinase CK2. *J. Biol. Chem.* **2000**, *275*, 14295–14306. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
71. Arrigoni, G.; Pagano, M.A.; Sarno, S.; Cesaro, L.; James, P.; Pinna, L.A. Mass spectrometry analysis of a protein kinase CK2beta subunit interactome isolated from mouse brain by affinity chromatography. *J. Proteome Res.* **2008**, *7*, 990–1000. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
72. Guerra, B.; Siemer, S.; Boldyreff, B.; Issinger, O.G. Protein kinase CK2: Evidence for a protein kinase CK2beta subunit fraction, devoid of the catalytic CK2alpha subunit, in mouse brain and testicles. *FEBS Lett.* **1999**, *462*, 353–357. [[Google Scholar](#)] [[CrossRef](#)]
73. Ruepp, A.; Waegele, B.; Lechner, M.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Mewes, H.-W. CORUM: The comprehensive

- resource of mammalian protein complexes--2009. *Nucleic Acids Res.* **2010**, 38, D497–D501. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
74. Villavicencio-Diaz, T.N.; Rodriguez-Ulloa, A.; Guirola-Cruz, O.; Perez-Riverol, Y. Bioinformatics tools for the functional interpretation of quantitative proteomics results. *Curr. Top. Med. Chem.* **2014**, 14, 435–449. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
75. Lou, D.Y.; Dominguez, I.; Toselli, P.; Landesman-Bollag, E.; O'Brien, C.; Seldin, D.C. The alpha catalytic subunit of protein kinase CK2 is required for mouse embryonic development. *Mol. Cell. Biol.* **2008**, 28, 131–139. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
76. Dominguez, I.; Degano, I.R.; Chea, K.; Cha, J.; Toselli, P.; Seldin, D.C. CK2 $\alpha$  is essential for embryonic morphogenesis. *Mol. Cell. Biochem.* **2011**, 356, 209–216. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
77. Dominguez, I.; Mizuno, J.; Wu, H.; Imbrie, G.A.; Symes, K.; Seldin, D.C. A role for CK2 $\alpha$ /beta in *Xenopus* early embryonic development. *Mol. Cell. Biochem.* **2005**, 274, 125–131. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
78. Bragdon, B.; Thinakaran, S.; Moseychuk, O.; King, D.; Young, K.; Litchfield, D.W.; Petersen, N.O.; Nohe, A. Casein Kinase 2  $\beta$ -Subunit Is a Regulator of Bone Morphogenetic Protein 2 Signaling. *Biophys. J.* **2010**, 99, 897–904. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
79. Liu, Y.; Holdbrooks, A.T.; De Sarno, P.; Rowse, A.L.; Yanagisawa, L.L.; McFarland, B.C.; Harrington, L.E.; Raman, C.; Sabbaj, S.; Benveniste, E.N.; et al. Therapeutic efficacy of suppressing the Jak/STAT pathway in multiple models of experimental autoimmune encephalomyelitis. *J. Immunol.* **2014**, 192, 59–72. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
80. Ulges, A.; Klein, M.; Reuter, S.; Gerlitzki, B.; Hoffmann, M.; Grebe, N.; Staudt, V.; Stergiou, N.; Bohn, T.; Brühl, T.-J.; et al. Protein kinase CK2 enables regulatory T cells to suppress excessive TH2 responses in vivo. *Nat. Immunol.* **2015**, 16, 267–275. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
81. Ampofo, E.; Rudzitis-Auth, J.; Dahmke, I.N.; Rössler, O.G.; Thiel, G.; Montenarh, M.; Menger, M.D.; Laschke, M.W. Inhibition of protein kinase CK2 suppresses tumor necrosis factor (TNF)- $\alpha$ -induced leukocyte-endothelial cell interaction. *Biochim. Biophys. Acta* **2015**, 1852, 2123–2136. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
82. Welker, S.; Götz, C.; Servas, C.; Laschke, M.W.; Menger, M.D.; Montenarh, M. Glucose regulates protein kinase CK2 in pancreatic  $\beta$ -cells and its interaction with PDX-1. *Int. J. Biochem. Cell Biol.* **2013**, 45, 2786–2795. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
83. Al Quobaili, F.; Montenarh, M. CK2 and the regulation of the carbohydrate metabolism. *Metabolism* **2012**, 61, 1512–1517. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
84. Lupp, S.; Götz, C.; Khadouma, S.; Horbach, T.; Dimova, E.Y.; Bohrer, A.-M.; Kietzmann, T.; Montenarh, M. The upstream stimulatory factor USF1 is regulated by protein kinase CK2 phosphorylation. *Cell. Signal.* **2014**, 26, 2809–2817. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]

85. Spohrer, S.; Dimova, E.Y.; Kietzmann, T.; Montenarh, M.; Götz, C. The nuclear fraction of protein kinase CK2 binds to the upstream stimulatory factors (USFs) in the absence of DNA. *Cell. Signal.* **2016**, *28*, 23–31. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
86. Zaman, M.S.; Johnson, A.J.; Bobek, G.; Kueh, S.; Kersaitis, C.; Bailey, T.D.; Buskila, Y.; Wu, M.J. Protein kinase CK2 regulates metal toxicity in neuronal cells. *Metallomics* **2016**, *8*, 82–90. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
87. Akkiraju, H.; Bonor, J.; Olli, K.; Bowen, C.; Bragdon, B.; Coombs, H.; Donahue, L.R.; Duncan, R.; Nohe, A. Systemic injection of CK2.3, a novel peptide acting downstream of bone morphogenetic protein receptor BMPRIa, leads to increased trabecular bone mass. *J. Orthop. Res.* **2015**, *33*, 208–215. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
88. Kahali, B.; Trott, R.; Paroush, Z.; Allada, R.; Bishop, C.P.; Bidwai, A.P. Drosophila CK2 phosphorylates Hairy and regulates its activity in vivo. *Biochem. Biophys. Res. Commun.* **2008**, *373*, 637–642. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
89. Kuntamalla, P.P.; Kunttas-Tatli, E.; Karandikar, U.; Bishop, C.P.; Bidwai, A.P. Drosophila protein kinase CK2 is rendered temperature-sensitive by mutations of highly conserved residues flanking the activation segment. *Mol. Cell. Biochem.* **2009**, *323*, 49–60. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
90. Ottaviani, D.; Marin, O.; Arrigoni, G.; Franchin, C.; Vilardell, J.; Sandre, M.; Li, W.; Parfitt, D.A.; Pinna, L.A.; Cheetham, M.E.; et al. Protein kinase CK2 modulates HSJ1 function through phosphorylation of the UIM2 domain. *Hum. Mol. Genet.* **2016**. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
91. Schwind, L.; Wilhelm, N.; Kartarius, S.; Montenarh, M.; Gorjup, E.; Götz, C. Protein kinase CK2 is necessary for the adipogenic differentiation of human mesenchymal stem cells. *Biochim. Biophys. Acta* **2015**, *1853*, 2207–2216. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
92. Herrmann, D.; Straubinger, M.; Hashemolhosseini, S. Protein kinase CK2 interacts at the neuromuscular synapse with Rapsyn, Rac1, 14-3-3 $\gamma$ , and Dok-7 proteins and phosphorylates the latter two. *J. Biol. Chem.* **2015**, *290*, 22370–22384. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
93. Cheusova, T.; Khan, M.A.; Schubert, S.W.; Gavin, A.-C.; Buchou, T.; Jacob, G.; Sticht, H.; Allende, J.; Boldyreff, B.; Brenner, H.R.; et al. Casein kinase 2-dependent serine phosphorylation of MuSK regulates acetylcholine receptor aggregation at the neuromuscular junction. *Genes Dev.* **2006**, *20*, 1800–1816. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
94. Hornbeck, P.V.; Kornhauser, J.M.; Tkachev, S.; Zhang, B.; Skrzypek, E.; Murray, B.; Latham, V.; Sullivan, M. PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **2012**, *40*, D261–D270. [[Google Scholar](#)] [[PubMed](#)]
95. Hornbeck, P.V.; Zhang, B.; Murray, B.; Kornhauser, J.M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **2015**, *43*, D512–D520. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]



96. Meng, R.; Al-Quobaili, F.; Müller, I.; Götz, C.; Thiel, G.; Montenarh, M. CK2 phosphorylation of Pdx-1 regulates its transcription factor activity. *Cell. Mol. Life Sci.* **2010**, *67*, 2481–2489. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
97. Venerando, A.; Franchin, C.; Cant, N.; Cozza, G.; Pagano, M.A.; Tosoni, K.; Al-Zahrani, A.; Arrigoni, G.; Ford, R.C.; Mehta, A.; et al. Detection of phospho-sites generated by protein kinase CK2 in CFTR: Mechanistic aspects of Thr1471 phosphorylation. *PLoS ONE* **2013**, *8*, e74232. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
98. Deutsch, E.W.; Csordas, A.; Sun, Z.; Jarnuczak, A.; Perez-Riverol, Y.; Ternent, T.; Campbell, D.S.; Bernal-Llinares, M.; Okuda, S.; Kawano, S.; et al. The ProteomeXchange consortium in 2017: Supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res.* **2017**, *45*, D1100–D1106. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
99. Franchin, C.; Salvi, M.; Arrigoni, G.; Pinna, L.A. Proteomics perturbations promoted by the protein kinase CK2 inhibitor quinalizarin. *Biochim. Biophys. Acta* **2015**, *1854*, 1676–1686. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
100. Rodríguez-Ulloa, A.; Ramos, Y.; Gil, J.; Perera, Y.; Castellanos-Serra, L.; García, Y.; Betancourt, L.; Besada, V.; González, L.J.; Fernández-de-Cossio, J.; et al. Proteomic profile regulated by the anticancer peptide CIGB-300 in non-small cell lung cancer (NSCLC) cells. *J. Proteome Res.* **2010**, *9*, 5473–5483. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
101. Le Bihan, T.; Hindle, M.; Martin, S.F.; Barrios-Llerena, M.E.; Krahmer, J.; Kis, K.; Millar, A.J.; van Ooijen, G. Label-free quantitative analysis of the casein kinase 2-responsive phosphoproteome of the marine minimal model species *Ostreococcus tauri*. *Proteomics* **2015**, *15*, 4135–4144. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
102. Franchin, C.; Cesaro, L.; Salvi, M.; Millionini, R.; Iori, E.; Cifani, P.; James, P.; Arrigoni, G.; Pinna, L. Quantitative analysis of a phosphoproteome readily altered by the protein kinase CK2 inhibitor quinalizarin in HEK-293T cells. *Biochim. Biophys. Acta* **2015**, *1854*, 609–623. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
103. Amanchy, R.; Periaswamy, B.; Mathivanan, S.; Reddy, R.; Tattikota, S.G.; Pandey, A. A curated compendium of phosphorylation motifs. *Nat. Biotechnol.* **2007**, *25*, 285–286. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
104. Jowsey, P.; Morrice, N.A.; Hastie, C.J.; McLauchlan, H.; Toth, R.; Rouse, J. Characterisation of the sites of DNA damage-induced 53BP1 phosphorylation catalysed by ATM and ATR. *DNA Repair. (Amst.)* **2007**, *6*, 1536–1544. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
105. Grove, B.D.; Bruchey, A.K. Intracellular distribution of gravin, a PKA and PKC binding protein, in vascular endothelial cells. *J. Vasc. Res.* **2001**, *38*, 163–175. [[Google Scholar](#)] [[CrossRef](#)]
106. Tao, J.; Wang, H.-Y.; Malbon, C.C. Protein kinase A regulates AKAP250 (gravin) scaffold binding to the beta2-adrenergic receptor. *EMBO J.* **2003**, *22*, 6419–6429. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
107. Rust, H.L.; Thompson, P.R. Kinase Consensus Sequences: A Breeding Ground for Crosstalk. *ACS Chem. Biol.* **2011**, *6*, 881–892. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]

108. Yamagata, K.; Daitoku, H.; Takahashi, Y.; Namiki, K.; Hisatake, K.; Kako, K.; Mukai, H.; Kasuya, Y.; Fukamizu, A. Arginine Methylation of FOXO Transcription Factors Inhibits Their Phosphorylation by Akt. *Mol. Cell* **2008**, *32*, 221–231. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
109. Sakamaki, J.-i.; Daitoku, H.; Ueno, K.; Hagiwara, A.; Yamagata, K.; Fukamizu, A. Arginine methylation of BCL-2 antagonist of cell death (BAD) counteracts its phosphorylation and inactivation by Akt. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6085–6090. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]

## 1.8 Supplemental Materials

The supplemental materials are provided within this thesis or as separate files whenever necessary due to their size (see Appendices).

**Supp. Table 1.1.** STRING network, function, and annotations of the top 50 CK2-related genes.

Sheet	Name	Content
1	STRING network	Top 50 related proteins according to STRING v10.0 analysis.
2	Protein function	Function of the CK2-related proteins downloaded from STRING v10.0 analysis.
3	Functional enrichment	Gene Ontology annotations (biological process, molecular function, and cellular component) and KEGG PATHWAY enrichment downloaded from STRING v10.0 analysis.
Appendix A		Supplemental Material – Chapter 1

**Supp. Table 1.2.** CK2-interacting proteins in the human interactome.

Sheet	Name	Content
1	Step 1 interactors of CK2 (direct)	Step 1 CK2 (CSNK2A1(GeneID:1457), CSNK2A2(GeneID:1459), CSNK2B(GeneID:1460)) interacting proteins retrieved from a human protein-protein interaction network built in Cytoscape from BioGRID data*
2	Step1, CK2 network	SIF format of the network in sheet#1
3	Step 2 interactors of CK2 (indirect)	Step 2 CK2 (CSNK2A1(GeneID:1457), CSNK2A2(GeneID:1459), CSNK2B(GeneID:1460)) interacting proteins retrieved from a human protein-protein interaction network built in Cytoscape from BioGRID data*
Appendix A		Supplemental Material – Chapter 1

**Supp. Table 1.3.** CK2-interacting hub proteins.

Sheet	Name	Content
1	CK2 hub interactors	Proteins that interact with CK2 and are highly connected in human protein-protein interaction networks.
Appendix A		Supplemental Material – Chapter 1

**Supp. Table 1.4.** CK2-interacting proteins of each subunit.

Sheet	Name	Content
1	CSNK2A1 interactors	CSNK2A1 interactors retrieved from SysBiomics database using BioGenet plugin of Cytoscape
2	CSNK2A2 interactors	CSNK2A2 interactors retrieved from SysBiomics database using BioGenet plugin of Cytoscape
3	CSNK2B interactors	CSNK2B interactors retrieved from SysBiomics database using BioGenet plugin of Cytoscape

4	Interactors intersection	Interactors intersection across the three subunits, includes the unique interactors
Appendix A		Supplemental Material – Chapter 1

**Supp. Table 1.5.** CK2-related GO annotations extracted from the literature and protein complex data.

Sheet	Name	Content
1	GO cellular component	Cellular components retrieved using GoPubmed text mining tool for the query: "Casein Kinase II"[mesh]*
2	GO biological process	Biological processes retrieved using GoPubmed text mining tool for the query: "Casein Kinase II"[mesh]*
3	GO molecular function	Molecular functions retrieved using GoPubmed text mining tool for the query: "Casein Kinase II"[mesh]*
4	CK2 complex	CORUM database information extracted for CSNK2A1, CSNK2A2, CSNK2B
Appendix A		Supplemental Material – Chapter 1

**Supp. Table 1.6.** PTM sites in the vicinity of CK2 phosphorylated sites.

Sheet	Name	Content
1	Kinases	Regulatory kinases of CK2 interactors and substrates extracted from PhosphoSitePlus.
2	Kinases, CK2 sites vicinity	Kinases that phosphorylate residues in the vicinity of CK2-phosphorylated sites. CK2 sites were retrieved for bovine CSNK2A1 (P68399), human CSNK2A1 (P68400), human CSNK2A2 (P19784), human CSNK2B (P67870), mouse CSNK2A1 (Q60737), rat CSNK2A1 (P19139).
3	PTM sites, CK2 sites vicinity	PTM sites in the vicinity of CK2-phosphorylated sites (-+7, +-36 distance window)
Appendix A		Supplemental Material – Chapter 1

**Supp. Table 1.7.** High-resolution images.

Content	
	High resolution images
Appendix A	
Supplemental Material – Chapter 1	

## Chapter 2

### 2 Protein Kinase CK2: Interplay Between Cellular Acetylation and Phosphorylation Networks.

#### **Proteoform diversity**

Protein CK2 target sites and residues in their vicinity can be targeted by different post-translational modification (PTM) types [1]. Furthermore, additional sources of biological variation exist that can modify the primary sequence of such sites, e.g., coding single-nucleotide polymorphisms (cSNP), mutations, and alternative splicing of RNA. These changes result from biological events and their biological relevance may not be fully grasped when considered as separate instances. This points to a need for studying CK2 targets, and those of other kinases, as proteoforms, i.e., “each individual molecular form of an expressed protein” [2]. Here, we were particularly interested in identifying the proteoforms of CK2 targets where other PTMs co-occur.

Briefly, the main sources of proteoform diversity at the DNA level are cSNPs and mutations whereas at the RNA level alternative splicing is the main source and to a lesser extent errors in translation. At the protein level, the modification of the lateral chain of amino acids is the main source of variation, which can occur co- and post-translational (i.e., PTM). An example of the complexity introduced by PTMs is the histone code, where several modifications co-occur at the tails of histones generating histone proteoforms with different functions. Other examples are the branching of monosaccharide and ubiquitin moieties attached to proteins. When combined in a single protein like the core histone H4, a study calculated that 13 of the most common PTMs and a single cSNP could give rise to more than 98,000 H4 proteoforms [2]. Considering that the number of PTM sites is between 0-90 modifications per protein the same study estimated that the number of proteoforms generated by any two modifications could be larger than  $1 \times 10^{27}$  [2]. However, such an explosion of proteoforms is not likely due to cellular regulation of protein copy number (1,000 copies of a proteins equals 1,000 proteoforms), abundance threshold (10,000 genes expressed at 100 copies per gene),

protein dynamic-range, cell type, spatial distribution, and enzymatic activity. A closer number based on database analyses is thought to be 6 million [3].

Notably, studying proteoforms remains a technological challenge, with proteoforms belonging to the same protein (proteoform families) being almost impossible to differentiate in some cases. In addition, proteomic techniques are starting to tackle the challenges of global PTM identification and discrimination of co-occurring PTMs [4]. Because of this, the number of proteoforms in each cell or cell population is a highly debated topic, with some researchers willing to consider the possibility of such high numbers and others considering that PTMs are mutually exclusive. However, given the functional diversity generated by proteoforms, it is essential to accept the fact that the primary sequence of a protein alone does not dictate its function and that protein networks in the cell are the results of functional associations between proteoform families.

### **PTM interplay and integration of signaling**

Biological systems are composed of modular and robust molecular networks that integrate and communicate information at different scales to create emergent properties. Such properties of complex systems change dynamically in response to the environment and cannot be predicted, even with an understanding of the parts alone [5]. Adopting this Systems Biology perspective posed a technological and scientific challenge and contributed to the development of OMICs and Bioinformatics.

The interaction of pathways to shape a specific response or biological process over time constitutes an example of an emergent property of systems [6]. At this level, PTM interplay, i.e., the directed positive/negative functional relations established between two modified amino acids in a protein, constitutes a pivotal regulatory mechanism of signal convergence [7]. A recent *in silico* study profiled global PTM coordination across protein-protein interaction networks (PPIN) highlighting around 400 “PTM spots” in 300 proteins [7]; this protein list contains hub proteins like P53 and CDK1, master regulators of signal transduction. Complementary studies have also been published, where PTM

interplay has been inferred based on amino acid co-evolution [8] and PTM conservation across functional domains [9].

Historically, PTM interplay has been associated with modifications targeting the tail of core histone proteins. Several models propose a histone “code”, which dictates what signaling complexes will be recruited to the DNA, modify the chromatin, and in turn gene expression [10–13]. However, this PTM interplay has also been described for non-histones proteins [14]. An example is the regulation of the stability of the tumor suppressor protein TP53 where different PTM combinations protect from or promote protein degradation, which in turn can dictate, for example, if a cell undergoes programmed cell death or not [15,16].

Briefly, interplaying PTMs can occur in the same or different proteins, cis- and trans-effect, respectively. Depending on the location of the sites in the protein the cis-effects can then be classified as adjacent or distant [10,17]. Functionally, the cis and trans-effects can be either direct or indirect; a direct effect involves only the PTMs whereas an indirect effect is mediated by other molecules through protein-protein interaction [10,17]. Another functional classification refers to the activating or inhibitory effect on one of the interplaying sites [10,17].

Recent studies by our group following a systems-approach have yielded important insights into the interplay between phosphorylation sites. The results obtained indicate a regulatory role of hierarchical phosphorylation in CK2-dependent phosphorylation such as priming of CK2-dependent phosphorylation by Proline-directed kinases [18]. Important insights were also obtained describing a regulatory role of CK2-dependent phosphorylation in modulating cell survival by attenuating caspase pathways through phosphorylation of residues adjacent to caspase cleavage sites [19–21] Altogether, these exemplify the intriguing connections between PTM interplay and kinase-dependent signaling.

Here, we continue the work of exploring the relevance of PTM interplay at CK2 target sites. We focus on describing a regulatory role between phosphorylation and lysine acetylation on CK2 signaling. As mentioned in the introductory part, lysine acetylation

can occur in neighboring residues of known CK2 target sites [1]. Lysine acetylation changes the local electrostatic properties of the sequence being acetylated by removing the positive charge of lysine [22]. Interestingly, positively charged residues, such as lysine, are considered negative determinants of CK2-dependent phosphorylation since CK2 acts upon acidic motifs in the substrate. Thus, we hypothesized that acetyllysine, occurring in the vicinity of the phosphorylatable residue modulates phosphorylation of substrates by protein kinase CK2. This functional association is intriguing since, like phosphorylation, lysine acetylation is dynamically maintained in the cell by well-characterized writers (lysine acetyltransferases) and erasers (lysine deacetylases) [22].

### **Global studies of PTMs**

Cell signaling involves the integration of different biological networks to adapt to external stimuli and perform cellular functions while maintaining homeostasis. Such networks are frequently found deregulated in cancer and other diseases, which has prompted extensive efforts to target their components. In this regard, proteomics has become an indispensable tool for understanding the impact of chemical probes/drugs that inhibit network effectors such as kinases and lysine deacetylases [23]. These studies provide invaluable information on the plasticity of cellular systems and how the phosphoproteome, acetylome, etc. dynamically respond and adapt to external perturbations. As mentioned before, an integrative view also demands an understanding of how such networks communicate to give rise to emergent properties of biological systems. Therefore, several studies aim to identify/predict PTM interplay by integrating the available proteomics data using a variety of bioinformatic tools and methods [9,14].

An extensive study by Scholz *et al.* published in 2015 profiled the changes in the acetylome of HeLa cells treated with 19 different deacetylase inhibitors, providing information on more than 8,000 acetylated sites [24]. Several such inhibitors are currently undergoing clinical trials [25]. The number of sites identified in this study compares to that of any given phosphoproteomic experiment [26] allowing a global assessment of the biological processes and subcellular compartments targeted by acetylation. For instance, it was found that treatment with nicotinamide, a pan-Sirtuin deacetylase inhibitor,

increased the acetylation of more than 500 sites belonging to proteins involved in transcription, RNA processing, gene expression, and chromatin modification, among others [24]. Other assessed compounds displayed high affinities for individual deacetylases causing the modulation of a specific set of biological processes, such as the regulation of microtubule polymerization by Tubacin and Bufexamac both of which are inhibitors of HDAC6 [24]. From the findings provided in this study, and others [27–29], it can be concluded that, just like phosphorylation, lysine acetylation has a global regulatory role.

By gathering information provided by acetylome and phosphoproteomic experiments, the literature, and PTM databases it comes as no surprise that an increasing number of proteins are being frequently reported as targets of different PTM types [1,30], much like histones and TP53. With the increasing flow of experimental information and advances in areas such as computational proteomics [4,31] integrative analysis of PTMs is becoming a more feasible strategy for uncovering proteome-wide PTM functional associations. In fact, a study by Grimes *et al.* in 2018 [32] used multiplexed proteomics to compare the phosphorylation, methylation, and acetylation patterns of 45 lung cancer cell lines to normal lung tissue and cell lines treated with anticancer drugs and integrated the PTM profiles obtained. The authors found that in proteins targeted by more than one PTM type the presence of those PTMs was inversely correlated, suggesting a pivotal functional role of this exclusivity in the functioning of signaling networks [32]. However, the latter is a highly debated topic. For instance, experimental evidence indicates that histone modifications can interplay positively and negatively. In fact, several interplay mechanisms have been observed (summarized in [12]): 1) competitive antagonism for modifications targeting the same residue, e.g., lysine acetylation, methylation, and ubiquitination; 2) dependence upon the addition of other modification in cis or trans, e.g., in yeast, methylation of H3K4 and H3K79 requires H2BK12 ubiquitination; and 3) cooperation between modifications, e.g., stronger binding of PHF8 to H3K4me3 when H3K9 and H3K14 are acetylated in cis.

As mentioned, treatment with deacetylase inhibitors increases the acetylation of hundreds of proteins in the cell [24,29]. Given the regulatory role of acetylation, a priori these



changes are likely to be sensed and acted upon by cellular signaling pathways of which kinases play prominent roles. Here, we make use of proteomics and bioinformatics to explore the connection between the acetyllysine and phospho-networks in U2OS (human osteosarcoma) and HeLa (cervix adenocarcinoma) cells treated with deacetylase inhibitors. Thus, the main goal of our study is to obtain experimental insights on the extent of the interplay between both networks by exploring the fraction of the phosphoproteome that is acetylated. It is anticipated that a study of this nature would likely provide information on an array of kinases and thus a wide range of signaling pathways. However, we focused on CK2-dependent signaling given its clinical relevance. Importantly, CK2 is a kinase with previous experimental connections to the regulation of the enzymatic activity of lysine deacetylases [33–37]. Regardless, the bioinformatics and experimental strategies devised in this study are of wide use and can be applied to understanding the functional association between any two types of modifications and/or regulatory protein of interest.

## 2.1 CK2-dependent signaling and lysine acetylation

Protein CK2 functional pleiotropy and the high frequency of the CK2 consensus motif throughout the proteome makes the CK2 literature broad and disperse. Given our hypothesis of a potential regulatory role for lysine acetylation proximal to a phosphorylation site matching the CK2 consensus, we initially looked for already established connections in the literature using the text mining tool Chilibot [38] and the Cytoscape v3.8.0 plugin stringApp v1.5.1 [39]. Using these tools, several instances were observed highlighting potential functional relationships between CK2 and lysine acetylation (Supp. Table 2.1, Supp. Fig. 2.1). In this regard, literature mining highlighted several functional links, for instance: CK2-SIRT1-FoxO3a acetylation in HCT116 and MCF 7 cells [40]; CK2-H3K56 acetylation in yeast [41]; CK2-Ikaros HDAC1 complex-JARID1B and H3K27 trimethylation and H3K9 acetylation [42]; CK2-P21-HDAC2 and KLF4 acetylation [37]; CK2-HDAC2 [43] (Supp. Table 2.1).

To further explore the extent of the association, the PubMed query: “ck2” AND “acetylation”, was graphically represented in a functional association STRING network (confidence cut-off score, i.e., the likelihood that a given interaction is biologically

meaningful, specific and reproducible, given the supporting evidence [39]): 0.7, maximum number of proteins: 100) followed by GO biological process enrichment analysis (Supp. Table 2.1, Supp. Fig. 2.1). As before, several meaningful associations were established i.e., 703 functional associations between the top 100 proteins (Supp. Table 2.1). Among the processes significantly enriched (FDR < 0.001) in the CK2 and acetylation search output (Supp. Table 2.1) we observed histone deacetylation (Supp. Fig. 2.1A), chromatin organization (Supp. Fig. 2.1B), and chromatin silencing (Supp. Fig. 2.1C). In this regard, CK2 has been shown to regulate several main effectors on these processes, such as erasers of lysine acetylation, including the lysine deacetylases HDAC1, HDAC2, HDAC3, and HDAC6 [33–37,44]. See Chapter 3 for a detailed summary of CK2's regulatory role in these processes.

Although experimental evidence positioned CK2 as a regulator of acetylation networks in the cell at the time of writing, no direct associations have been reported describing the role of neighboring cis-acting interplay between CK2-dependent phosphorylation and acetylation at the site level. However, an intriguing finding in a study of HDAC2 acetylation in response to hypertrophic stresses is that the substitution of the acetylation residue K75 by R75 in HDAC2 resulted in a significant decrease in CK2-dependent phosphorylation at Ser394 site [45]. However, it is not clear if this constitutes a direct or indirect interplay. Here, we explore the co-occurrence of acetylation and phosphorylation by CK2 in cis and the presence of acetylation sites in the phosphoproteome.

## 2.2 Tentative interplay between CK2-dependent phosphorylation and other PTM types

A total of 267 known human CK2 substrates from the PhosphoSitePlus (PSP) database [46] were found annotated in predicted within-protein (cis) PTM interplay data retrieved from the database PTMcode v2 [47] (Table 2.1). A total of 469 acetylation sites (known sites propagated to orthologs in Eukaryota) were predicted to interplay with the CK2 target sites (Table 2.1). The tentative interplay in the close vicinity of the CK2 substrates was determined by selecting either +/-4 and +/-7 sequence window (Table 2.1). Similar results to the analysis of known PTM sites in the vicinity of human CK2 substrates (see Chapter 1) were obtained regarding the number and types of PTM represented (Table

2.1). A limitation of the analysis is that PTMcode v2 data set is from 2014; since then many more modification sites have been added to the databases. However, not many PTM databases exist for PTM interplay. Another limitation is the incomplete annotation of CK2 substrates in PSP. Despite this limitation, since the number of CK2 substrates in PSP is in the hundreds, the dataset is sufficient to estimate the extent of the tentative interplay.

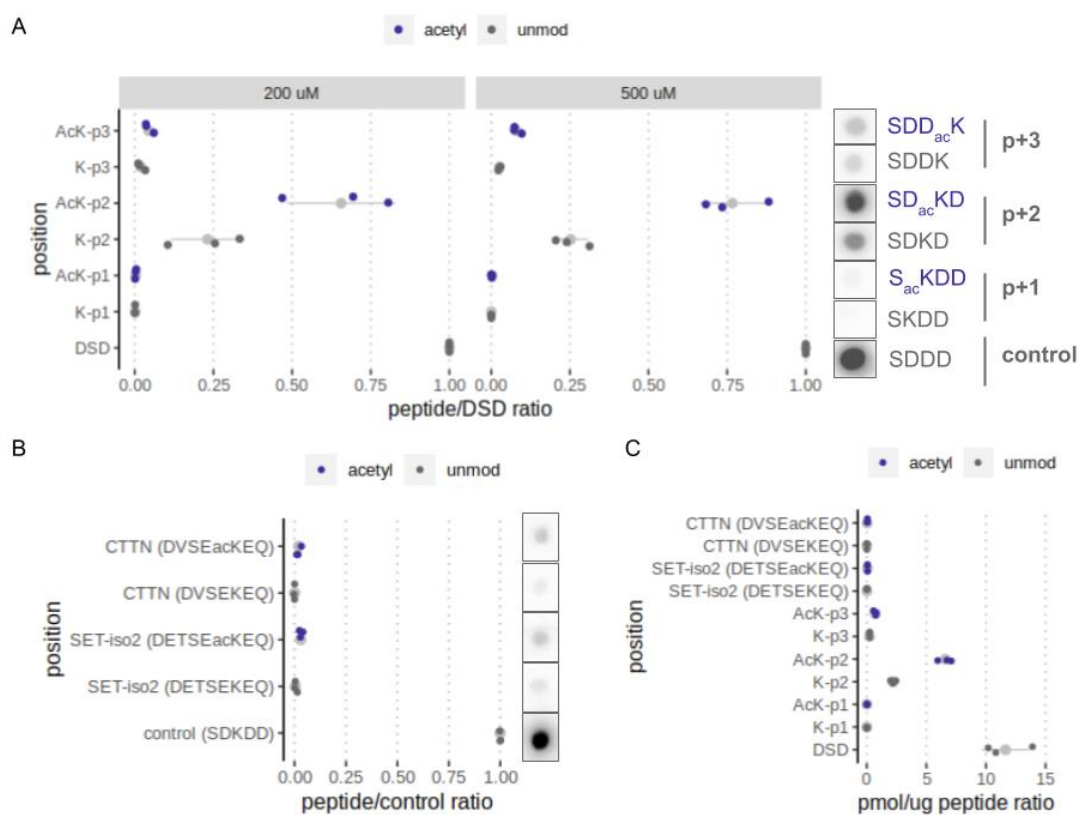
**Table 2.1.** PTM interplay annotated for known human CK2 substrates, retrieved from PSP, in the database PTMcode v2.

PTM type	No. interplaying sites	+/-4 amino acids	+/-7 amino acids
Phosphorylation	4836	202	292
Ubiquitination	536	1	5
Acetylation	469	2	10
Methylation	39	1	2
Sumoylation	39	-	2
Proteolytic Cleavage	27	-	-
N-linked Glycosylation	26	-	-
Nitrosylation	23	-	1
O-linked Glycosylation	19	2	5
Neddylation	14	-	1
O-glcnaC Glycosylation	14	2	3
Oxidation	8	-	1
Caspase Cleavage Aspartic Acid	7	-	-
O-galnaC Glycosylation	4	-	-
Hydroxylation	1	-	-
Myristoylation	1	1	1
Nitration	1	-	-

## 2.3 Phosphorylation of in solution peptides matching the CK2 motif with lysine and acetyllysine as determinants

In solution peptides derived from the CK2 peptide substrate DSD (RRRDDDSDDD) containing either lysine or acetyllysine at positions +1, +2, or +3, downstream of the

phosphoacceptor site, were phosphorylated *in vitro* by CK2 at two different concentrations (200 and 500  $\mu$ M) in triplicate. It was observed that both lysine and acetyllysine residues behaved as negative determinants at positions +1, completely abrogating phosphorylation compared to control (Fig. 2.1). The presence of lysine at position +3 also dramatically inhibited phosphorylation, with a mean value of 2.27 % phosphorylation compared to control (Fig. 2.1). Acetylation of lysine at position +3 showed modestly increased phosphorylation compared to the unmodified lysine variant, with a mean value of 6.26 % respect to control (Fig. 2.1). Overall, both lysine and acetyllysine behaved as negative determinants at positions +1 and +3. Conversely, acetyllysine, and lysine at position +2 rescued phosphorylation to 71.13 % and 24.19 % of the control, respectively (Fig. 2.1). This indicated that acetyllysine and lysine are tolerated for CK2-dependent phosphorylation at position +2, with acetylation of the lysine promoting phosphorylation as compared to the unmodified lysine. Peptide quality was assessed by MALDI (Supp. Fig. 2.2).



**Figure 2.1.** *In vitro* phosphorylation of peptides containing K and AcK (acetyllysine) by CK2. A) Phosphorylation of CK2 peptide substrate variants (200  $\mu$ M, 500  $\mu$ M);  $\gamma$ -<sup>32</sup>P-

ATP incorporation. B) Phosphorylation of endogenous sequences (500  $\mu\text{M}$ );  $\gamma\text{-}^{32}\text{P}\text{-ATP}$  incorporation. C) Phosphorylation stoichiometry in pmol of phosphate per  $\mu\text{g}$  of peptide. p: position, ac: acetylation, DSD: CK2 peptide substrate, grey point: mean.

Phosphorylation of in solution peptides representing two endogenous sequences containing lysine at position +2, Ser33 of CTTN (cortactin), and Thr23/Ser24 of SET-iso 2 (protein SET), was negligible compared to the peptide substrate containing lysine at +2 (Fig. 2.1). The phosphorylation of these endogenous sequences was increased by a modest degree by acetyllysine (Fig. 2.1). CTTN was also immunoprecipitated from U2OS cells but it was not possible to detect endogenous phosphorylation at the sites of interest by LC-MS/MS analysis [48] (Supp. Fig. 2.3).

Since studies with peptides revealed that lysine is tolerated and that acetylation of lysine could enhance CK2 phosphorylation when lysine is at the +2 position, we were interested in determining how common lysine appears in this position in the known CK2 target sites reported in the databases or in phosphoproteomic studies. Accordingly, linear patterns, referred to as K motif variants, were generated matching the sequences studied with in solution peptides (Table 2.2).

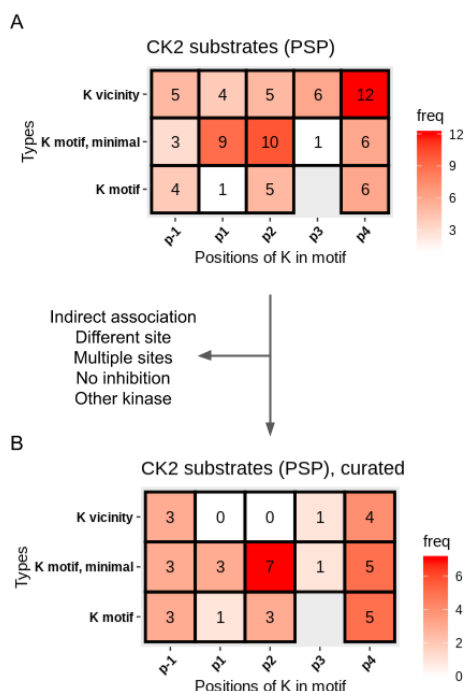
**Table 2.2.** Linear patterns (K motifs) matching the sequences studied in the in solution peptides.

Position	K motif variants		
	K motif	K motif, extended	K motif, minimal
p+1	[ST]K[DE][DE]	[ST]K[DE][DE][DE]	[ST]K.[DE]
p+2	[ST][DE]K[DE]	[ST][DE]K[DE][DE]	[ST]{P}K[DE]
p+3	[ST][DE][DE]K	[ST][DE][DE]K[DE]	[ST]{P}[DE]K

ScanProsite format: . Any amino acid. []: either amino acid. {}: any amino acid but, p: position.

The linear patterns were used for scanning the CK2 substrates annotated in the database PhosphoSitePlus (September 2019). After manual curation of the results several known CK2 targets matching the patterns (K motifs) were observed (Fig. 2.2, Supp. Table 2.2). Initially, a few instances of sites not matching the K motifs but containing Lys in the vicinity were identified. However, after manual curation most of these sites were

eliminated based on insufficient experimental information to confidently identify them as CK2 target sites (Fig. 2.2). The remaining instance containing Lys at position +3 but not matching the K motif or K motif minimal is an example of hierarchical phosphorylation pattern for CK2 (Supp. Table 2.2). Most of the studies describing the CK2 target sites identified as K motif hits apply *in vitro* phosphorylation and mutation of the amino acids of interest. The observed proportion of K motif variant hits was significantly different (Chi-squared test for given probabilities P value 0.015, Cohen's  $w = 0.7483$  (large effect)). The higher proportion of sites matching the K motif variants was observed for position +2 compared to +1 (Chi-squared test for given probabilities P value 0.1088, Cohen's  $w = 0.4286$  (medium effect)) and +3 (Chi-squared test for given probabilities P value 0.006656, Cohen's  $w = 0.8182$  (large effect)). The functions of the K motif variant hits were found to be diverse including transcription regulation, cytoskeleton organization, apoptosis, mRNA splicing, etc., (Supp. Table 2.2). The p-1 and p+4 positions were included to determine the number of K motif and K vicinity hits in residues that are not considered determinants of CK2 phosphorylation. After manual curation, a higher number of occurrences was observed for these positions compared to p+1 and p+3 and comparable to p+2 (Fig. 2.2, Supp. Table 2.2).



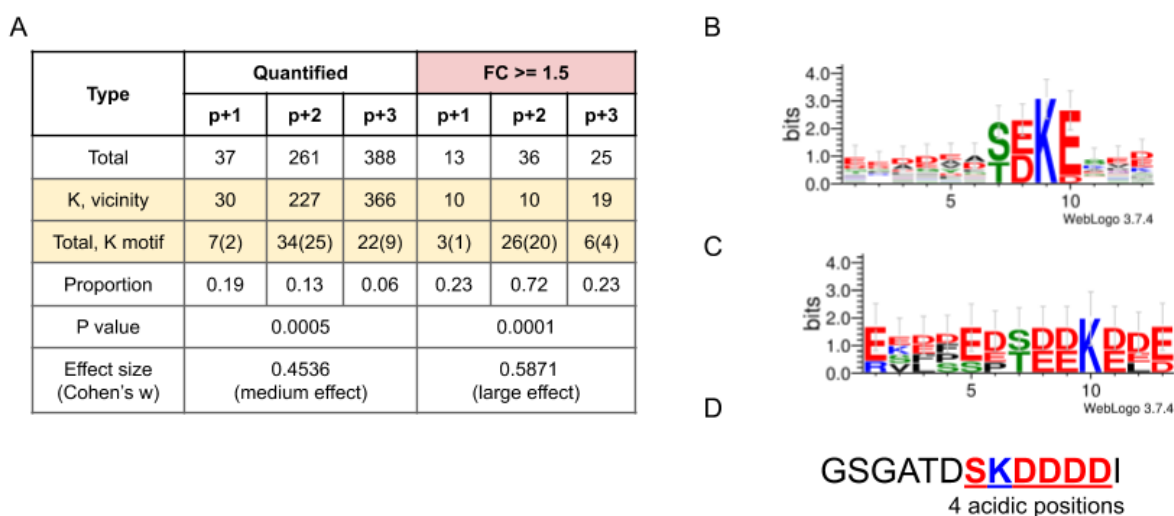
**Figure 2.2.** Known CK2 target sites in PhosphoSitePlus (PSP) containing Lys in the vicinity of the phosphorylatable site. A) Minimal and extended K motif hits or sites containing Lys (K) in the vicinity but not matching K motif variants. B) Hits after manual curation removing inconsistencies such as evidence where the site reported does not match the site of interest or no direct association to CK2. p: position.

## 2.4 Analysis of high-throughput datasets resulting from CK2 manipulation

### 2.4.1 *In silico* analysis of *in vitro* high-throughput CK2 kinase assay

Since lysine and acetyllysine at position +2 were tolerated by CK2-dependent phosphorylation and acetyllysine enhanced phosphorylation compared to lysine it was important to further investigate their relevance for CK2 phosphorylation events. The occurrence of phosphorylation at these K motif variants was further explored in three data sets obtained from a global screening of *in vitro* CK2 substrates by an integrated phosphoproteomic workflow [49]. As a result, phosphorylation by CK2 of sites matching all K motif variants was observed in the data sets analyzed (Supp. Table 2.3). In this regard, phosphorylation of Ser24 of SET-isoform 2 by CK2 was observed (Supp. Table 2.3).

The proportion of K motif hits among the sites containing Lys in the vicinity of a site phosphorylated by CK2 (fold change  $\geq 1.5$ ) was significantly different between variants (Fig. 2.3A), P value = 0.0001 (3-sample test for equality of proportions without continuity correction), Cohen  $w = 0.5871$ . The proportions were calculated after filtering ambiguous peptide hits (Supp. Table 2.3). As expected from the known CK2 substrate analysis, the observed proportion of +2 hits was larger than that of +1 (P value = 0.003, Cohen's  $w = 0.5157$ ) and +3 (P value = 0.0003, Cohen's  $w = 0.5012$ ) hits. These differences do not stem from the observed differences in the coverage of variants after quantification (P value = 0.0005, Cohen's  $w = 0.4536$ ) as a larger number of +3 hits were quantified compared to +2 (P value = 0.001, Cohen's  $w = 0.3935$ ) (Fig. 2.3A). Contrary to what was observed after filtering for CK2 substrates, fold change  $> 1.5$ , where the number of +2 K motif hits (26) was larger than that for +3 (6 sites) and +1 (3 sites) variants, in that order (Fig. 2.3A).



**Figure 2.3.** Sites phosphorylated *in vitro* by CK2 identified by Bian *et al.* 2013 [49] that contained Lys in the vicinity of the phosphorylatable residue. A) Total number of hits quantified in the workflow evaluation, Jurkat cells and HeLa cells data sets. CK2 sites, fold change (FC)  $> 1.5$ . P values were calculated using a 3-sample test for equality of proportions without continuity correction. In parenthesis hits for K motif extended. B) Weblogo representing the hits of K motif +2 variant. C) Weblogo representing the hits of K motif +3 variant. D) Sequence found phosphorylated for K motif +1 variant.

A weblogo representation of the K motif hits, not including minimal hits (Supp. Table 2.3), showed for the +2 variant an equal conservation of Asp and Glu residues at position

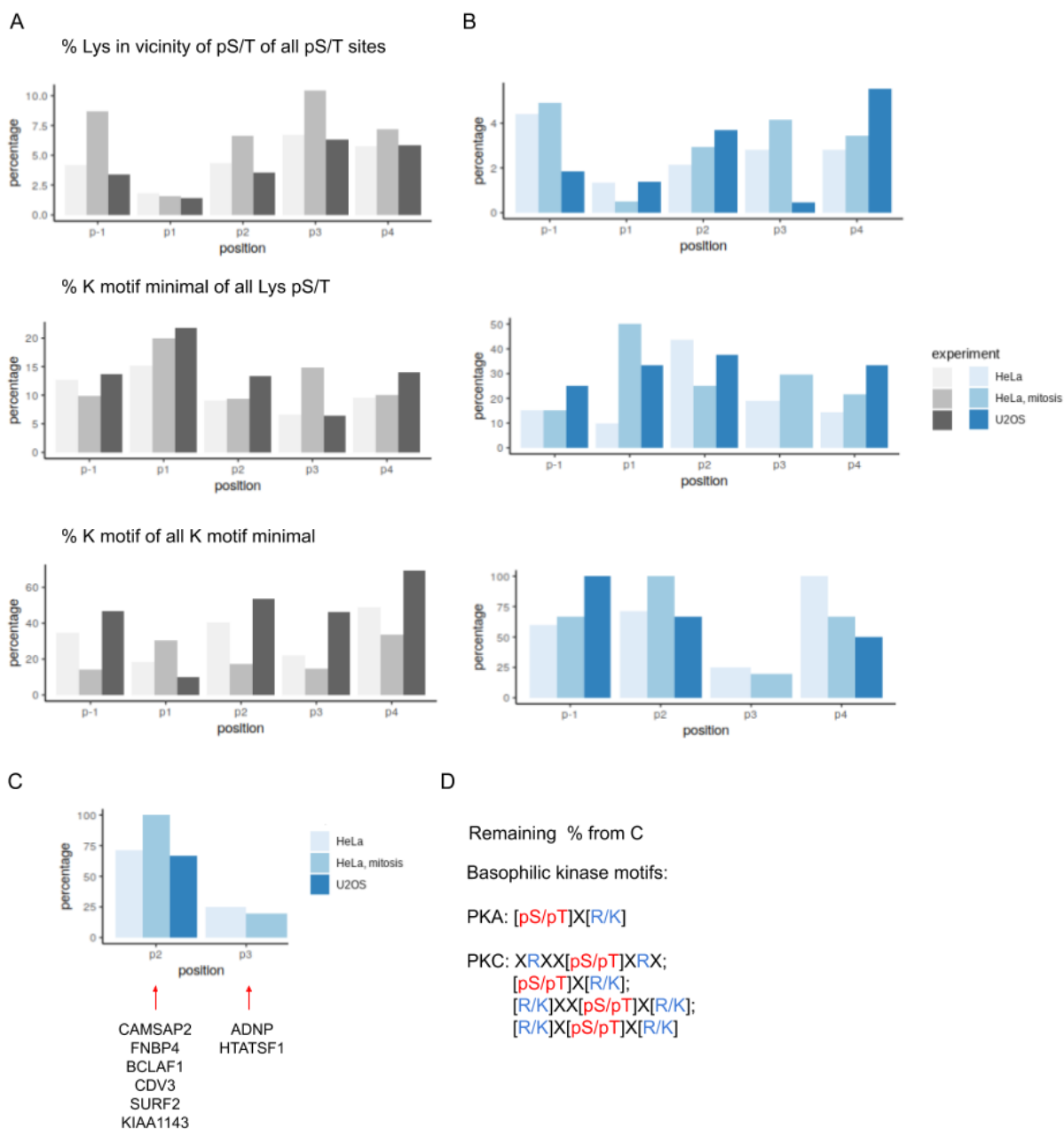


+1 and a higher conservation for Glu at position +3 (Fig. 2.3B). For the +3 hits equal conservation of Asp and Glu was observed at positions +1, +2, and +4 (Fig. 2.3C). Special attention was given to the +1 hits as a complete abrogation of phosphorylation was observed for this K motif variant in the peptide assays (Fig. 2.1). Noticeably, for the +3 (4 sites) and +1 (1 site) hits (Fig. 2.3A) the presence of four acidic positions downstream of the phosphorylated site by CK2 (Fig. 2.3C-D) was observed compared to two acidic positions conserved for the +2 variant (Fig. 2.3B).

#### 2.4.2 *In silico* analysis of CK2 inhibition experiments by CX-4945 treatment

Data from phosphoproteomic experiments profiling the response of HeLa and U2OS cells to treatment with the ATP-competitive CK2 inhibitor CX-4945 [50] in a range of concentrations was retrieved from the literature (HeLa mitosis, 5  $\mu$ M CX-4945 for 45 min [51]) or generated by our group (HeLa cells 20  $\mu$ M CX-4945 for 1-4 h [52] and U2OS cells, 30  $\mu$ M CX-4945 for 4 h (<https://ir.lib.uwo.ca/etd/5164>)) and analyzed for K motif patterns occurrence (Supp. Table 2.4). From all the Ser and Thr phosphosites identified 1.4-10.4 % from all data sets contained a Lys residue in the vicinity of the phosphorylatable site, p-1-p+4 (Fig. 2.4A, top). Of these sites between 6.4-21.7% matched the K motif minimal across all positions (Fig. 2.4A, middle) and between 10-60% of the minimal hits matched the K motif patterns (Fig. 2.4A bottom). After quantification and filtering (fold change  $\leq$  -2, P value  $\leq$  0.05), 0.5-5.5 % of all inhibited sites contained Lys in the vicinity of the phosphorylated residue (Fig. 2.4B, top). From these, 10-50 % sites matched the K motif minimal patterns (Fig. 2.4B, middle) and between 25-100 % of the minimal hits matched the K motif patterns (Fig. 2.4B bottom). In the latter case, for positions -1, +2, and +4 the percentage was 60-100 % whereas for +3 was 20-25 % (Supp. Table 2.4). Noticeably, inhibition of phosphosites quantified in singly-phosphorylated peptides was observed for the +2 and +3 variants of K motif (no K motif minimal) but no for +1 (Fig. 2.4B, Supp. Table 2.4). As before, the functions of the tentative K motif substrates are diverse including transcriptional regulation, apoptosis, cell cycle regulation, etc. (Supp. Table 2.4). The inhibition observed for sites matching the K motif minimal patterns could result, at least in part,

from downstream and/or off-target effects of CX-4945 treatment. Among the downstream pathways that could be affected are those regulated by basophilic kinases such as PKA and PKC [17], as the consensus motif for these kinases can overlap with the K motif patterns (Fig. 2.4D).



**Figure 2.4.** Phosphosites identified and quantified in three different phosphoproteomic studies of CX-4945 inhibition that match the K motif variants. A) Identified phosphosites. B) Inhibited phosphosites, fold change  $\leq -2$ ; singly phosphorylated peptides. The HeLa mitosis dataset does not have aggregated sites, meaning that the same

site found in two different peptide charge states was not quantified as one. C) Tentative CK2 target sites matching K motif variants. D) Examples of basophilic kinases consensus. p: position.

### 2.4.3 Asp-N digestion experiment

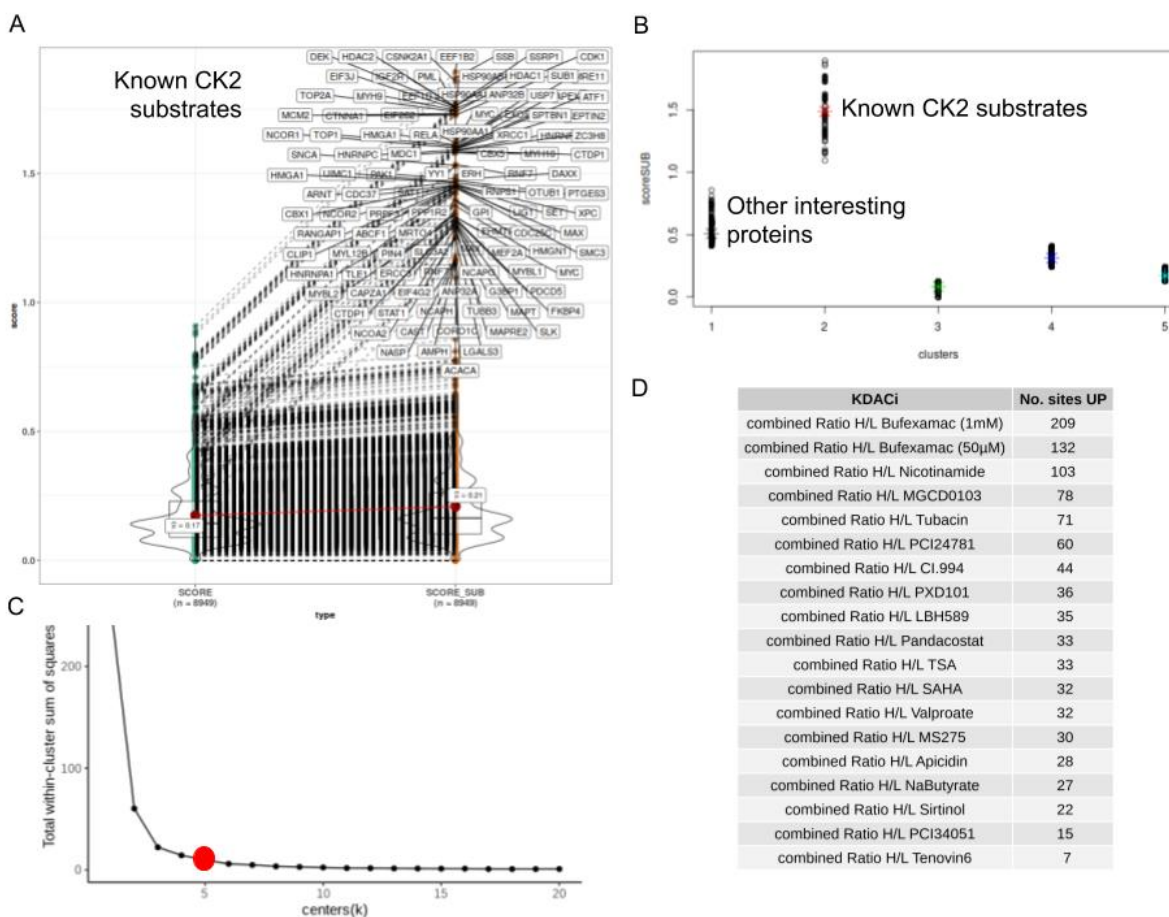
The experiments analyzed rely on trypsin digestion for generating phosphopeptides during protein digestion. Trypsin cuts towards the C-terminal of the amino acids Lys and Arg generating what are referred to as tryptic peptides. To expand the search for K motif a double-SILAC (Stable Isotope Labeling by/with Amino acids in Cell culture) phosphoproteomic study in U2OS cells was performed after 6 h treatment with 30  $\mu$ M CX-4945 using the endoproteinase Asp-N as the enzyme for protein digestion [53] (Appendix B). The endoproteinase Asp-N, as per its name, cuts proteins endogenously towards the N-terminal of Asp residues. Since the K motif pattern was designed to match Asp residues in several positions this experiment was of great value to uncover additional hits. Also, since the K motif pattern also includes K, the SILAC approach is feasible for quantification as both Lys and Arg are labelled. As a result, phosphorylation of Ser24 of SET, a hit of K motif position +2 that was tested in the peptide assays (Fig. 2.1B), was detected but did not show any changes after treatment compared to control (Table 2.3). The adjacent phosphosite, Thr23 of SET, was also quantified but inhibition was not observed (Table 2.3). Importantly, only the singly modified peptides were identified, not the doubly-phosphorylated, if occurring in the cell. Other K motif hits quantified and found non inhibited were: Ser375 of PACSIN2 (p+2) and Ser302 of GBF1 (p+1) (Table 2.3). The absence of inhibition of position +2 hits after CX-4945 treatment does not necessarily reflect the expected outcome upon CK2 manipulation. However, not all CK2 sites should be expected to change upon treatment with a CK2 inhibitor in a cell line given the regulatory pathways of protein phosphorylation in place. Interestingly, inhibition of a site with an Arg at position +2, Ser371 of EZH2, but not for Arg position +1 was observed. Finally, inhibition of phosphosites located in the vicinity of K motif hits was observed for Ser160 and Ser158 of BAZ1B of, ~2 fold change (Table 2.3). This is interesting as CK2-dependent signaling could not only be directly targeting the Ser161 of BAZ1B in this specific context, but neighboring residues (Table 2.3).

**Table 2.3.** K motif hits represented in the Asp-N phosphoproteomic dataset.

Ratio H/L	Site	Motif	Gene	Modified sequence	Category
-1.02	Ser375	SEKD	PACSIN2	_DEDDTGSTVS(ph)EK_	K motif
1.01	Ser24	SEKE	SET	_DETS(ph)EKEQQAIEHI_	
1.00	Ser24	SEKE	SET	_DGADETS(ph)EKEQQAIEHI_	
-1.13	Ser302	SKED	GBF1	_DSGLEFSSQTTS(ph)KE_	
-1.01	Thr23	TSEK	SET	_DET(ph)SEKEQQAIEHI_	K motif in vicinity
-1.84	Ser160	SSDK	BAZ1B	_DSPS(ph)SDKENSSQIAQ_	
-1.73	Ser158	SPSS	BAZ1B	_DS(ph)PSSDKENSSQIAQ_	
1.23	Ser1236	STSS	RBBP6	_DYTS(ph)TSSTGGSPVRKSEKT_	
1.03	Ser911	SRDD	TMEM94	_DDLNQVS(ph)R_	R motif
-1.27	Ser178	SRED	TUBGCP5	_DQQPLS(ph)RE_	
-2.92	Ser371	SDRE	EZH2	_DTDS(ph)DREAGTETGGENN_	

#### 2.4.4 Acetylated proteins functionally linked to CK2

After determining the specificity of CK2 against peptides containing acetylated K motif variants the functional link between CK2 and acetylation was explored by applying an integrative prioritization algorithm (see Methods and Materials section). As a result, more than 180 acetylated-proteins identified in the study by Scholz *et al.* 2015 [24] were functionally linked to CK2, including 124 known CK2 substrates (Fig. 2.5A-B) and more than 60 other proteins of interest (Fig. 2.5B-C).



**Figure 2.5.** Acetyl-proteins identified by Scholtz *et al.* 2015 functionally linked to CK2. A) Outliers detection (outlier.coef = 5, Tukey's method). B) K-means clustering of scores, centers (k) = 5. Cluster 2: 124 acetyl proteins (CK2 substrates); Cluster 1: acetyl proteins 204 (score > 0.46); Cluster 3-5: not enough information to link to CK2. C) Optimal k determined by Elbow method. D) KDAC inhibitors profiled by Scholz *et al.* 2015 [24] that can be used to explore the link between CK2-dependent phosphorylation and lysine acetylation in the cell. UP: increased acetylation. Sequence windows were used to calculate frequencies; fold change cut-off  $\geq 1.5$ .

Next, the quantitative data provided by Scholz *et al.* 2015 were assessed to find the treatments that induced an increased acetylation of the proteins functionally linked to CK2. From this analysis, it was determined that treatment with the lysine deacetylase inhibitors Bufexamac (HDAC6 inhibitor) and Nicotinamide (pan-SIRT inhibitor) increased the acetylation of the acetyl-proteins linked to CK2 (Fig. 2.5D). The rest of the inhibitors were found equally represented. Among them Trichostatin A (pan-HDAC inhibitor) was chosen because of the scholarly literature available. The compound

MGCD0103 (HDAC1, HDAC2, HDAC3, and HDAC11 inhibitor) could also be of interest for future studies (Fig. 2.5D).

## 2.5 K motif representation in the proteome and modification networks

### 2.5.1 Acetylation motif discovery

Previously we determined *in vitro* that the K motif variant hits are targeted by phosphorylation. Here, we determined if the K motif variants can also be targeted by acetylation *in vivo* by exploring the acetylation sites identified in Scholz *et al* 2015. For this motif, discovery was performed with the tool MoMo v5.0.5 [54] from the MeMe suite [55,56]. As a result, a total of 15 motifs out of the 34 discovered significantly enriched (adjusted P value < 0.01, Supp. Table 2.5) showed acetylated Lys close to negatively charged residues (Table 2.4). This indicated that acetylation of sites conforming to the K motif variants *in vivo* is possible. The same analysis could have been performed for all the lysine acetylation sequences reported in the PSP database, however, the input dataset will be heterogeneous. On the other hand, by using the data from Scholz *et al.* 2015 the results were circumscribed to an *in vivo* study with enough information accessible (e.g., MS raw data) to assess the quality of the data.

**Table 2.4.** Significant motifs in an acidic environment discovered on acetylation sites identified in Scholz *et al.* 2015 [24] using MoMo v5.0.5.

Id	Motif	Regexp	Score	Fg_match	Bg_match	Fg/Bg	Adjusted_p-value
1	xxxxxE_K_xxxxKK	....EK....KK	122.2	24	1	24	0.00037
2	xxxxxE_K_xxxCxx	....EK...C..	103.2	32	3	10.7	9.10E-05
6	xxxMxD_K_xxxxxx	...M.DK.....	42.47	31	6	5.2	0.0083
7	xxxQxx_K_Dxxxxx	...Q..KD.....	35.3	50	10	5	3.10E-05
8	xxxTxD_K_xxxxxx	...T.DK.....	62.84	48	10	4.8	9.30E-05
10	xxxxxE_K_xxxxxR	....EK....R	95.27	107	27	4	3.30E-10
13	xxxxxE_K_xrxxx	....EK..R..	48.46	99	27	3.7	1.40E-08
14	xxxxxD_K_xrxxx	....DK..R..	45.42	52	14	3.7	0.00056
15	xxxxxE_K_lxxxx	....EK.I...	78.15	104	29	3.6	7.60E-09
16	xxxxxG_K_Dxxxxx	....GKD.....	46.89	88	27	3.3	1.90E-06
17	xxxxxE_K_xxxxxK	....EK....K	63.62	196	61	3.2	1.40E-15
18	xxxxxD_K_xxxxKx	....DK....K.	53.48	89	28	3.2	2.60E-06
23	xxxxxE_K_xxxxxx	....EK.....	32.81	1372	769	1.8	1.70E-42

25	xxxxxD_K_xxxxxx	.....DK.....	25.46	803	466	1.7	1.00E-20
31	xxxxxx_K_Dxxxxx	.....KD.....	14.95	710	504	1.4	1.70E-07

## 2.5.2 K motif proteome-wide acetylation

The Scholz *et al.* 2015 data set, HeLa cells acetylome enrichment, was scanned for K motif hits. Hits from all three K motif variants were observed among the identified or increased acetylation sites (fold change > 1.5) after treatment with 1 or more of the 19 deacetylase inhibitors tested in the study (Table 2.5). No significant differences were observed between K motif variants for the proportion of identified (P value = 0.05827, Cohen's w = 0.176) or increased (P value = 0.8623, Cohen's w = 0.05853) acetylated hits among sites with neighboring Ser/Thr residues (Table 2.5). Thus, contrary to CK2-dependent phosphorylation, lysine acetylation near a phosphorylatable Ser/Thr site does not seem to occur preferentially at K motif +2 variant, or any other variant(s). This analysis is a snapshot and the results could change for acetylome data sets with a higher coverage increasing the power to detect differences. It can also depend on the biological context e.g., different treatment type, treatment duration, and cell line.

**Table 2.5.** Acetylation sites identified in Scholz *et al.* 2015 [24] matching the K motif variants and total number of acetylation sites in the vicinity of Ser/Thr residues; unique sequences.

Position	K motif	# Sequences (Up)	K Motif, extended	# Sequences (Up)
p+1	[ST]K[DE][DE]	48 (27) S/T vicinity: 1340*	[ST]K[DE][DE][DE]	3(2)
p+2	[ST][DE]K[DE]	70 (27) S/T vicinity: 1274*	[ST][DE]K[DE][DE]	11(5)
p+3	[ST][DE][DE]K	56 (30) S/T vicinity: 1294*	[ST][DE][DE]K[DE]	6(3)
P value**		0.05827 (0.8623)	P value**	0.07685 (not enough hits)
Effect size		0.176, small effect (0.05853, no effect)	Effect size	0.5001, large effect (not enough hits)

Up: acetylation upregulated by treatment with deacetylase inhibitors (fold change > 1.5). \*Vicinity: [ST]K, [ST].K, [ST]..K. \*\* P values were calculated using a 3-sample test for equality of proportions without continuity correction. Cohen's w effect size.

To further explore the occurrence *in vivo/in vitro* of lysine acetylation in K motif variants, acetylation data were retrieved from the PSP database (September 2019) (Table 2.6). As before, no preference was observed in the acetylation of the K motif variants (P value = 0.2126, Cohen's w = 0.1272). Interestingly, around 23-36% of the acetylated K motif hits retrieved from PSP are also targeted by phosphorylation, no co-occurrence implied (Supp. Table 2.6). Contrary to what was observed in CK2 phosphorylation and CX-4945 inhibition data (Fig. 2.3A, Fig. 2.4B) no differences were observed in the occurrence of phosphorylation between variants (Table 2.6). This suggests that other kinases besides CK2 also target these sites in the cell; this is expected, from the number of basophilic kinases that exist (Fig. 2.4D). More importantly, 3 and 4 K motif extended hits for positions +1 and +2, respectively, were found annotated as acetylated and as phosphorylated (Table 2.6). Based on the peptide assay (Fig. 2.1) and the *in vitro* CK2 assay phosphoproteomic study (Fig. 2.3) these are candidate targets of CK2 (Supp. Table 2.6). As before, our results could change with the addition of new data to the PSP database.

**Table 2.6.** Acetylation sites annotated in the PSP database (September 2019) conforming to K motif variants and corresponding phosphorylation data, unique sites.

Position	K motif	# Acetylation sites	K motif, extended	# Acetylation sites
p+1	[ST]K[DE][DE]	63 S/T vicinity: 2695* Phospho: 23 (36.51 %)**	[ST]K[DE][DE][DE]	5 Phospho: 3
p+2	[ST][DE]K[DE]	57 S/T vicinity: 2744* Phospho: 18 (31.58 %)**	[ST][DE]K[DE][DE]	6 Phospho: 4
p+3	[ST][DE][DE]K	46 S/T vicinity: 2752* Phospho: 11 (23.91 %)**	[ST][DE][DE]K[DE]	3 Phospho: 0
P value***		Acetyl: 0.2126 Phospho: 0.1068	P value***	not enough hits
Effect size		Acetyl: 0.1351, small effect Phospho: 0.2919, small effect	Effect size	not enough hits

\*Vicinity: [ST]K, [ST].K, [ST]..K; \*\* Overlap of acetylation with phosphorylation (does not imply co-occurrence). \*\*\* P values were calculated using a 3-sample test for equality of proportions without continuity correction. Cohen's w effect size.



### 2.5.3 K motif proteome-wide phosphorylation

The proportion of K motif variants annotated as phosphorylated in the PSP database among all Ser/Thr phosphosites containing lysine in the vicinity was determined. As a result, more than 200 hits matching each variant were found (Table 2.7, Supp. Table 2.7). A small but significant difference between proportions was observed for K motif (P value = 1.962e-13, Cohen's w = 0.2568) and K motif extended (P value = 0.001, Cohen's w = 0.2814) hits. The same small difference was observed when directly comparing the number of hits for all variants independent of the Ser/Thr phosphosite background (Chi-squared test for given probabilities P value = 0.04, Cohen's w = 0.09411). Importantly, > 40 phosphosites matching any K motif extended variants (Table 2.7) were identified and these are tentative CK2 targets (Supp. Table 2.7). Again, results could change with the addition of new data to the PSP database.

**Table 2.7.** K motif representation in PSP phosphorylation data, unique sites.

Position	K motif	# Phospho sites	K motif, extended	# Phospho sites
+1	[ST]K[DE][DE]	236 S/T vicinity: 8,509	[ST]K[DE][DE][DE]	41
+2	[ST][DE]K[DE]	278 S/T vicinity: 11,594	[ST][DE]K[DE][DE]	63
+3	[ST][DE][DE]K	224 S/T vicinity: 15,646	[ST][DE][DE]K[DE]	41
P value*		1.962e-13	P value***	0.001
Effect size		0.2568, small effect	Effect size	0.2814, small

K vicinity: [ST]K, [ST].K, [ST].K. \* P values were calculated using a 3-sample test for equality of proportions without continuity correction. Cohen's w effect size.

Phosphorylation and acetylation information was also retrieved from the database iPTMnet [57] (September 2019) and combined with the Scholz *et al.* 2015 (acetylation) data to expand the list of K motif hits targeted by acetylation and phosphorylation (Table 2.8). As with the PSP analysis, more than a dozen hits were found to be targeted by these modifications although the existence of both modifications does not imply co-occurrence (Table 2.8, Supp. Table 2.8). Of note, this analysis was carried because iPTMnet includes data from PSP and other resources further increasing the wealth of data analyzed. In

addition, these hits were also shown to be targeted by other PTMs including methylation, sumoylation, and ubiquitination (Table 2.8, Supp. Table 2.8). The kinase-substrate relationships retrieved provided information for a small subset of sites (Table 2.8, Supp. Table 2.8). A few sites matched predicted cis interplaying sites annotated in the database PTMcode v2 (October 2019) (Table 2.8). Results could change with the addition of new data to the databases iPTMnet and PTMcode v2 as new potential sites could be added.

**Table 2.8.** K motif hits annotated phosphorylated in the database iPTMnet (September 2019) and reported acetylated in iPTMnet and/or the Scholz *et al.* 2015 data [24]; corresponding PTM interplay from PTMcode v2. Unique sites.

Position	Phospho-acetyl* (iPTMnet)	Phospho-acetyl* (Scholz <i>et al.</i> 2015)	Kinase-substrate relationships	PTM co-regulation (PTMcode v2)
p+1	Phospho-acetyl: 22 Methylation: 1 Sumoylation: 2 Ubiquitination: 12	Phospho-acetyl: 13 Sumoylation: 1 Ubiquitination: 11	CHEK1: YBX3 T112, YBX1 T80, YBX2 T115	1 (Structural distance, coevolution)
p+2	Phospho-acetyl: 17 Methylation: 1 Ubiquitination: 10	Phospho-acetyl: 11 Ubiquitination: 8	CSNK2A1: APEX1 S123, MAPRE2 S236	4 (Structural distance, coevolution)
p+3	Phospho-acetyl: 12 Sumoylation: 3 Ubiquitination: 6	Phospho-acetyl: 8 Sumoylation: 2 Ubiquitination: 5	-	3 (Structural distance, coevolution)

\*Overlap of acetylation with phosphorylation (does not imply co-occurrence)

A similar number of K motif hits reported phosphorylated was found annotated in the iPTMnet databases as compared to PSP (Table 2.7, Table 2.9). These results could change with the addition of new data to the databases iPTMnet and PTMcode v2. Additional PTM types were also observed targeting the Ser/Thr and K in the K motifs (Table 2.9). Furthermore, an additional number of enzyme-substrate relationships was uncovered (Table 2.9). Results could change with the addition of new data to the databases iPTMnet and PTMcode v2.

**Table 2.9.** PTM types targeting the K motif hits annotated in the database iPTMnet (unique sites).

PTM type	Position			Amino acid modified
	p+1	p+2	p+3	
Acetylation	61	56	44	K

Methylation	4	6	4	
Sumoylation	24	1	34	
Ubiquitination	143	109	83	
Phosphorylation	228	271	221	S/T
O-glycosylation	1	6	1	
Enzyme	14	5	7	Target K or S/T

The fact that many other modification types were found to target the K motif (Table 2.9) suggests that several of the endogenous sequences conforming to the motifs are hubs for signal integration in the cell. This was supported by the retrieval of the corresponding enzyme-substrate relationships from PSP and iPTMnet, with 27 kinases identified (Table 2.10, Supp. Table 2.9). CK2 was found to phosphorylate hits of position +2 and +1. Regarding the latter, other acidophilic kinases were also found represented at position +1 pointing to a possible overlap of consensus which is especially relevant for non-extended K motif variants (Fig. 2.3D, Fig. 2.4C). Basophilic kinases were found to phosphorylate sites from all three variants as the lysine residue in the K motif variants provide the required determinants (Table 2.10, Fig. 2.4D).

**Table 2.10.** Kinase-substrate relationships corresponding to K motif hits annotated phosphorylated in the PSP and iPTMnet databases (September 2019).

Position	Enzyme
p+1	CDK5, CHEK1, FAM20C, GSK3B, MARK1, PKN1, PLK1**, PLK2**, PRKAA1, PRKACA*, RPS6KA5, RPS6KB1, SGK1, TRPM7, CSNK2A1**
p+2	AKT1*, CSNK2A1**, GRK2, PRKACA*, PRKCA*
p+3	CAMK1*, EEF2K, FAM20C, PAK1, LRRK2, PRKACA, PKA

\*basophilic kinase; \*\*Acidophilic kinase

#### 2.5.4 K motif representation in the human proteome

To determine the proteome-wide occurrence of the K motif the human proteome was scanned for the K motif linear patterns using FIMO v5.0.5 [58] and the motifs discovered with MoMo v5.0.5 using the Scholz *et al.* 2015 dataset as input (Supp. Table 2.5). As a

result, > 100 K motif hits were identified with FIMO (log-likelihood ratio score) as a significant match (q-value < 0.05, i.e., the minimal FDR threshold at which the *P*-value is deemed significant) to the sequences reported acetylated *in vivo* in the HeLa cells by Scholz *et al.* 2015 (Table 2.11, Supp. Table 2.10). This indicated that several K motif hits in the proteome are potential candidates for acetylation *in vivo* as they are a significant match for *in vivo* acetylation sequences. Interestingly, the best proteome match retrieved for the discovered motif xxxMxD\_K\_xxxxxx (Table 2.4) in Scholz *et al.* 2015 was KLAMSDKELMKAA (adjusted P value 0.0083), which matches the K motif variant at position +2. According to our results, the proportion of candidate acetylation hits in the proteome for the K motif +1, +2, and +3 variants is 1:7:19.25 (Chi-squared test for given probabilities P value < 2.2e-16, Cohen's w = 0.8361).

**Table 2.11.** K motif representation on the human proteome with FIMO v5.0.5 by using the MoMo v5.0.5 motifs discovered in the acetylation sites identified in Scholz *et al.* 2015 [24] as input; unique sequences. UniProtKB human sequences reviewed and isoforms April 26, 2019.

Position	K motif	# hits proteome (q-value* < 0.05)	K motif, extended	# Sequences proteome (q-value* < 0.05)
1	[ST]K[DE][DE]	8	[ST]K[DE][DE][DE]	2
2	[ST][DE]K[DE]	56	[ST][DE]K[DE][DE]	12
3	[ST][DE][DE]K	154	[ST][DE][DE]K[DE]	17

\* q-value: The q-value of the motif occurrence. The minimum False Discovery Rate (FDR) required to consider this match significant. FIMO estimates q-values from all the match p-values using the method proposed by Benjamini & Hochberg.

The probability score of finding hits matching the K motif and K motif extended (additional acidic position in the pattern, Table 2.2) by chance in any given human protein (as calculated by ELM database using the amino acid probabilities from UniProt [59]) was determined to be 0.00019 and 0.00003, respectively. Meanwhile, the probability of occurrence by chance of the CK2 pattern annotated in ELM, ...([ST])..E, = 0.0145681 (MOD\_CK2\_1, [ELME000064](#)). This means that the ELM CK2 pattern is 76 and 485 times more likely to appear by chance in any human protein than the K motif and K motif extended, respectively.

To identify all possible K motif hits in the human proteome, all the reviewed canonical and isoform sequences of UniProtKB [60] database (42,371 SwissProt sequences, Sep 2019) were scanned against the K motif linear patterns (Table 2.2) using ScanProsite [61] and the patmatdb tool from EMBOSS suite [62]. As a result, > 3,000 and > 500 sites matching the K motif and the K motif extended variants were identified, respectively (Table 2.12). Thus, the observed probability of finding a K motif or K motif extended hit in a protein sequence was 0.24916 and 0.03677, respectively. The values determined were ~1,000 times higher than the expected probabilities. Furthermore, the number of K motif hits equaled the number of K motif matches in a random proteome (shuffle k-mer 20, like k-mer 1/2) (Table 2.12). A very small yet significant difference was observed between the number of K motif variant hits (Chi-squared test for given probabilities P value = 1.545e-05, Cohen's w = 0.04581) (Table 2.12). This was due to a smaller number of hits for position +3 when compared with +1 and +2 (Pairwise comparisons using chi-squared tests FDR adjusted P value 3.6e-05 and 0.0003, respectively). No differences were observed in the number of K motif extended hits (Chi-squared test for given probabilities P value = 0.2552, Cohen's w = 0.04187) (Table 2.12). Regarding specific sequences likely to be relevant for acetylation and CK2-dependent phosphorylation *in vivo* 495 hits conforming to SDKE were retrieved.

**Table 2.12.** K motif representation on the human proteome using ScanProsite and EMBOSS, unique sites.

Position	K motif	# hits proteome (shuffle k-mer 20)	K motif, extended	# hits proteome (shuffle k-mer 20)
p+1	[ST]K[DE][DE]	3,658 (3,654)	[ST]K[DE][DE][DE]	506 (506)
p+2	[ST][DE]K[DE]	3,606 (3,606)	[ST][DE]K[DE][DE]	502 (502)
p+3	[ST][DE][DE]K	3,293 (3,292)	[ST][DE][DE]K[DE]	550 (550)
Total		10,557 (10,552)		1,558 (1,558)
P value*		1.545e-05		0.2552
Effect size		0.04581, no effect		0.04187, no effect

\* P values were calculated using a Chi-squared test for given probabilities. Cohen's w effect size.

### 2.5.5 K motif boundaries and cellular conditions

Given the thousands of candidate hits retrieved for each K motif variant, it is important to be able to determine their boundary and cellular conditions to find those that constitute “functional hits”. To study the boundary and cellular context of the K motif hits the SLiMSearch4 framework [63] was used which allows proteome-wide discovery of functional motifs with an expected number of instances returned (E) < 10,000 and localized to regions with a IUPred disordered score (based on energy estimation approach)  $\geq 0.4$ . This score threshold is set as default by SLiMSearch for motifs as it allows for a locally more ordered tendency compared to other applications of IUPred like disordered binding regions [63]. As a result, > 700 and > 100 hits on canonical sequences for K motif and K motif extended position, respectively, localized to intrinsically disordered regions in the proteome (Table 2.13, Supp. Table 2.11).

**Table 2.13.** K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off  $\geq 0.4$ ).

Position	K motif	# Sites proteome (canonical)	K motif, extended	# Sites proteome (canonical)
p+1	[ST]K[DE][DE]	723 hits in 655 proteins	[ST]K[DE][DE][DE]	127 hits in 124 proteins
p+2	[ST][DE]K[DE]	833 hits in 734 proteins	[ST][DE]K[DE][DE]	146 hits in 143 proteins
p+3	[ST][DE][DE]K	721 hits in 635 proteins	[ST][DE][DE]K[DE]	156 hits in 134 proteins

Several warnings were raised by SLiMSearch4 across all positions. After filtering, hundreds of hits remained for K motif and K motif extended variants (Table 2.14, Supp. Table 2.11). The warnings represented instances for example where the sites localized to a structural domain in the protein, the site has low surface accessibility, the protein constraining the site is secreted, etc. (Supp. Table 2.11).

**Table 2.14.** Warnings detected on K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off  $\geq 0.4$ ).

Position	K motif	# Sites warnings (keep)	K motif, extended	# Sites warnings (keep)
p+1	[ST]K[DE][DE]	200 (523) Domain: 166 Localisation: 42 Surface accessibility: 12	[ST]K[DE][DE][DE]	34 (93) Domain: 28 Localisation: 4 Surface accessibility: 2

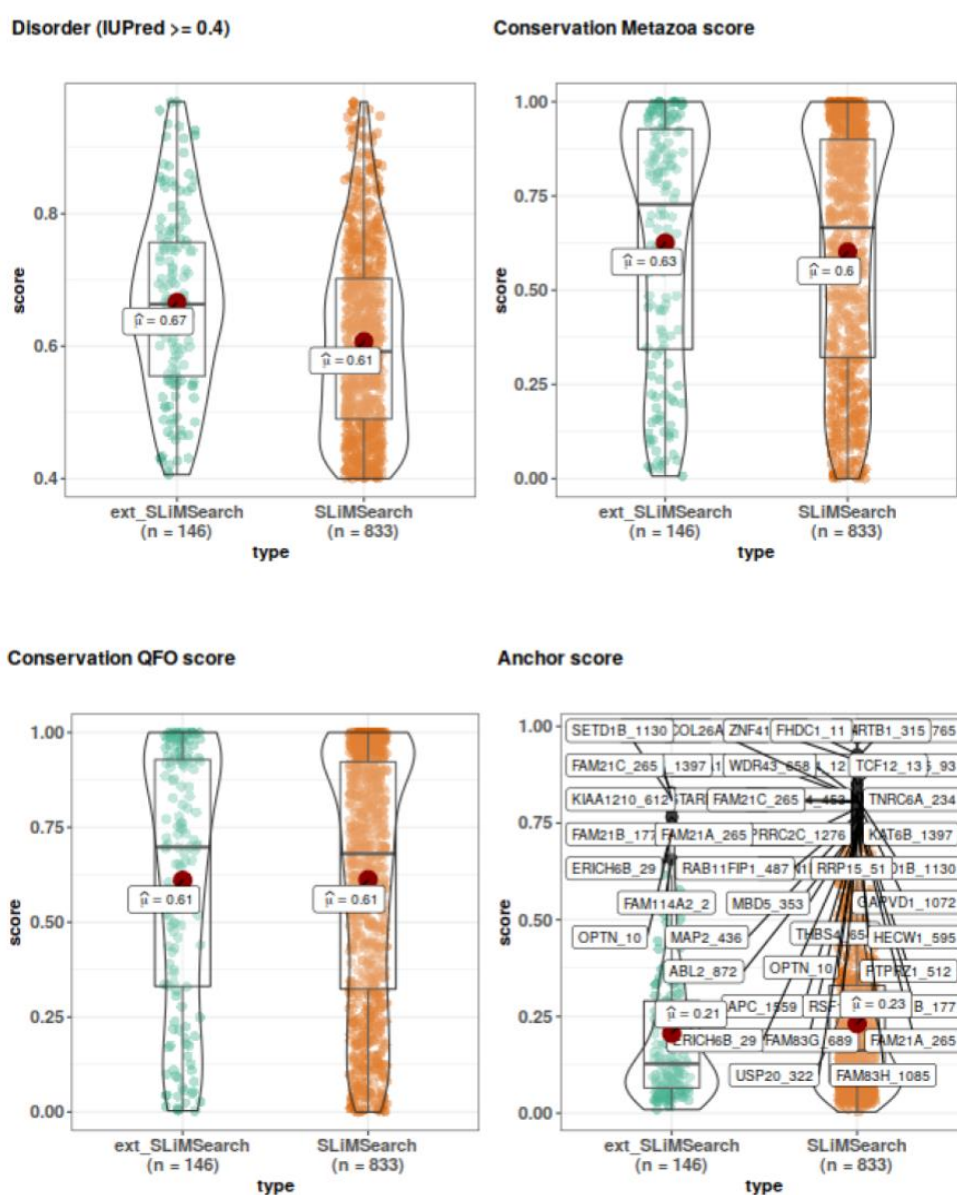
		Topology: 20		Topology: 4
p+2	[ST][DE]K[DE]	216 (617) Domain: 170 Localisation: 36 Surface accessibility: 21 Topology: 37	[ST][DE]K[DE][DE]	31 (115) Domain: 26 Localisation: 3 Surface accessibility: 3 Topology: 8
p+3	[ST][DE][DE]K	180 (541) Domain: 151 Localisation: 36 Surface accessibility: 2 Topology: 18	[ST][DE][DE]K[DE]	42 (114) Domain: 35 Localisation: 9 Surface accessibility: 1 Topology: 3

### 2.5.5.1 Conservation and accessibility

The K motif hits returned by SLiMSearch4 (IUPred  $\geq 0.4$ ) are deemed to have less propensity to be in a globular region of the protein and are considered highly accessible (Fig. 2.6, Supp. Fig. 2.4, Supp. Table 2.11). The mean disorder score for the hits returned was between 0.58-0.67 for the K motif and K motif extended hits, higher than the cut-off of 0.4 (Fig. 2.6, Supp. Fig. 2.4). Additional parameters were also explored, i.e., sequence conservation, surface accessibility, and propensity of folding upon binding (Fig. 2.6, Supp. Fig. 2.4). Regarding conservation, the analysis performed by SLiMSearch4 computed the conservation of the hits in metazoans and orthologs (Quest For Orthologs dataset), values representing conservation over large evolutionary distances [63] (Fig. 2.6, Supp. Fig. 2.4). Some of the hits for which the conservation scores were calculated were found conserved (low score); however, the mean conservation value was between 0.6-0.74 (Fig. 2.6, Supp. Fig. 2.4) indicating that the majority of these sites are not conserved over large evolutionary distances. However, for the K motif extended hits +3, 7 low score outliers (outliers coefficient (Tukey's method) = 1.5.) were found pointing to their conservation in Metazoan (Supp. Fig. 2.4). The propensity to fold upon binding was determined by the anchor score, a higher score implies a higher propensity to folding upon binding i.e., a transition between disorder-to-order upon binding [64]. These binding events are relevant in cell signaling and regulatory processes since the coupling of folding and binding can control binding specificity without increasing affinity and provides flexibility for the binding site to adopt different conformations when faced with

different interactors [64]. A mean anchor score between 0.17-0.23 was observed across all variants meaning that most of the hits do not have a propensity to fold upon binding (Fig. 2.6, Supp. Fig. 2.4). However, several outliers (outliers coefficient = 1.5, Tukey's method) were identified with anchor scores  $\sim 0.75$  across all variants (Fig. 2.6, Supp. Fig. 2.4) suggesting these might localize to disordered binding regions. Finally, not enough information was available to confidently assess surface accessibility of K motif hits (Supp. Table 2.11) due to the lack of protein structure information [63].

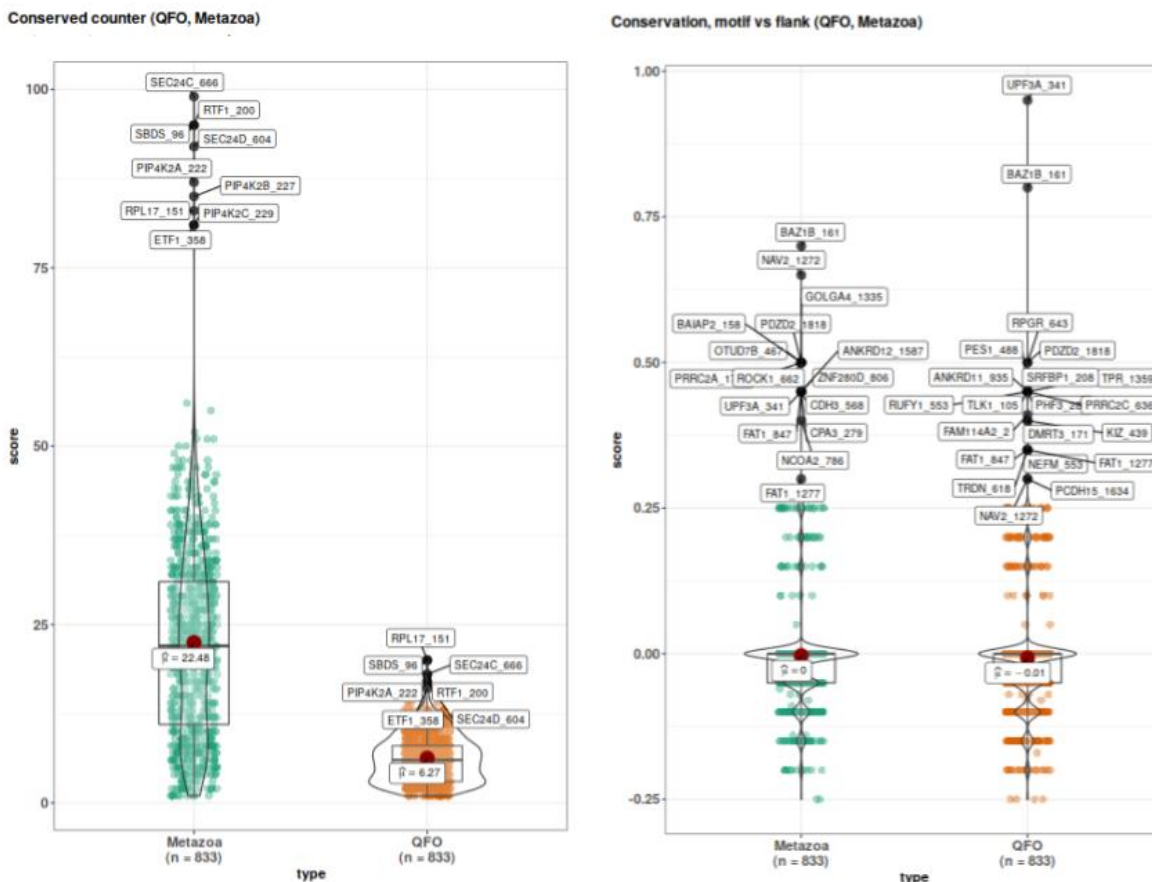
p+2





**Figure 2.6.** SLiMSearch4 results for the K motif variant +2 (K motif extended (ext\_SLiMSearch) and all hits (SLiMSearch)) showing the distribution of the disordered, conservation, and Anchor scores for the hits found. Left-Right: Disorder score (IUPred). Conservation in Metazoa. Conservation in Quest for Orthologs (QFO), Anchor score (disorder-order transition upon binding). Outliers coefficient (Tukey's method) = 1.5.

The conserved counter score, i.e., the number of species in which the motif consensus is present at the same position as the query species motif, was analyzed for outliers (Fig. 2.7, Supp. Fig. 2.5). As a result, nine hits conserved in more than 75 species in Metazoa and seven in more than 10 orthologs were identified (Fig. 2.7, Supp. Fig. 2.5). The mean number of species where the hit position was conserved was 22 and 6 for Metazoan and orthologs, respectively. In addition, the conservation of the K motif variants relative to their flanking regions in Metazoa and orthologs was also analyzed (Fig. 2.7, Supp. Fig. 2.5). This is calculated in SLiMSeach by determining the significantly conserved residues in the defined positions vs the significantly conserved residues in the flanks, +/- 10 amino acids [63]. The mean motif vs flank conservation score was zero meaning that most of the hits do not have conserved residues or have the same proportion of conserved residues between the motif sequence and the flanking region (Fig. 2.7, Supp. Fig. 2.5). Importantly, several outliers with higher conservation inside the motif vs flanks. Among them Ser161 of BAZ1B was observed, a K motif extended hit at position +2 (Fig. 2.7), known to be phosphorylated in cells and neighboring other modified residues.



**Figure 2.7.** SLiMSearch4 hits for K motif +2 variant conserved counter (Outliers coefficient = 1.5) and conservation, motif v flank scores (Outliers coefficient (Tukey's method) = 5).

### 2.5.5.2 Functional features

An array of sequence and boundary properties of the K motif were previously calculated with SLiMSearch4 suggesting a functional relevance for several of the K motif hits across all variants. Thus, the next analysis was focused on those sequence features and properties that provide context on the function of the K motif hits. For this the SLiMSearch4 output was explored for evidence of occurring PTMs (complementary to PSP and iPTMnet analyzes), functional domains, motifs, mutagenesis, SNPs, etc. (Table 2.15, Supp. Table 2.11). Of notice was the finding that sequences conforming to the K motif are targets of SNPs and mutations in the cell (Table 2.15, Supp. Table 2.11).

**Table 2.15.** SLiMSearch4 functional features on K motif hits.

	Position	

PTM type	p+1	p+2	p+3	AA modified
Acetylation	15	16	12	K
Methylation	2	3	2	
Sumoylation	-	1	-	
Ubiquitination	21	14	14	
Phosphorylation	60	83	61	S/T
O-glycosylation	-	1	---	
Enzyme	1	1	-	K, S/T, or D/E
Other	1	5	3	
Domain	29	31	25	
Motif	2	1	1	
Structure	7	16	7	
SNP	33	40	25	
Mutagenesis	2	3	0	
Region	24	37	25	
Topology	11	13	9	
Secondary Structure	4	13	5	
Isoform	31	26	24	
Switch	0	1	0	

The domain, motif, and secondary structure information provided by SLiMSearch4 (Table 2.15) was extended to all the K motif hits found in the human proteome scan using ScanProsite with information retrieved from UniProtKB and PDB [65] databases. The K motif hits were found to overlap in three cases with nuclear localization signals and with > 200 domains (Supp. Table 2.12). Also, several types of secondary structure were also identified including unordered regions (Supp. Fig. 2.6), as expected from SLiMSearch4 analysis, but also strand, helix, and turn (Supp. Fig. 2.6).

### 2.5.5.3 Mutations

Since SLiMSearch4 returned information regarding mutagenesis (UniProt: mutated residues which alter function) and SNPs (dbSNP, 1000genomes and UniProt: single nucleotide polymorphism with disease association and genotype information). The hits in which mutations impact the local charge of the K motif were determined.

Correspondingly, several hits were identified where Lys was substituted by Glu and vice versa in all K motif variants (Table 2.16, Supp. Table 2.11). Substitutions of Lys by non-charged amino acids like Ile and Gln were also detected.

**Table 2.16.** Mutations targeting the positions in the K motif returned by SLiMSearch4.

No. hits	Substitution	Protein and site targeted
p +1		
7	S->P:3; S->C:2; S->G:1; S->N:1	FCRL4_510; IRX2_199; NOP56_513; SRCAP_1844; TINF2_305; TRPM2_663; ZNF638_1528
10	T->A:3; T->I:3; T->N:2; T->M:1; T->S:1	ALOX12_568; CHCHD4_16; CMYA5_3027; CWF19L2_278; KDELC1_498; L3MBTL3_183; NFKB1_489; UBE3A_174; ZCCHC6_666; ZDHHC23_238
11	D->N:4; D->E:3; D->G:2; D->A:1; D->V:1	ARHGAP11A_755; DNHD1_620; EPB41L5_578; FAM154A_340; HERC5_230; MIA3_876; OS9_302; PODXL_546; TCP1_537; VCAN_1351; ZNF407_828
15	E->K:5; E->D:3; E->G:3; E->Q:2; E->*:1; E->V:1	EIF4ENIF1_31; EML6_488; FBR3_263; HMG1_70; HNF1B_257; ICT1_169; KIAA2013_628; LRRFIP1_801; MAP2_763; MYH2_1267; ORC3_507; OTUD7A_571; RIF1_1879; SLC8A1_690; TCAP_102
15	K->E:8; K->R:2; K->T:2; K->M:1; K->N:1; K->Q:1	CCAR2_584; CDC5L_625; CENPC_439; HECW1_482; KIF20B_734; LAD1_78; MCM3_176; MSH6_269; POLK_724; PSMC2_12; RREB1_59; SRSF6_265; TTN_16066; WDR49_548; ZNF445_757
p +2		
8	S->C:2; S->N:2; S->G:1; S->L:1; S->P:1; S->Y:1	ALAS1_65; ANKRD12_1758; DMRT3_171; KTN1_1319; LRR1Q1_1536; MXRA5_750; SYCP2L_422; TRDN_339
8	T->M:5; T->I:3	CDH3_568; GABPA_450; KAT6B_1397; KCNH2_895; SEC24D_604; SURF2_195; TMEM209_230; USP35_688
11	D->E:6; D->G:3; D->Y:2	CDYL2_163; CENPJ_759; EVPLL_70; GAPT_107; IFIT3_393; KIF1A_1555; PPIG_442; PRRC2A_1765; RABGAP1L_464; SDC2_107; ZBTB41_324

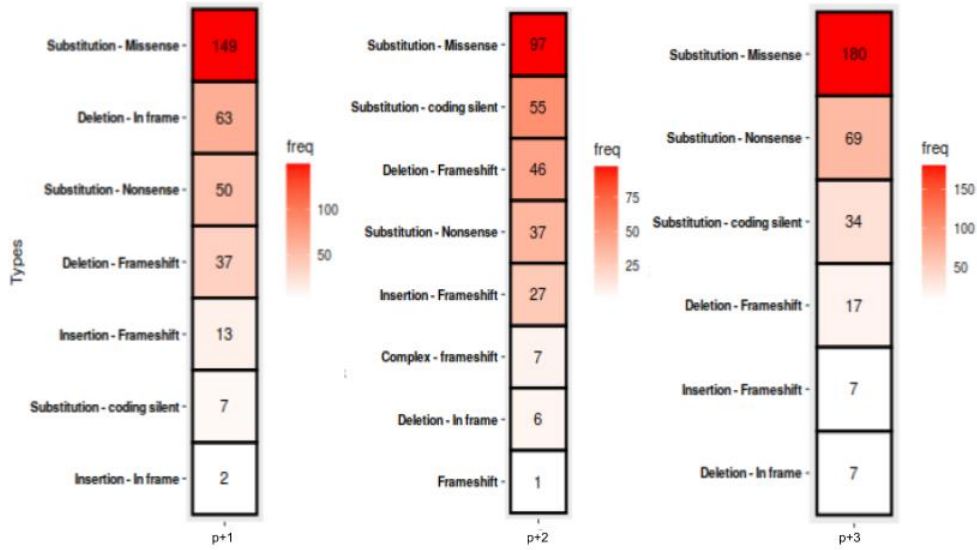
13	E->K:5; E->G:4; E->D:2; E->Q:1; E->V:1	KIAA0101_76; KIF4B_953; LRRD1_2; MYZAP_334; OPTN_10; RNF113A_124; RSBN1L_21; RSF1_643; SETD1B_1065; SMARCA2_992; TIAM2_1314; TSHZ2_662; UBE2E1_22
13	K->N:6; K->E:3; K->R:3; K->I:1	ARID4A_651; C1orf168_398; C2orf44_550; CARD10_512; CCAR2_285; DSC3_408; FBXO40_246; FREM3_460; GOLGA4_1335; LDHA_220; PJA2_447; RMDN3_189; RPL6_235
p+3		
8	S->C:2; S->N:2; S->G:1; S->L:1; S->P:1; S->Y:1	ALAS1_65; ANKRD12_1758; DMRT3_171; KTN1_1319; LRR1Q1_1536; MXRA5_750; SYCP2L_422; TRDN_339
8	T->M:5; T->I:3	CDH3_568; GABPA_450; KAT6B_1397; KCNH2_895; SEC24D_604; SURF2_195; TMEM209_230; USP35_688
11	D->E:6; D->G:3; D->Y:2	CDYL2_163; CENPJ_759; EVPLL_70; GAPT_107; IFIT3_393; KIF1A_1555; PPIG_442; PRRC2A_1765; RABGAP1L_464; SDC2_107; ZBTB41_324
13	E->K:5; E->G:4; E->D:2; E->Q:1; E->V:1	KIAA0101_76; KIF4B_953; LRRD1_2; MYZAP_334; OPTN_10; RNF113A_124; RSBN1L_21; RSF1_643; SETD1B_1065; SMARCA2_992; TIAM2_1314; TSHZ2_662; UBE2E1_22
13	K->N:6; K->E:3; K->R:3; K->I:1	ARID4A_651; C1orf168_398; C2orf44_550; CARD10_512; CCAR2_285; DSC3_408; FBXO40_246; FREM3_460; GOLGA4_1335; LDHA_220; PJA2_447; RMDN3_189; RPL6_235

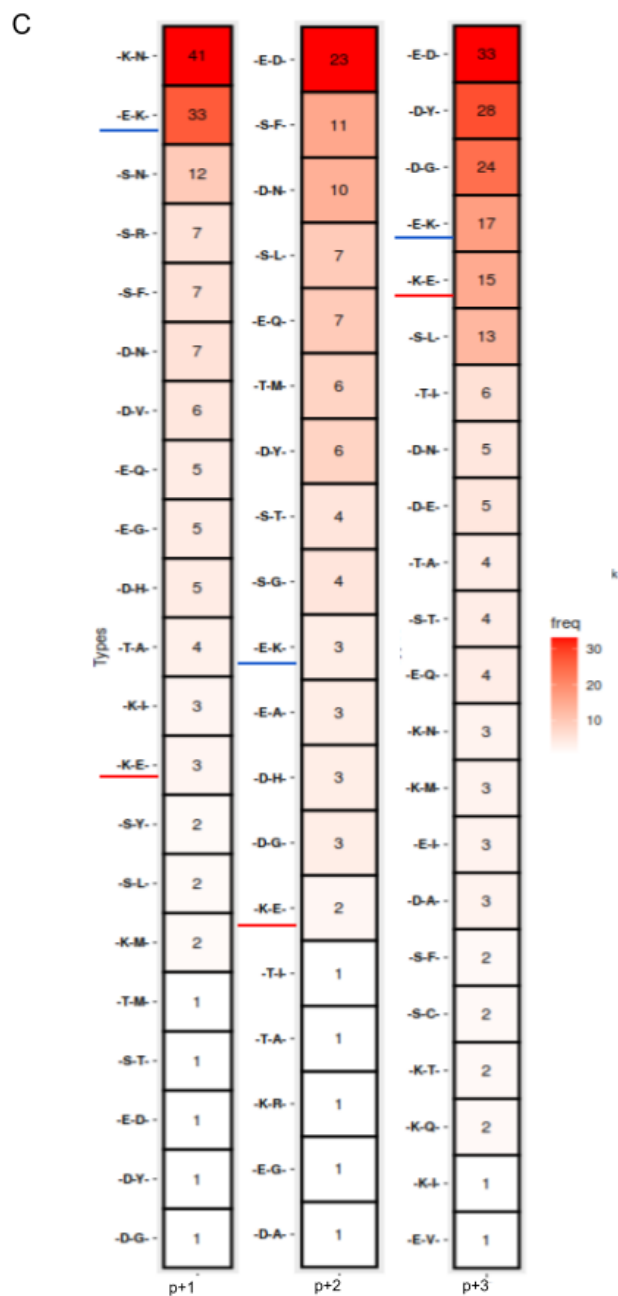
To expand our understanding on the substitutions that could occur inside the K motif hits data from the database COSMIC [66] (the Catalogue Of Somatic Mutations In Cancer) (September, 2019) was also retrieved and analyzed for hits. More than 100 mutation instances in cancer were identified targeting all K motif variants (Fig. 2.8A) including missense substitutions (Fig. 2.8B). According to FATHMM prediction provided by COSMIC most of such mutations were classified as pathogenic and a lower number as neutral (Supp. Table 2.12); furthermore, some of the mutations were annotated as contributing to resistance (Supp. Table 2.12). As with SLiMSearch4 output, the COSMIC data was searched for mutations affecting the local charge of K motif hits (Fig. 2.8C). The mutations retrieved represent different tumor origins: primary, secondary, recurrent, and metastasis; some are non-specified (Supp. Table 2.12). The primary site and histology of the cancer were diverse (Supp. Table 2.12). Additional mutation data was gathered (Supp. Table 2.12).

A

Position	No. mutations, COSMIC curated	No. sites (No. proteins)
p+1	321	118 (85)
p+2	276	101 (83)
p+3	314	128 (80)

B





**Figure 2.8.** Curated mutations annotated in COSMIC (February 2020) on sites matching the K motif. A) Number of mutations per K motif variants. B) Mutation description. C) Missense substitution, unique mutation id.

A total of 3, 2, and 15 missense substitution mutations annotated in COSMIC matching the K motif variants +1, +2, and +3, respectively, were found to change Lys by Glu.

These results were of special interest for +1 and +3 hits where the presence of Lys abrogates phosphorylation by CK2. Although there is no direct evidence of the relevance

of these mutations at the functional molecular level some functional remarks can be made. For example, regarding position +1 mutations impacting the residue Lys1994 of ATM were observed in chronic lymphocytic leukemia. ATM is a Ser/Thr kinase that senses DNA damage, the residue Lys1994 is localized to the FAT domain, a regulatory domain that binds the kinase domain in the absence of DNA damage and thus inhibiting its activity [67]. Interestingly, Lys1994 is in the vicinity of main phosphorylation sites Ser1981 and Thr1989 [67] and has been identified to be ubiquitinated [68]. Another mutational example of interest retrieved from COSMIC and expanded upon from the literature are the variants observed for the protein PIK3CA (Phosphoinositide-3-kinase) generated by the substitution of Gln546 and neighboring residues in the sequence  $^{544}\text{TEQE}_{547}$  to  $^{544}\text{TEKE}_{547}$  and  $^{544}\text{TEEE}_{547}$ . These mutations localize to the helical domain of PIK3CA; as a result, the helical domain is regarded as a “hotspot” for mutations, which are considered to impact binding to the SH2 domain of PI3KCA resulting in activation of the kinase [69]. The impact of this substitutions on phosphorylation and if phosphorylation occurs at Thr544 site (evidence of phosphorylation in the GPMdb database [70]) is not known. However, since PIK3CA is functionally linked to CK2 as a regulator of the AKT/mTOR pathway [71] it raises questions about the relevance of such mutations in the K motif hits and the rewiring of phosphorylation.

#### 2.5.5.4 Regulatory sites

The K motif hits were searched for regulatory sites annotated in the PSP database (September 2019). A total of 3, 8, and 1 regulatory site were observed for +1, +2, and +3 hits, respectively (Supp. Table 2.13). This is a relevant finding as it implies that they are among the few highly studied modification sites in the proteome, likely due to their biological role or the role of the proteins that contain them. The list included 10 phosphosites as well as one sumoylation site and one methylation site. Interestingly, for +2 hits the regulatory sites observed are involved in regulating the activity of the corresponding proteins by inducing protein stability, inducing activity, and regulation of molecular association (Supp. Table 2.13). Their phosphorylation is known to impact transcription, RNA splicing, endocytosis, and translation processes (Supp. Table 2.13). Complementary, a total of 88, 90, and 74 genes which contain mutations that have been



causally implicated in cancer processes and thus driving cancer annotated in the COSMIC Cancer Gene Census [72] were observed for +1, +2, and +3 variants, respectively (February 2020).

### 2.5.5.5 Sites in pathways

Additional important regulatory information retrieved for the K motif hits is their involvement in pathways as annotated in the database Reactome [73]. Regarding the +2 hits the sites K164 of DDX58 and K788 of NCOA2 were annotated to a total of 24 reactions in 15 pathways linked to innate immune system and the sumoylation of NCOA2, respectively (Supp. Table 2.13).

### 2.5.5.6 Relationship to lysine acetylation writers and erasers

The protein-protein interaction data of the proteins containing K motif hits was retrieved using the PSICQUIC services [74] in R [75] and searched for lysine acetyltransferases (KAT) and lysine deacetylases (KDAC) interactors. As a result, 411 and 396 unique instances of gene-KAT and gene-KDAC interactions, respectively, and 199 unique gene-KDAC-KAT interaction pairs were observed (Table 2.17, Supp. Table 2.13). This amounted to a total of 57 combinations of KDACs-KATs represented in the hit proteins (Table 2.17, Supp. Table 2.13). Importantly, for all three K motif variants interactions were observed with both KDACs and KATs (Table 2.17, Supp. Table 2.13).

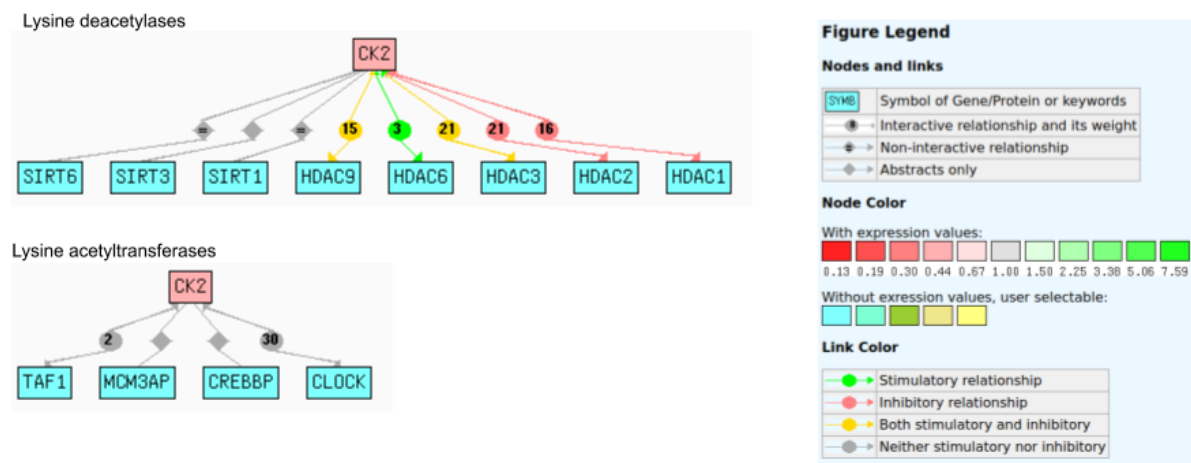
**Table 2.17.** Enzyme-substrate relationships represented in the positions matched in the K motif (iPTMnet).

KDAC-KAT	Gene-KDAC-KAT	Gene-KAT	Gene-KDAC	p+1 hits	p+2 hits	p+3 hits
57	199	411	396	KDAC: 77 KAT: 125 Both: 45	KDAC: 76 KAT: 126 Both: 95	KDAC: 104 KAT: 122 Both: 43

KAT: lysine acetyltransferase; KDAC: lysine deacetylase

For position +2 the KDACs identified as interactors were: HDAC1, HDAC2, HDAC3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9, HDAC10, SIRT1, SIRT3, SIRT4, and SIRT6. The KATs identified were: CLOCK, CREBBP, EP300, GTF3C4, HAT1, KAT14, KAT2A, KAT2B, KAT5, KAT6A, KAT6B, KAT7, KAT8, MCM3AP, NCOA1, NCOA2, NCOA3, TAF1. Interestingly the functional association between CK2

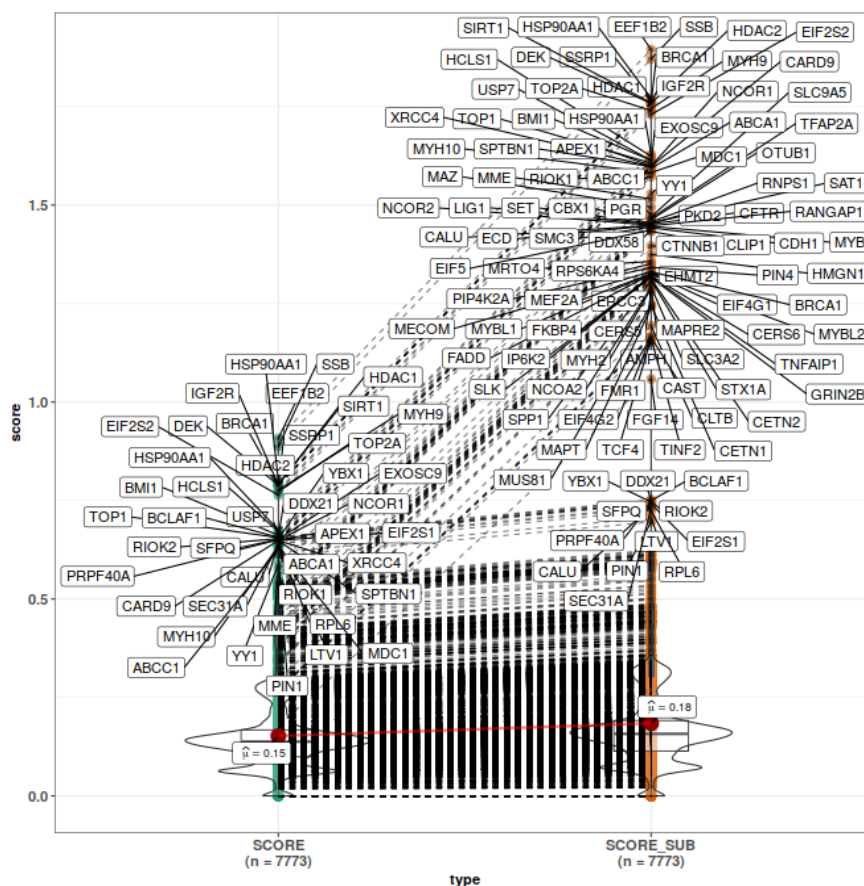
a number of these erasers and writes has been described in the literature (see Chapter 3). The corresponding relationships were extracted from the literature using Chilobot (Fig. 2.9).



**Figure 2.9.** Literature relationships between CK2 and lysine deacetylases and acetyltransferases extracted using Chilobot.

### 2.5.6 Functional link to CK2

The functional association of the K motif hits to CK2 was assessed according to CK2 substrates, CK2 interactors, CK2 neighborhood, mentions in the CK2 literature, co-regulation with CK2, similarity with GO biological processes annotated for CK2, changes when CK2 is manipulated, and conforming to CK2 motif while localizing to subcellular compartments where CK2 is found. More than 320 human proteins conforming to the K motif linear patterns across all variants were found linked to CK2 including 152 known CK2 substrates and more than 170 other closely functionally linked proteins (outlier.coef = 6, Tukey's method) (Fig. 2.10). Collectively, these observations highlight a need for understanding how CK2-dependent signaling integrates at the hit sites in these proteins.



**Figure 2.10.** K motif hits in the human proteome functionally linked to CK2. “SCORE” (0-1), all resources values are averaged, and “SCORE\_SUB”, all resources values are averaged except for the CK2 targets which is summed to the average score allowing for direct identification of known CK2 substrates as outliers. Higher ranking proteins, CK2 substrates: SCORE\_SUB > 1). Resources: CK2 substrates, CK2 interactors, CK2 neighborhood, mentions in the CK2 literature, co-regulation with CK2, similarity with GO biological processes annotated for CK2, changes when CK2 is manipulated, and conforming to CK2 motif while localizing to subcellular compartments where CK2 is found. Outliers coefficient = 6, determined by Tukey’s method.

### 2.5.7 Candidate K motif sites for CK2-dependent phosphorylation

Following the described analysis, a list of known and candidate CK2 substrates conforming to the K motif variant +2 and the K motif extended variants +1 and +3 (with four or more negative residues) was generated. This list contained sites matching the criteria from the curated known CK2 substrates retrieved from PSP, the *in vitro* hits phosphorylated by CK2, the sites inhibited by CX-4945, the sites annotated phosphorylated and acetylated in PSP or just phosphorylated (Supp. Table 2.14).

Functional information about the sites can be accessed in the Supplemental Material. A

total of 113 unique sites was selected (Supp. Table 2.14) with the corresponding 109 proteins displaying diverse molecular activities (Table 2.18, Supp. Table 2.14) including DNA (cytosine-5-)-methyltransferase, histone acetyltransferase, histone methyltransferase, kinase, and nuclease activities.

**Table 2.18.** Summary of GO molecular function, level 5, classification corresponding to the candidate CK2 substrates conforming to the K motif variant.

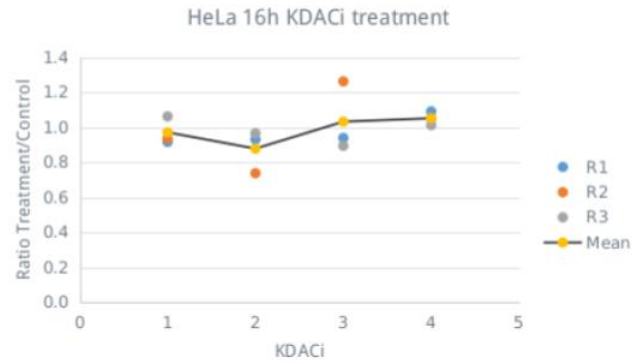
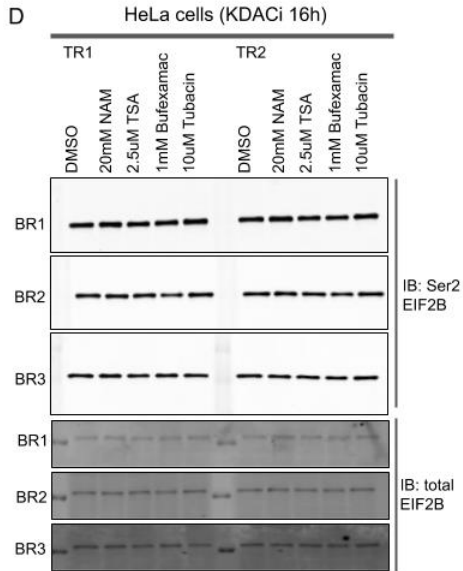
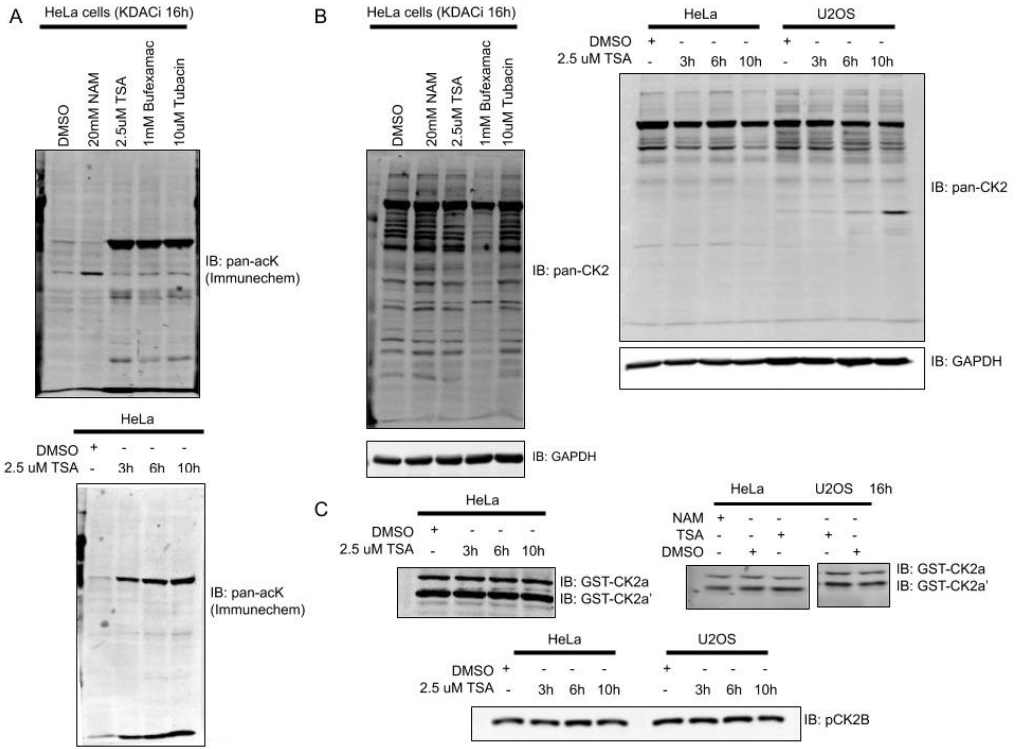
ID	Description	Count	GeneRatio	geneID
GO:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific	4	4/107	TFAP2A/PRDM2/ARNTL/REB1
GO:0016301	kinase activity	4	4/107	RIOK1/PKM/PRKG1/MAST2
GO:0008168	methyltransferase activity	3	3/107	FTSJ3/PRDM2/DNMT1
GO:0001227	DNA-binding transcription repressor activity, RNA polymerase II-specific	2	2/107	TFAP2A/ZEB1
GO:0004518	nuclease activity	2	2/107	APEX1/EXOSC9
GO:0003755	peptidyl-prolyl cis-trans isomerase activity	2	2/107	PPIG/PPIL4
GO:0030374	nuclear receptor transcription coactivator activity	1	1/107	CCAR1
GO:0016615	malate dehydrogenase activity	1	1/107	MDH1
GO:0016747	transferase activity, transferring acyl groups other than amino-acyl groups	1	1/107	KAT6A
GO:0042578	phosphoric ester hydrolase activity	1	1/107	APEX1
GO:1990380	Lys48-specific deubiquitinase activity	1	1/107	OTUD7B
GO:0009982	pseudouridine synthase activity	1	1/107	PUS7L
GO:0003909	DNA ligase activity	1	1/107	LIG1
GO:0042054	histone methyltransferase activity	1	1/107	PRDM2
GO:0003886	DNA (cytosine-5-)-methyltransferase activity	1	1/107	DNMT1
GO:0008649	rRNA methyltransferase activity	1	1/107	FTSJ3
GO:0046943	carboxylic acid transmembrane transporter activity	1	1/107	SLC1A5
GO:0005216	ion channel activity	1	1/107	TRPV5
GO:0099516	ion antiporter activity	1	1/107	SLC9A5
GO:0015186	L-glutamine transmembrane transporter activity	1	1/107	SLC1A5
GO:0015103	inorganic anion transmembrane transporter activity	1	1/107	SLC4A7
GO:0015399	primary active transmembrane transporter activity	1	1/107	RALBP1
GO:0035064	methylated histone binding	1	1/107	CDYL2
GO:0005070	SH3/SH2 adaptor activity	1	1/107	SKAP1
GO:0048018	receptor ligand activity	1	1/107	GDF6
GO:0017112	Rab guanyl-nucleotide exchange factor activity	1	1/107	DENND1A
GO:0005246	calcium channel regulator activity	1	1/107	PRKG1
GO:0030414	peptidase inhibitor activity	1	1/107	APLP2
GO:0005096	GTPase activator activity	1	1/107	RALBP1
GO:0005125	cytokine activity	1	1/107	GDF6

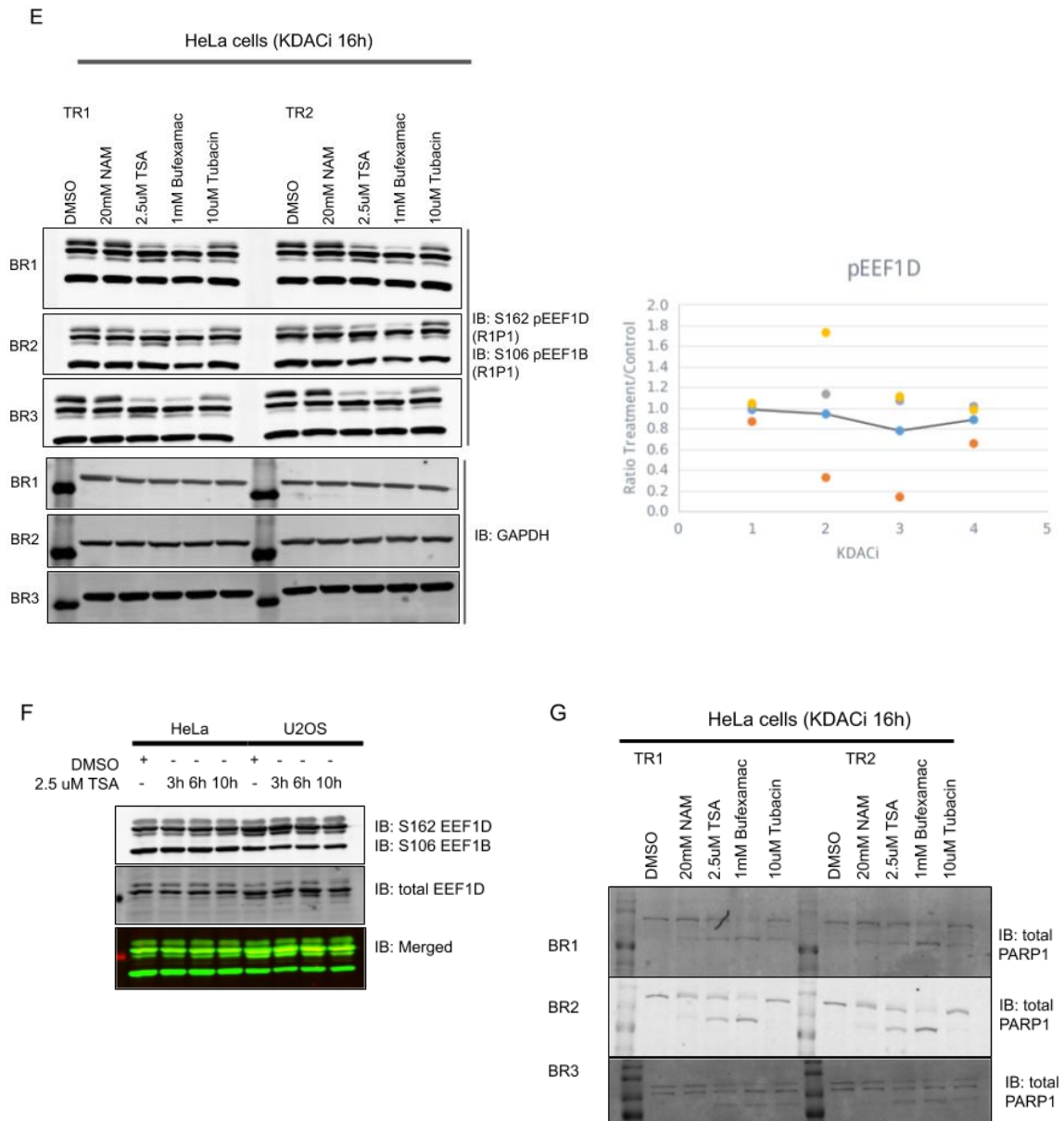
## 2.6 K motif as a ground for PTM interplay

To study the co-occurrence of phosphorylation and acetylation in sites conforming to the K motif we performed immunoblotting and phosphoproteomic studies followed by functional analysis.

### 2.6.1 Modulation of CK2-dependent signaling by lysine deacetylase inhibitors

First, an appropriate treatment to increase acetylation in U2OS and HeLa cells was found by treating cells for 3 to 16h with the lysine deacetylase inhibitors (KDACi) Nicotinamide (NAM), Trichostatin A (TSA), Bufexamac (Bx), and Tubacin (Fig. 2.11A). Then the modulation of CK2-dependent signaling was evaluated using immunoblotting (Fig. 11B-F) and phosphoproteomics (Fig. 2.16-17). Interestingly, a modulation on CK2-dependent signaling was observed at 10 and 16 h of treatment by immunoblotting with the CK2 motif ((pS/pT)DXE) antibody (Fig. 2.11B, Fig. 2.13). No changes were observed in the total protein levels of CK2alpha and CK2alpha' (Fig. 2.11C) or in the phosphorylation by CK2 of the N-terminal of CSNK2B at 16h and 10h of treatment (Fig. 2.11C). No changes were observed in the phosphorylation levels of Ser162 of EEF1D and Ser2 of EIF2B bonafide substrates of CK2 at 16 h of treatment (Fig. 2.11D-F). An induction of apoptosis was observed by monitoring PARP1 cleavage at 16 h, which was more marked in the TSA and Bx treatments (Fig. 2.11G).



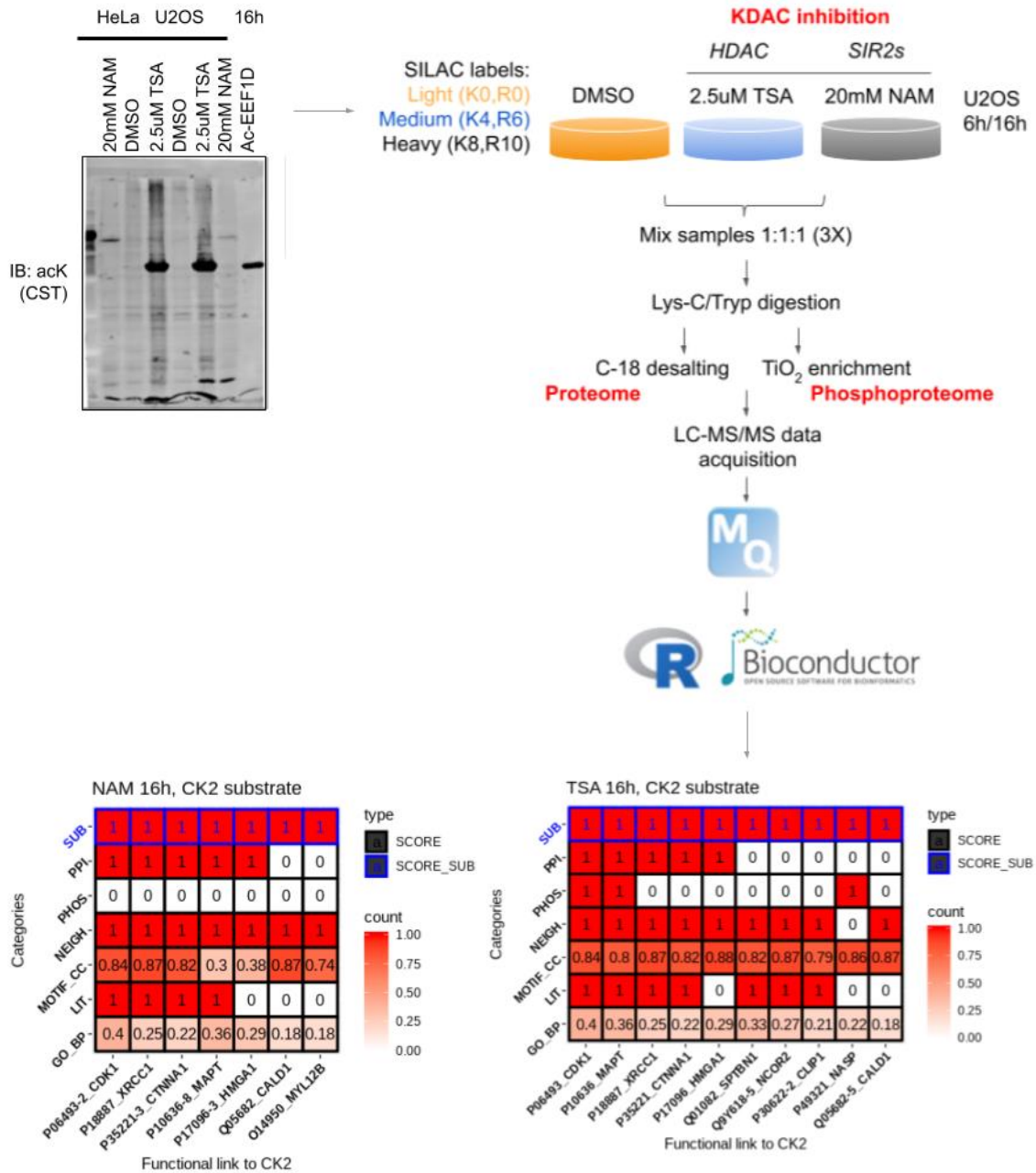


**Figure 2.11.** Immunoblots of lysine deacetylase inhibitor (KDACi) treatment of U2OS and HeLa cells at different time points monitoring total protein acetylation status and the modulation of CK2 subunits and CK2 substrates. A) Increase in acetylation levels in HeLa cells at 3, 6, 10, and 16 h using pan-AcK antibody (Immunechem). B) Impact on CK2 substrates in HeLa and U2OS cells at 3, 6, 10, and 16 h using the pan-CK2 motif ((pS/pT)DXE) antibody. C) Impact on CK2 levels in HeLa and U2OS cells at 3, 6, 10, and 16 h using GST-CK2 antibody. Impact on CSNK2B N-terminal phosphorylation at 3, 6, and 10 h using phosphor-Ser2/3/8 CSNK2B antibody. D) Impact on the bona fide CK2 substrate Ser2 of EIF2B in HeLa cells at 16 h using phosphor-Ser2 and total (Novus) EIF2S2 antibodies. E) Impact on the phosphorylation of the bona fide CK2 target site Ser162 of EEF1D and total protein levels of EEF1D in HeLa and U2OS cells at 3, 6, and 10 h. Detected using phospho-Ser162 and total EEF1D (YenZym) antibodies. G) Total

PARP1 levels in HeLa cells at 16 h using total PARP1 (CST) antibody. Nicotinamide (NAM), Trichostatin A (TSA), Bufexamac (Bx). All immunoblots were performed in three biological replicates (BR1-3) and the signal was detected in a LiCor Odyssey Infrared Imaging System managed by the Odyssey v3.0 software. Quantification of EIF2S2 and EEF1D signals was performed by using the up-bottom background method in Odyssey v3.0 software. Lane normalization for quantification was performed using total GAPDH (Millipore) antibody or REVERT 700 total protein stain (LiCor) as indicated by LiCor.

Since changes were observed with the anti-pan-CK2 substrate antibody the phosphoproteome of isotopically-labeled U2OS cells in response to TSA and NAM treatment at 6 and 16 h was profiled (Fig. 2.12). Overall changes to several sites corresponding to proteins functionally linked to CK2 were observed which was marked at 16 h (Fig. 2.12).

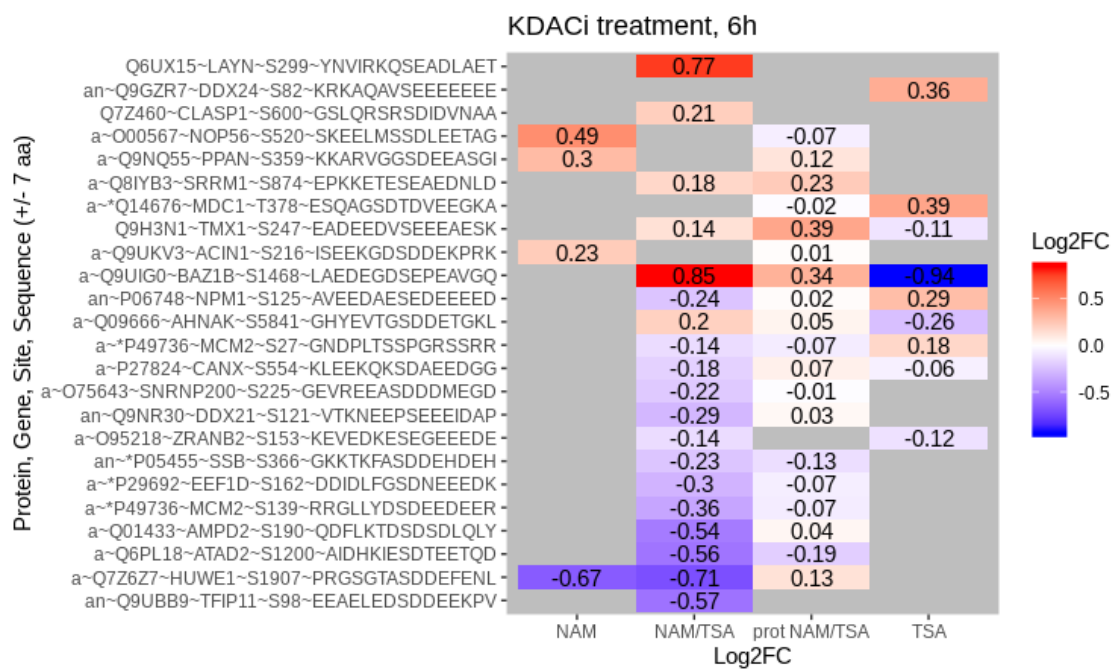


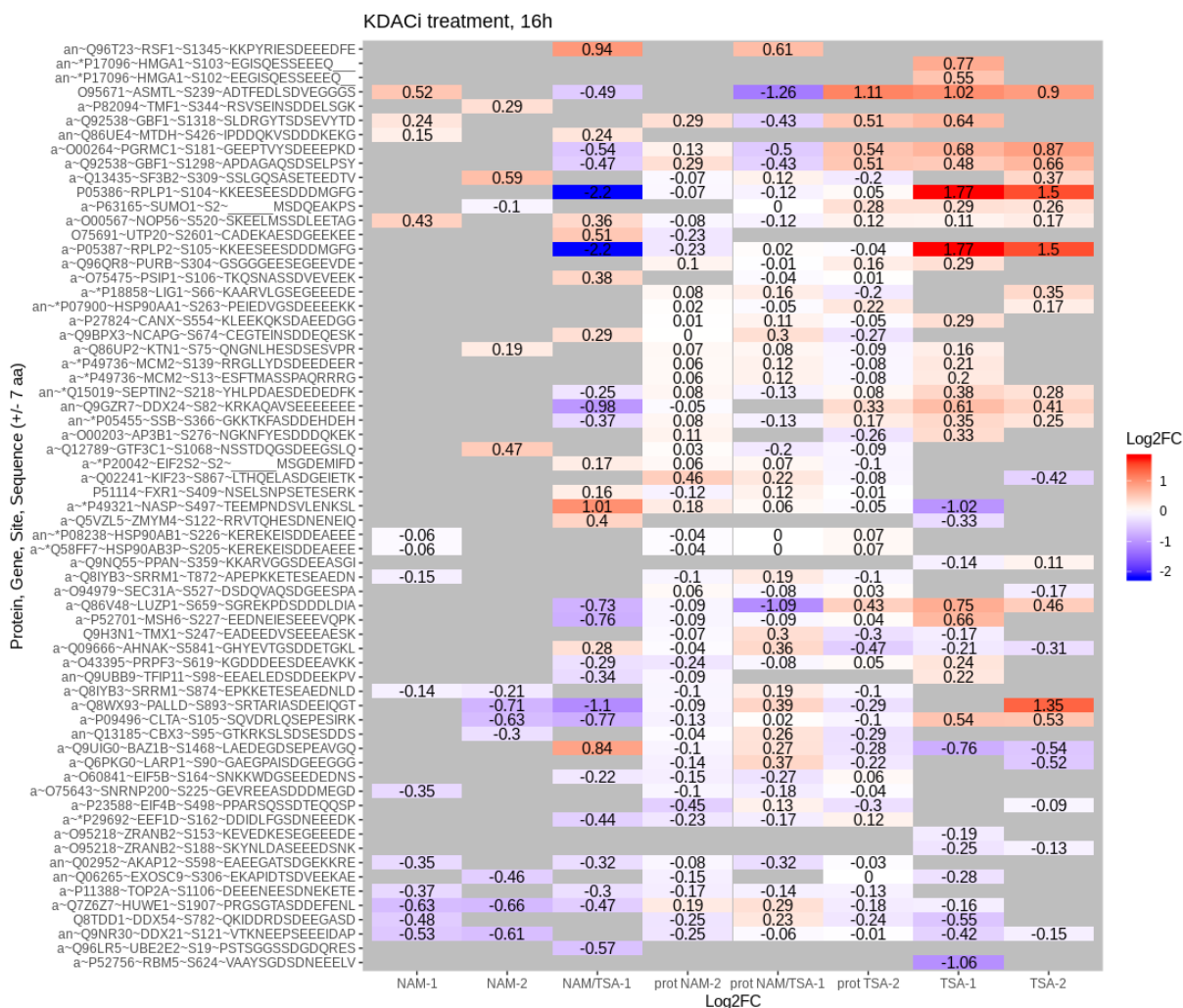


**Figure 2.12.** SILAC workflow of lysine deacetylase (KDAC) inhibitor treatment of U2OS cells. Increased acetylation levels in HeLa and U2OS cells after 16 h of treatment. Followed by triple/double-SILAC phosphoproteomic profiling of modified sites and identification of proteins functionally linked to CK2 whose phosphorylation was impacted at 16 h of treatment. Nicotinamide (NAM) and Trichostatin A (TSA).

Many known and tentative CK2 sites were quantified in three or more replicates out of four in U2OS cells treated with 2.5  $\mu$ M TSA or 20 mM NAM for 6 and 16 h (Fig. 2.13,

Supp. Table 2.15). As indicated, most of the sites studied belong to known acetyl-proteins (iPTMnet database) (Fig. 2.13, Supp. Table 2.15). Several of the quantified sites contain one or more known acetylation sites nearby, +/-20 amino acids (Fig. 2.13, Supp. Table 2.15).



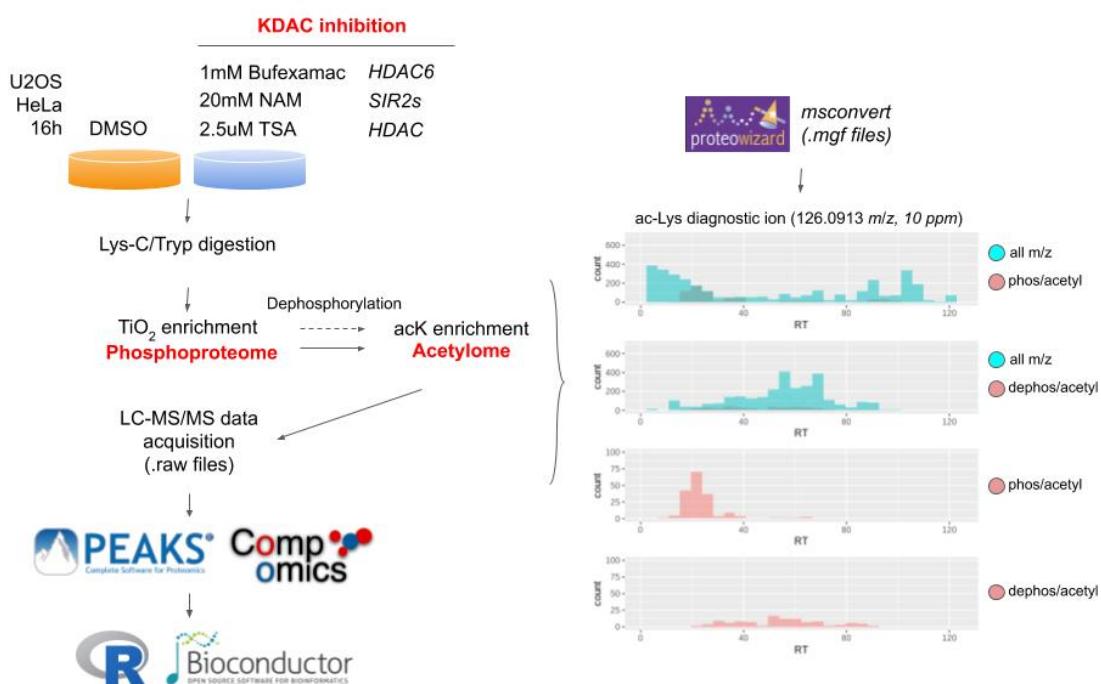


**Figure 2.13.** Known and tentative CK2 target sites and substrates quantified in U2OS cells treated with the lysine deacetylase inhibitors (KDACi) Trichostatin A (TSA) and Nicotinamide (NAM) for 6 (top) and 16 h (bottom) across different experiments. 6 h experiment, 20 mM NAM vs control condition (M/L triple-SILAC) and 2.5  $\mu$ M TSA vs control condition (H/L triple-SILAC). 16 h experiments, -1: 20 mM NAM vs control condition (M/L triple-SILAC) and 2.5  $\mu$ M TSA vs control condition (H/L triple-SILAC); -2: 20 mM NAM or 2.5  $\mu$ M TSA (M/L double-SILAC). The phosphosites and proteins were quantified if identified in three or more replicates (FDR adjusted p value < 0.05). \*: known CK2 site, a: acetyl-protein, n: nearby acetylation site (+/- 20 aa) to the phosphorylation site studied, prot: proteins, L: light label, M: medium label, and H: heavy label.

## 2.6.2 Co-occurrence of lysine acetylation and phosphorylation in neighboring residues

Next the acetylated fraction of the phosphoproteome was determined for finding K motif instances *in vivo* where both the phosphorylatable and acetylation site are modified. For

this the same unlabeled HeLa and U2OS protein samples collected after 16 h treatment with NAM, TSA, or Bx were analyzed using two different strategies: 1) phosphopeptide enrichment (phosphoproteome) followed by acetyl enrichment (acetyl fraction of the phosphoproteome); 2) phosphopeptide enrichment followed by dephosphorylation (dephosphorylated phosphoproteome) and acetyl enrichment (acetyl fraction of the dephosphorylated phosphoproteome) (Fig. 2.14, Supp. Table 2.16). After dephosphorylation of the samples we observed a shift towards higher retention times, which was indicative of a decrease in the polarity of the samples being analyzed by removal of the phosphate group (Fig. 2.14). Particularly, we observed this change in the subgroup of  $m/z$  spectra containing the  $m/z$  diagnostic ion for acetyllysine: 126.0913  $m/z$  [76,77] (Fig. 2.14), indicative that in the samples analyzed these peptides contained both phosphorylation and acetylation moieties.



**Figure 2.14.** Label-free workflow to study the acetylated fraction of the phosphoproteome.

### 2.6.2.1 Biological processes represented

To assess the validity of our experiments on providing relevant functional information regarding the acetylation of the phosphoproteome *in vivo*, the overlap between biological

processes represented in our study and the Scholz *et al.* 2015 was determined. Noticeably, an extensive overlap of many processes found represented in both studies was observed (Fig. 2.15, Supp. Table 2.17). Furthermore, like in Scholz *et al.* 2015 study the acetylation of HIF1A was observed in the Bx treatment. Expressly, acetylation of Lys251 in concert with phosphorylation of either Ser247 or Ser244 residues was observed. These sites localize to the second PAS domain of HIF1A, i.e., PAS-B, necessary for protein heterodimerization [78]. Phosphorylation of Ser247 by CK1 has been described in the literature and decreases HIF1A binding to HIF1B (ARNT) affecting HIF1A induced gene expression [78].

TSA treatment (BP: biological process)			NAM treatment (BP: biological process)			Bx treatment (BP: biological process)		
GO BP classification	U2OS	HeLa	GO BP classification	U2OS	HeLa	GO BP classification	U2OS	HeLa
Transcription	2	10	Transcription	3	5	Protein localization	28	24
RNA metabolism	32	45	RNA metabolism	36	47	Cytoskeleton	20	24
Gene expression	47	58	Gene expression	44	65	Microtubuli	10	11
Chromatin modification	10	20	Chromatin modification	20	15	Vesicle mediated transport	14	11
DNA metabolism	16	20	DNA metabolism	20	24	Cell cycle	17	20
Chromosome organization	14	27	Chromosome organization	21	22	Mitosis	13	24
Cytoskeleton organization	18	24	Cytoskeleton organization	20	33	Apoptosis	11	18
Apoptosis	8	15	Apoptosis	8	20	Metal	1	2
Mitosis	11	22	Mitosis	13	24	Autophagy	1	3

D

<p>Bufexamac treatment Scholz <i>et al.</i> 2015:</p> <ul style="list-style-type: none"> <li>• Increased levels of HIF1A</li> <li>• Acetylation of Lys709 of HIF1A</li> </ul>	<p>1mM Bx HeLa:</p> <p>HIF1A_HUMAN Hypoxia-inducible factor 1-alpha (Q16665): NH2-TFLS&lt;p&gt;RHSLDMK&lt;ace&gt;-COOH</p> <ul style="list-style-type: none"> <li>• acK: Lys251 confident, reported ubiquitinated (iPTMnet).</li> <li>• pST: Ser244 Doubtful; can be Ser247 (by CSNK1D, iPTMnet, PSP).</li> </ul>
---	---

**Figure 2.15.** Functional classification of phosphoproteins containing acetylation sites in HeLa and USOS cells, biological processes observed in common in Scholz *et al.* study and our study. A) Trichostatin A (TSA) treatment. B) Nicotinamide (NAM treatment). C) Bufexamac (Bx) treatment. D) Modulation of HIF1A modification in Scholz *et al.* 2015 study [24] and the Bufexamac treatment.

### 2.6.2.2 Histone code represented

Since the “histone code” is the most studied example of PTM interplay the sites identified were search for histone modifications as validation of the experimental method implemented. As expected, we observed several previously described histone modifications (Fig. 2.16). What is more, several unannotated acetylation and phosphorylation sites and their combination were also identified. This consistent with the increasing number of phosphorylation and acetylation sites reported to date in the databases and intriguing considering their co-occurrence and the lack of information available on this subject (Fig. 2.16).

Histone	PTM	Sequence	Source	Inhibitor
H1FX	Acetylation of K:195;Phosphorylation of T:189	NH2-T<p>AAAGGK<ace>K-COOH	UT	Q92522
H1FOO	Acetylation of K:281,287;Phosphorylation of T:297	NH2-K<ace>VVAKAK<ace>APKAGQGPNP<p>K-COOH	HT	Q81ZA3
H1FOO	Acetylation of K:142,148;Phosphorylation of T:158	NH2-K<ace>VVAKAK<ace>APKAGQGPNP<p>K-COOH	HT	Q81ZA3-2
HIST1H1A	Acetylation of K:137,138;Phosphorylation of S:136;Phosphorylation of T:133	NH2-AT<p>GAS<p>K<ace>K<ace>LKK-COOH	HT	Q02539
H1FX	Acetylation of K:146;Phosphorylation of S:154	NH2-K<ace>AAPGAAGS<p>RRADK-COOH	HT	Q92522
HIST1H1A	Acetylation of K:137;Phosphorylation of S:136;Phosphorylation of T:128	NH2-VAT<p>KTKATGAS<p>K<ace>K-COOH	HT	Q02539
HIST1H2BK	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	O60814
HIST1H2BJ	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	P06899
HIST1H2BO	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	P23527
HIST1H2BB	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	P33778
H2BFS	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	P57053
HIST1H2BD	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	P58876
HIST1H2BC	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	P62807
HIST2H2BE	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q16778
HIST2H2BF	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q5QNW6
HIST2H2BF	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q5QNW6-2
HIST3H2BB	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q8N257
HIST1H2BH	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q93079
HIST1H2BN	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q99877
HIST1H2BL	Acetylation of K:16;Phosphorylation of S:7	NH2-S<p>APAPKKGSK<ace>K-COOH	UN	Q99880
H1FX	Acetylation of K:23;Phosphorylation of S:33	NH2-VTK<ace>AGGSAALSPS<p>K-COOH	HN	Q92522
HIST1H1A	Acetylation of K:23;Phosphorylation of S:34	NH2-K<ace>AKKPAKAAAAS<p>K-COOH	HN	Q02539
HIST1H1A	Acetylation of K:131;Phosphorylation of S:136;Phosphorylation of T:128	NH2-VAT<p>KTK<ace>ATGAS<p>KK-COOH	HN	Q02539
H1FX	Acetylation of K:23;Phosphorylation of S:33	NH2-VTK<ace>AGGSAALSPS<p>K-COOH	UB	Q92522
H1FX	Acetylation of K:23;Phosphorylation of S:31	NH2-VTK<ace>AGGSAALS<p>PSK-COOH	UB	Q92522
HIST1H1A	Acetylation of K:137;Phosphorylation of T:133,128	NH2-VAT<p>KTKAT<p>GASK<ace>K-COOH	UB	Q02539
HIST1H1A	Acetylation of K:147;Phosphorylation of T:143	NH2-KAT<p>GASK<ace>KSVKTPK-COOH	UB	Q02539
H1FOO	Acetylation of K:327;Phosphorylation of S:335	NH2-TEAPKGPGRK<ace>AGLPIKAS<p>SSK-COOH	UB	Q81ZA3
H1FOO	Acetylation of K:188;Phosphorylation of S:196	NH2-TEAPKGPGRK<ace>AGLPIKAS<p>SSK-COOH	UB	Q81ZA3-2
HIST1H1T	Acetylation of K:131;Phosphorylation of S:138,141;Phosphorylation of T:132	NH2-SVSAK<ace>T<p>KKLVLS<p>RDS<p>K-COOH	UB	P22492
HIST1H1T	Acetylation of K:131,125;Phosphorylation of S:127,129;Phosphorylation of T:132	NH2-AK<ace>KS<p>VS<p>AK<ace>T<p>KK-COOH	UB	P22492
H1FX	Acetylation of K:23;Phosphorylation of S:33	NH2-VTK<ace>AGGSAALSPS<p>K-COOH	HB	Q92522
HIST1H1D	Acetylation of K:202,188;Phosphorylation of S:189	NH2-AAK<ace>S<p>PAKAKAPKPKAAK<ace>PK-COOH	HB	P16402
H1FX	Acetylation of K:23;Phosphorylation of S:27	NH2-VTK<ace>AGGS<p>AALSPSK-COOH	HB	Q92522
HIST1H1E	Acetylation of K:140,148;Phosphorylation of S:150;Phosphorylation of T:142	NH2-K<ace>AT<p>GAATPK<ace>KS<p>AKK-COOH	HB	P10412
HIST1H1E	Acetylation of K:21,17;Phosphorylation of T:4	NH2-ET<p>AAPAAPAPAEK<ace>TPVK<ace>K-COOH	HB	P10412
HIST1H1D	Acetylation of K:207,202,214;Phosphorylation of T:211	NH2-AAK<ace>PKSGK<ace>PKVT<p>KAK<ace>K-COOH	HB	P16402

**Figure 2.16.** Histone code observed. Red: histone modifications annotated in iPTMnet and PSP databases and Scholz *et al.* 2015 study [24].

### 2.6.2.3 Known modification sites, enzyme-substrate relationships, and functional classification

The known modifications annotated to the identified sites (including ambiguously localized sites) were retrieved from the iPTMnet and PSP database (Table 2.19) as validation of the experimental method implemented. Hundreds of known sites including

phosphorylation and acetylation but also other PTM types such as methylation and ubiquitination that overlap at the identified sites were also identified (Table 2.19). In addition, 29 enzyme-substrate relationships (iPTMnet database) and 9 regulatory sites (PSP database) corresponding to the identified sites were retrieved (Table 2.19, Supp. Table 2.18). Out of the 29 enzyme-substrate relationships 28 were kinase-substrate (19 kinases) and 1 acetyltransferase-substrate hits. The latter represented the autocatalytic lysine acetylation of EP300 lysine acetyltransferase at the Lys1560 residue (Supp. Table 2.18). This residue is in one of the most highly acetylated regions of EP300 and is likely responsive to SIRT2 RNAi-mediated depletion [29,79]. The kinase-substrate relationship retrieved indicated that three of the sites identified in the context of phosphorylation and acetylation regulate cell cycle, i.e., Ser1112 of RBL2 phosphorylated by CDK2 [80,81], S878 of KIF20A phosphorylated by AURKB [82], and S247 of HIF1A phosphorylated by CSNK1D [83]. Overall, the impact on the function of these regulatory sites on the corresponding proteins spanned regulation of molecular association, induction/inhibition of activity, change of protein conformation, alteration of receptor desensitization, receptor internalization, and protein stabilization (Supp. Table 2.18). What is more, modification at some of these sites was observed to affect protein-protein interaction (Supp. Table 2.18). For example, phosphorylation of Ser1112 of RBL2 disrupts the interaction with E2F4 [80,81]. Phosphorylation of Ser153 of PEBP1 disrupts interaction with RAF1 and induces interaction with GRK2 [84,85]. Phosphorylation of Ser140 of CHEK2 kinase induces interaction with PLK1 kinase [86]. Finally, as mentioned before, phosphorylation of HIF1A at Ser247 disrupts the interaction with ARNT [83].

**Table 2.19.** Known modifications annotated to the identified phospho-acetyl sites retrieved from iPTMnet (December 2019) and PSP (September 2019) databases.

Cell line	Trichostatin A (2.5 $\mu$ M)	Nicotinamide (20 mM)	Bufexamac (1mM)
U2OS (iPTMnet)	Phosphorylation: 70 Ubiquitination: 34 Acetylation: 17 Sumoylation: 5 Methylation: 3 O-Glycosylation: 1	Phosphorylation: 106 Ubiquitination: 27 Acetylation: 26 Methylation: 4 Sumoylation: 4 O-Glycosylation: 1	Phosphorylation: 120 Ubiquitination: 49 Acetylation: 27 Sumoylation: 7 Methylation: 4 O-Glycosylation: 2

HeLa (iPTMnet)	Phosphorylation: 141 Ubiquitination: 40 Acetylation: 13 Sumoylation: 4 Methylation: 1 O-Glycosylation: 1	Phosphorylation: 121 Ubiquitination: 59 Acetylation: 22 Methylation: 6 O-Glycosylation: 3 Sumoylation: 1	Phosphorylation: 182 Ubiquitination: 70 Acetylation: 42 Methylation: 8 Sumoylation: 5 O-Glycosylation: 3
Total (iPTMnet and PSP)	Acetylation: 141 Phosphorylation: 566 Methylation: 24 O-Glycosylation: 1 Sumoylation: 19 Ubiquitination: 248		
Enzymes	Kinases: AURKB, CDK2, CHEK2, NEK9, PRKACA, PRKCA, PRKCB, PRKCD, PRKCG, PRKCZ, STK4, IKBKE, CDK1, CDK12, CDK2, CHEK2, CSNK1D, MST1, and STK4 Lysine acetyltransferases: EP300		

The proteins containing the known sites that were found to co-occur with acetylation or phosphorylation were classified according to their GO molecular function, level 5 (Table 2.20, Supp. Table 2.17). Of notice, several writers, erasers, and readers associated with protein modification were observed including 33 Ser/Thr protein kinases and 4 lysine-acetylated histone binding. Importantly, all three subgroups of methylation writers, erasers, and readers were identified with 21 methyltransferases and three histone demethylases identified (Table 2.20). The latter highlights an important link between different modifications and their writers for example between acetylation and methylation which target Lys.

**Table 2.20.** GO molecular function classification, level 5, associated with writers, erasers, and readers of containing co-occurring phosphorylation and lysine acetylation sites.

GO ID	Description	Ratio	Genes
Methylation			
GO:0008168	methyltransferase activity	21/1097	ASH1L, CMTR1, KMT2A*, KMT2B, KMT2C*, KMT5C*, METTL18, METTL25, MRM3*, NSD2*, NSD3, NSUN3*, PRDM1, PRDM10, PRDM16, SETD1A, SETD2*, SETDB2, SMYD4*, TRMT1*, ZCCHC4
GO:0032454	histone demethylase activity (H3-K9 specific)	3/1097	JMJD1C, KDM3A*, PHF2*
GO:0042054	histone methyltransferase activity	11/1097	ASH1L, KMT2A*, KMT2B, KMT2C*, KMT5C*, NSD2*, NSD3, PRDM16, SETD1A, SETD2*, SETDB2



GO:0035064	methylated histone binding	10/1097	ATRX*, CBX8*, CDYL2, CHD1*, CHD5, CHD8*, PHF1, PHF2*, TDRD3*, ZMYND8*
GO:1990226	histone methyltransferase binding	1/1097	NOP56*
Acetylation			
GO:0070577	lysine-acetylated histone binding	4/1097	BRD4*, KMT2A*, PHIP*, ZMYND8*
GO:0035035	histone acetyltransferase binding	3/1097	HIF1A*, MYOCD*, SP1*
GO:0042826	histone deacetylase binding	9/1097	CHD4*, HIF1A*, MYOCD*, NIPBL*, PKN2, SP1*, TOP2B, UHRF1BP1, WDTC1
Phosphorylation			
GO:0004674	protein serine/threonine kinase activity	33/1097	ACVR1*, AKAP13*, ALPK2*, BMPR2, BRD4*, CDC42BPB*, CDK12*, CHEK2*, CSNK1A1, CSNK1G1, CSNK1G2, CSNK1G3, LRRK1, MAP3K10, MAP4K3*, MAST1*, NEK5, NEK9*, PKN2, PRKAA1*, PRKG1, PRPF4B*, RPS6KA1*, RPS6KA2, SLK*, SRPK1, STK10*, STK33*, TAOK1*, TLK1, TOP1*, TTN*, WNK1
GO:0004713	protein tyrosine kinase activity	7/1097	BAZ1B*, EPHB2, FER, FGFR1, IGF2R, KDR*, TTN*
GO:0035173	histone kinase activity	2/1097	BAZ1B*, PRKAA1*

\*: Genes with known modification sites co-occurring with acetylation and phosphorylation.

Along with several enzymes found targeted by acetylation and phosphorylation additional subunits of protein complexes linked to transcriptional regulation and chromatin organization were also observed dually modified (Table 2.21, Supp. Table 2.17). A summary of CK2 regulatory roles on these complexes, and chromatin organization in general, can be found in Chapter 3.

**Table 2.21.** Protein complexes represented among the proteins containing the identified phospho-acetyl sites.

Description	No.	Gene	Keyword
GO:0005667: transcription factor complex	11	ARNT2, EP300*, ERCC4, GTF3C6, HIF1A*, NCOA6*, RB1*, SKOR1, TAF2, TAF3, ZFH3	Transcriptional regulation
GO:0035097: histone methyltransferase complex	11	C17orf49, CHD8*, HCFC2, JARID2, KANSL1*, KMT2A*, KMT2B,	Histone methylation

		KMT2C*, NCOA6*, PHF1, SETD1A	
GO:0044798: nuclear transcription factor complex	6	ERCC4, GTF3C6, HIF1A*, RB1*, TAF2, TAF3	Transcriptional regulation
GO:0055029: nuclear DNA-directed RNA polymerase complex	6	CDC73, ERCC4, POLR3A*, POLR3GL*, TAF2, TAF3	Transcription
GO:0000428: DNA-directed RNA polymerase complex	6	CDC73, ERCC4, POLR3A*, POLR3GL*, TAF2, TAF3	Transcription
GO:0008278: cohesin complex	5	CDCA5*, SGO2*, SMC1A*, SMC3*, STAG1	Chromatin cohesion
GO:0031519: PcG protein complex	5	CBX8*, JARID2, PHC1*, PHF1, RYBP*	Transcriptional repression
GO:0000123: histone acetyltransferase complex	5	DMAP1*, EP300*, JADE3, KANSL1*, TAF2	Histone acetylation
GO:0016581: NuRD complex	3	CHD4*, CHD5, MTA3	Transcriptional repression
GO:0031391: Elg1 RFC-like complex	2	ATAD5*, RFC1	Telomere length regulation
GO:0072357: PTW/PP1 phosphatase complex	1	PPP1R10*	Chromatin structure control, cell cycle progression
GO:0009330: DNA topoisomerase complex (ATP-hydrolyzing)	1	TOP1*	Supercoiling and torsional tension of DNA release
GO:0034991: nuclear meiotic cohesin complex	1	SMC3*	Chromatin cohesion
GO:0031011: Ino80 complex	1	INO80*	Chromatin remodeling
GO:0070522: ERCC4-ERCC1 complex	1	ERCC4	Nucleotide-excision repair
GO:0005663: DNA replication factor C complex	1	RFC1	PCNA to DNA loading
GO:0031390: Ctf18 RFC-like complex	1	CHTF18*	PCNA to DNA loading

GO:0033503: HULC complex	1	RNF20*	Histone ubiquitination
GO:0000347: THO complex	1	THOC2*	Transcription to cytoplasmic mRNA export coupling
GO:0005697: telomerase holoenzyme complex	1	SMG6*	Telomere elongation
GO:0030870: Mre11 complex	1	SP100	Meiotic recombination, DNA repair and checkpoint signaling
GO:0032116: SMC loading complex	1	NIPBL*	Chromosome (SMC) complex to DNA loading
GO:0071204: histone pre-mRNA 3'end processing complex	1	YBX1*	Histone pre-mRNAs processing

\*: Genes with known modification sites co-occurring with acetylation and phosphorylation.

### 2.6.3 *In vivo* phospho-acetyl K motif hits

All the bioinformatic analyses conducted pointed to the acetylation and phosphorylation of the K motif variants *in vivo*. After performing dual enrichment and functionally validating the sites identified the corresponding window sequences were scanned against the CK2 consensus and the K motif variants (Table 2.22) to identify hits that occur *in vivo* in HeLa and U2OS treated with NAM, TSA, and Bx. As a result, several phospho-acetyl peptides were observed matching the linear patterns (Table 2.22, Supp. Table 2.19). No sites were found matching the +1 extended variant pattern either phosphorylated and/or acetylated at the motif (Table 2.22). Conversely, several hits were observed for +2 and +3 variants both for the K motif and K motif extended patterns (Table 2.22, Supp. Table 2.19). Besides K motif hits a number of sites matching the CK2 consensus and the CK2 consensus extended were observed targeted by phosphorylation and in the vicinity of an acetylation site (Table 2.22, Supp. Table 2.19); 13 of these sites were found annotated as phosphorylated in the PSP and iPTMnet databases (Supp. Table 2.19) and three sites i.e. Ser520 of NOP56 (ribosomal subunit biogenesis), Ser102 of SLIRP (nuclear receptor corepressor), and Ser381 of CWC22 (component of the spliceosome), were found phosphorylated *in vitro* by CK2 [49]. The sites discovered here are novel candidates for studying the interplay between CK2-dependent phosphorylation

and lysine acetylation. For example, the ATP-dependent RNA helicase DDX18, has a role in ESC self-renewal by antagonizing the polycomb repressor complex PRC2 maintaining the integrity at actively transcribed rDNA loci [87]. The residue Ser136 of DDX18, matches the CK2 consensus (SEEE); the neighboring residue Lys129 was found acetylated. Another interesting candidate is PPP1R10 (PNUTS) a scaffold protein of the PTW/PP1 phosphatase complex which functions in chromatin organization. The residue Ser124 of PPP1R10 matches the extended CK2 consensus (SEDEE) and it is in the region of interaction with TOX4 a core member of the complex and in the TFIIS N-terminal domain. The neighboring sites K122, K118, and K131 could be targets of acetylation. Finally, among the hits, three CK2 interactors were found: HDGFL2, NOP56, and SLIRP; and six proteins that belong to the CK2 neighborhood: BDP1, ATRX, CHD5, POLR3A, CORIN, and ALPL (Supp. Table 2.19). The proteins identified mentioned in the literature with CK2 were BDP1, DNMT1, and PPFIBP1 (Supp. Table 2.19).

**Table 2.22.** Phosphorylated and acetylated sites that conform to CK2 consensus and K motif hits identified *in vivo* in peptides with co-occurring phosphorylation and acetylation; includes canonical and isoform sequences.

Group	Type	Pattern	No. sites
CK2	CK2 minimal	^.{7}[ST][^PKR].[DE]	555
	CK2 consensus	^.{7}[ST][DE].[DE]	100
	CK2 extended	^.{7}[ST][DE][DE][DE]	14
	CK2 extended	^.{7}[ST][DE][DE][DE][DE]	4
K motif	K: K motif, p+1	^.{6}[ST]K[DE][DE]	14
	K: K motif, p+2	^.{5}[ST][DE]K[DE]	25
	K: K motif, p+3	^.{4}[ST][DE][DE]K	15
	K: K motif extended, p+1	^.{6}[ST]K[DE][DE][DE]	0
	K: K motif extended, p+2	^.{5}[ST][DE]K[DE][DE]	3
	K: K motif extended, p+3	^.{4}[ST][DE][DE]K[DE]	3
	S/T: K motif, p+1	^.{7}[ST]K[DE][DE]	18
	S/T: K motif, p+2	^.{7}[ST][DE]K[DE]	24
	S/T: K motif, p+3	^.{7}[ST][DE][DE]K	13

	S/T: K motif extended, p+1	$\wedge.\{7\}[ST]K[DE][DE][DE]$	0
	S/T: K motif extended, p+2	$\wedge.\{7\}[ST][DE]K[DE][DE]$	3
	S/T: K motif extended, p+3	$\wedge.\{8\}[ST][DE][DE]K[DE]$	4

The function of the protein containing phosphorylated and acetylated sites in sequences matching the K motif variants was investigated (Table 2.23). As well as the functional features of the corresponding proteins (Table 2.23, Supp. Table 2.19). Interestingly, several K motif hits were identified containing known modification sites annotated in PSP and/or iPTMnet databases.

**Table 2.23.** Co-occurring phosphorylation and acetylation sites *in vivo* at K motif variants (the table includes ambiguously localized sites).

Data set	Sites	Function (UniProtKB)
p +2		
UT	CYLC2: T213*;K215 (TEKD)	Structural constituent of cytoskeleton
	CYLC2: T308;K310* (TEKE)	
	FAM133A: T167;K169 (TEKE)	-
HT	ANO1: S235;K237 (SDKD)	Calcium-activated chloride channel, modified sites are on cytoplasmic domain (IPR031287)
	AZI2: S55;K57; (SEKE)	Activates serine/threonine-protein kinase TBK1 and facilitates its oligomerization. Enhances the phosphorylation of NF-kappa-B p65 subunit RELA by TBK1.
UN	SAMSN1: S10;K12 (SEKE)	Immune response stimulates HDAC1 activity (20478393).
HN	CHD5: S1460;K1462 (SEKE)	Chromatin-remodeling protein part of a NuRD-like, complex composed at least of GATAD2B, HDAC1, HDAC2 and MTA3.
	DOPEY1: S1281;K1283 (SEKE)	Traffic between late Golgi and early endosomes.
UB	PPFIBP1: T547*;K549* (TEKE)	Disassembly of focal adhesions.
	LHX5: S162;K164 (SDKE)	DNA-binding transcription factor activity, RNA polymerase II-specific. Regulation of neuronal differentiation.
HB	NOL8: S1008;K1010 (SEKEE)	Regulation of gene expression at the post-transcriptional level or in ribosome biogenesis in cancer cells.
	EXOSC1: T111;K113* (TEKD)	Exonucleolytic catabolism of deadenylated mRNA and rRNA biogenesis.

	MAP7D3: T181;K183 (TDKE)	Promotes the assembly and stability of microtubules.
	POLR3A: S63;K65* (SEKD)	Largest and catalytic core component of RNA polymerase III which synthesizes small RNAs, such as 5S rRNA and tRNAs.
p +3		
HB	ANKRD36: S1101;K1104 (SDEKD)	Interactor of IKZF1 (Ikaros).
HN	BOD1L1: S1040;K1043 (SDDKD)	Cellular response to DNA damage stimulus.
HT	POLR3GL: T141;K144 (TEDKEE)	DNA-dependent RNA polymerase catalyzes transcription by RNA polymerase III.

U: U2OS, H: HeLa, T: Trichostatin A, N: Nicotinamide, B: Bufexamac. \* known modification sites.

The boundaries and cellular properties of the K motif hits were searched for information provided in the SLiMSearch4 analysis (Table 24). This revealed that several of the sites identified are conserved across Metazoa and QFO and that the sites do not have the propensity to fold upon binding. The rest of the hits identified *in vivo* were not returned by SLiMSearch4 possibly due to their IUPred score not meeting the 0.40 cutoff selected as default to perform the analysis.

**Table 2.24.** K motif hits boundaries and cellular properties determined by SLiMSearch4.

Sites	Description
CYLC2: T213 (TEKD)	IUPred: 0.579; Anchor: 0.046; SA: -1; QFO: 0.4286048, 0.27975; Metazoa: 0.5391182, 0.30075; Phospho: Yes; Flank vs motif: 0 (Metazoa) and 0 (QFO)
FAM133A: T167 (TEKE)	IUPred: 0.571; Anchor: 0.061; SA: -1; QFO: 0.5975777, 0.22975; Metazoa: 0.7148451, 0.167; Phospho: No; Flank vs motif: -0.05 (Metazoa) and -0.15 (QFO)
SAMSN1: S10 (SEKE)	IUPred: 0.817; Anchor: 0.333; SA: -1; QFO: 0.7115836503, 0.259; Metazoa: 0.9574487466, 0.149; Phospho: No; Flank vs motif: 0 (Metazoa) and 0 (QFO)
DOPEY1: S1281 (SEKE)	IUPred: 0.456; Anchor: 0.217; SA: -1; QFO: 0.218416177, 0.199; Metazoa: 0.9979192631, 0.01325; Phospho: No; Flank vs motif: 0 (Metazoa) and 0 (QFO)
PPFIBP1: T547 (TEKE)	IUPred: 0.62; Anchor: 0.34; SA: -1; QFO: 0.172981865, 0.35475; Metazoa: 0.4867627347, 0.2425; Phospho: Yes; Flank vs motif: 0.2 (Metazoa) and 0 (QFO)
LHX5: S162 (SDKE)	IUPred: 0.788; Anchor: 0.105; SA: -1; QFO: 0.539949871, 0.2535; Metazoa: 0.9918983853, 0.01025; Phospho: No; Flank vs motif: 0 (Metazoa) and -0.1 (QFO)
NOL8: S1008 (SEKEE)	IUPred: 0.568; Anchor: 0.093; SA: -1; QFO: 0.7633968875, 0.21808; Metazoa: 0.8420275525, 0.13712; Phospho: No; Flank vs motif: -0.05 (Metazoa) and -0.05 (QFO)
ANKRD36: S1101 (SDEKD)	IUPred: 0.446; Anchor: 0.465; SA: -1; QFO: 1, 0; Metazoa: 1, 0; Phospho: No
BOD1L1: S1040 (SDDKD)	IUPred: 0.612; Anchor: 0.095; SA: -1; QFO: 0.300831324, 0.316; Metazoa: 0.647191683, 0.326; Phospho: No; Compositionally biased region:Lys-rich (315-1046)

POLR3GL: T141 (TEDKEE)	IUPred: 0.618; Anchor: 0.057; SA: -1; QFO: 0.9368245236, 0.11975; Metazoa: 0.854150482, 0.201; Phospho: No; Domain: RNA_pol_3_Rpc31 (1-218); Compositionally biased region: Glu-rich (142-191)
------------------------	--

## 2.7 Methods and Materials

**CK2 substrates.** The known CK2 substrates were downloaded from the database PhosphoSitePlus (PSP) [46] (September 2019) by querying substrates of: “CK2a1”, “CK2a2”, “CK2b”. The CK2 substrates annotated in PSP matching the K motif were manually curated to remove entries with not enough supporting evidence including sites identified as substrates in multiply phosphorylated sequences and *in vivo* sites changing upon treatment with non-specific CK2 inhibitors. The known CK2 substrates were downloaded from the database iPTMnet [57] (September 2019) by querying: “csnk2a1”, “csnk2a2”, “csnk2b”. The data sets were combined to generate a list of unique substrates (see Appendix C).

**PTM interplay.** The within protein post-translational modifications (PTM) interplay data were retrieved from a local copy of the database PTMcode v2 (October 2019), containing 15,988,098 intra-protein PTM associations. The overlap between the known CK2 sites on PSP and the interplaying sites on the database PTMcode v2 [47] at a distance of +/-4 and +/-7 amino acids was determined. Potential phospho-acetyl interplaying sites were also determined by overlapping K motif hits information retrieved from the iPTMnet database (phospho-Ser/Thr, see PTM data retrieval below) and Scholz *et al.* 2015 (acetyl-Lys, see Acetyl-protein dataset below) study on the acetylome. Since the Lys site represented in the K motif can harbor other modification types, tentative interplay information for these modifications (e.g., phospho-ubiquitination or phospho-methylation) was also retrieved (see PTM data retrieval below). All data processing and analysis steps were performed programmatically using R (see Appendix C).

**PTM data, kinase-substrate relationship, and regulatory sites retrieval.** I retrieved the PTM information and kinase-substrate relationships programmatically from the iPTMnet database using R’s package iptmnetr v0.1.8 [88]. Information was also downloaded from the PSP database (September 2019) and accessed locally using bash and R (see Appendix C).

**Protein functional classification.** I retrieved the protein functions for the CK2 substrates and other proteins of interest from the UniProt database [60]. GO classification/enrichment (GO BP, GO CC, and GO MF) [89] was performed using Bioconductor's packages clusterProfiler v3.14.3 [90] and org.Hs.eg.db v3.10.0 [91] using human genome as background (see Appendix C). GO classification/enrichment (GO BP, GO CC, and GO MF) was also performed using DAVID v6.8 [92]; the human genome was set as background.

**In solution peptide synthesis.** Peptides containing the K motif (Table 2.25) were synthesized by the group of Dr. Shawn Li as described in St-Denis *et al.* 2015 [18]. A Trp residue was added to the N-terminal for peptide quantification at 280 nm. Three Arg residues were added to the N-terminal for binding to P81 Ion Exchange Cellulose Chromatography Paper (Reaction Biology Crop., IEP-01; VWR, 21426-202; Fisher Scientific, 05-717-2B (WHT3698325)). The CK2 peptide substrate RRRDSDSD or DSD (EZBiolabs [21]) was used as control. The quality of the peptides was assessed by MALDI Mass Spectrometry (London Regional Proteomics Centre).

**Table 2.25.** K motif in solution peptides.

Sample ID	MW	Sequence	Type
D12	1690.79	WRRRADETSEKEQ	K p+2, SET-iso 2 Thr23/Ser24
E1	1732.79	WRRRADETSE <sup>x</sup> EQ	ac-K p+2, SET-iso 2 Thr23/Ser24
E2	1701.86	WRRRVNDVSEKEQ	K p+2, Ser33 CTTN
E3	1743.86	WRRRVNDVSE <sup>x</sup> EQ	ac-K p+2, Ser33 CTTN
E10	1649.65	WRRRADDSDKDDD	K p+1
E11	1649.65	WRRRADDSDKDD	K p+2
E12	1649.65	WRRRADDSDDKD	K p+3
F1	1691.65	WRRRADDSD <sup>x</sup> DDD	ac-K p+1
F2	1691.65	WRRRADDSD <sup>x</sup> DD	ac-K p+2
F3	1691.65	WRRRADDSD <sup>x</sup> D	ac-K p+3

**In vitro phosphorylation of in solution peptides.** The synthesized K motif peptides were dissolved in an adequate solvent (DMSO or distilled water) according to the peptide charge and quantified using a NanoDrop. A 2.5 mM peptide stock was prepared for each



peptide. The radioactive CK2 kinase assays were carried as indicated in Turowec *et al.* 2010 [93] for 10 minutes in triplicate for peptides concentrations of 200 and 500  $\mu\text{M}$ . The reactions were started by adding 1:5,000 final dilution of untagged CK2 $\alpha$  (specific activity: 424.3 nmol/min/mg, 19.15 mg/ml). Reactions were carried out by adding 2  $\mu\text{L}$  of 5x Kinase Activation Buffer (250 mM Tris, pH 7.5, 750 mM NaCl, 50 mM  $\text{MgCl}_2$ , 500  $\mu\text{M}$  ATP) spiked with 0.2  $\mu\text{Ci}/\mu\text{L}$   $\gamma$ - $^{32}\text{P}$ -ATP (3000 Ci/mmol, PerkinElmer BLU002A250UC), 6  $\mu\text{L}$  of peptide, and 2  $\mu\text{L}$  of kinase in CK2 dilution buffer (1 mg/ml bovine serum albumin, 5 mM MOPS, pH 7.0, 200 mM NaCl) for a total reaction volume of 10  $\mu\text{L}$ . Negative control reactions were performed by adding water instead of peptide substrate, 5x Kinase Activation Buffer, or kinase. Reactions were stopped by spotting 4  $\mu\text{L}$  of the reaction onto a 1 x 1 cm squares of P81 Ion Exchange Cellulose Chromatography Paper followed by 4 x 5 min washes in a shaker with 1 % phosphoric acid (at least 10 ml of 1% phosphoric acid/paper) and 1 x 5 min in a shake with 95 % ethanol. The papers were dried for 15 min under a heat lamp and exposed for 20 min to a PhosphorImager screen followed by scanning in a STORM Phosphorimager (Molecular Dynamics) and quantification using ImageQuant TL software (Amersham Biosciences). Quantification of incorporation of  $\gamma$ - $^{32}\text{P}$  was enabled by spotting a serial dilution of the 5x Kinase Activation Buffer as standards.

**K motif linear patterns.** Several linear patterns were deduced from the peptide phosphorylation data, represented as regular expressions or ScanProsite patterns, and used for scanning PTM data, phosphoproteome and acetylome sequence window information, and the human proteome, for finding relevant hits. Example of regular expressions: [ST][DE]K[DE] or [ST][^PRK]K[DE]; example of ScanProSite: [ST]{^PRK}K[DE].

**Cell culture.** U2OS (human osteosarcoma) and HeLa (cervix adenocarcinoma) cells were cultured in 10 or 15 cm dishes (TPP, FroggaBio) at 5 %  $\text{CO}_2$  in Dulbecco's modified eagle medium (DMEM), supplemented with 10 % Fetal Bovine Serum (FBS, HyClone), 100 U/ml Penicillin, and 0.1 mg/mL Streptomycin. Cells were maintained with 2.5 mg/mL plasmocin (InvivoGen, ant-mpp) to prevent mycoplasma contamination and tested for mycoplasma using the MycoAlert<sup>TM</sup> Assay Control Set (Lonza).

**Cell treatment and lysis.** For immunoblotting and immunoprecipitation assays the cells were grown up to 80-90% confluency and treated with lysine deacetylase inhibitors for different periods of time as indicated. After treatment, the culture media was removed, and the culture dishes were washed twice with ice-cold Phosphate Buffered Saline (PBS). Cells were harvested in 300  $\mu$ L-1 mL RIPA (radioimmunoprecipitation assay buffer) lysis buffer (150 mM NaCl, 1.0 % NP-40, 0.5 % sodium deoxycholate, 0.1 % SDS, 50 mM Tris pH 7.5) supplemented with protease (1 mM PMSF, 7  $\mu$ g/mL PepA, 20  $\mu$ g/mL Leupeptin, 2.9  $\mu$ g/mL Aprotinin) and phosphatase (1  $\mu$ M Okadaic acid, 1  $\mu$ M Microcystin, 5 mM NaF, 1 mM Sodium orthovanadate) inhibitors. The cells were scraped from the culture plate or lifted using (PBS)-EDTA. The cell suspension was incubated on ice for 30 min vortexing every 5 min. Samples were spun at 15,000 g for 15 or at 3,500 for 30 min at 4 °C. Protein concentration was determined by BCA. For immunoblotting 10  $\mu$ g was loaded onto a 10% SDS-PAGE gel for separation and transfer to a PVDF membrane. For detections with pan-antibodies between 25-50  $\mu$ g of protein lysate was loaded into the gel. Unused lysates were frozen at -80 °C until further use. At least three biological replicates were obtained for each experiment.

**Lysine deacetylase inhibitors.** Nicotinamide (Sigma, N0636), Trichostatin A (Sigma, T8552), Tubacin (Sigma, SML0065), Bufexamac (Sigma, B0760), Tubastatin A (Selleckchem, S8049). The inhibitors were used at concentrations: 20 mM Nicotinamide (NAM) in distilled water, 2  $\mu$ M Trichostatin A (TSA) in DMSO, 10  $\mu$ M Tubacin in DMSO, 1 mM Bufexamac DMSO, and 10  $\mu$ M Tubastatin A in distilled water.

**Antibodies.** The antibodies used and the corresponding experimental details are shown in Table 2.26.

**Table 2.26.** Antibodies used.

Antibody	Description	Experimental details
<b>Primary antibodies</b>		
pSer2 EIF2S2	Rabbit polyclonal.	1:10,000 dilution in 3 % BSA-TBST. Immunogen: ac-pS-GDEMIFDPTMSKC-amide peptide
Total EIF2S2	Mouse monoclonal (clone 2F3, Novus).	1:500 dilution in 3 % BSA-PBST

pSer2/3/8 CSNK2B	Rabbit polyclonal.	1:10,000 dilution in 3 % BSA-TBST. Immunogen: ac-pS-pS-pS-EEVSMISWFC-amide and ac-pS-pS-pS-EEV-pS-MISWFC-amide peptides, 1:1.
pSer162 EEF1D	Rabbit polyclonal (YenZym Antibodies, LLC, San Francisco, CA).	1:20,000 dilution in 3 % BSA-TBST. Fraction R1p1.
Total EEF1D	Rabbit polyclonal (YenZym Antibodies, LLC, San Francisco, CA).	1:6,000 dilution in 3 % BSA-TBST. Cross-reactive fraction from pSer162 EEF1D immunization. Fraction R2p2.
Total CK2 (CSNK2A1 and CSNK2A2)	Rabbit polyclonal (BAbCO, 7592773)	1:2,000 dilution in 3 % BSA-TBST. Immunogen: GST-Fusion CK2a
Total GAPDH	Mouse monoclonal (clone 6C5, Millipore)	1:1,000 dilution in 1 % BSA-PBST.
Total PARP (full length: 116 kDa; large fragment: 89 kDa)	Rabbit polyclonal (#9542, CST)	1:1,000 dilution in 3 % BSA-TBST.
Total CTTN	Mouse monoclonal (clone H-5, sc-5579)	1:200 dilution in 3% BSA-PBST, immunoprecipitation [1-2 µg per 100-500 µg of total protein (1 ml of cell lysate)]
<b>Primary antibodies, pan-specific</b>		
Phospho-CK2 Substrate (pS/pT)DXE.	Rabbit polyclonal mix (MultiMab™ #8738, CST)	1:1,000 dilution in 3 % BSA-TBST.
Acetylated-Lysine Antibody	Rabbit polyclonal (#9441, CST)	1:1,000 dilution in 3 % BSA-TBST. Immunogen: synthetic acetylated lysine-containing peptide.
Acetyl Lysine Antibody	Rabbit polyclonal (ICP0380, immunechem)	1:1,000 dilution in 3 % BSA-TBST. Immunogen: Acetylated KLH conjugates.
<b>Secondary antibodies</b>		
IRDye® 800CW Goat anti-Rabbit	Goat (926-32210, LICOR).	1:10,000 dilution in 3 % BSA-TBST or LICOR blocking buffer. Immunogen: Rabbit IgG.
IRDye® 680RD Goat anti-Mouse IgG (H + L)	Goat (926-68070, LICOR)	1:40,000 dilution in 1% BSA-PBST or LICOR blocking buffer. Immunogen: Mouse IgG paraproteins

**Immunoblotting.** After SDS-PAGE separation the cassette(s) was assembled, and the transfer performed at 0.4 mA for 1 h. The run was stopped, and the membrane(s) air-dried, and then re-wetted in methanol for 1 min. Blocking was performed using the LICOR blocking buffer or BSA-TBS/PBS for 1 h at room temperature. After 3 washes with TBST/PBST five minutes each, the primary antibody was added in the correct dilution on the blocking buffer and incubated for 1 h at room temperature or overnight at 4 °C. After incubation with the primary antibody, the washes were repeated, and the corresponding secondary antibody was added in the blocking buffer and incubated for 1

h. After incubation with the secondary antibody, the washes were repeated with a final wash with TBS/PBS. The membranes were scanned using a LiCor Odyssey Infrared Imaging System managed by the Odyssey v3.0 software. Quantification was performed by using the up-bottom background method in Odyssey v3.0 software. Lane normalization was assessed either by detection with the total GAPDH antibody or by staining with Revert 700 Total Protein Stain for Western Blot Normalization (LiCor). Lane normalization was performed as indicated by LiCor in the REVERT protocol.

**Src substrate cortactin (CTTN) immunoprecipitation.** U2OS cell lysates were prepared in a RIPA buffer supplemented with protease and phosphatase inhibitors. Then increasing amounts (2, 4, 8, 16  $\mu$ g) of the mouse monoclonal H-5 antibody (see antibody table) were tested against 1 mg of total U2OS protein lysate. The lysates were mixed with the corresponding amount of H-5 antibody and incubated overnight at 4 °C to allow for CTTN-H-5 complex formation. Next the CTTN-H-5 complex was precipitated by incubation with 30  $\mu$ l of A/G UltraLink beads (Thermo. Prod #53132, Lot # NH174763) for 1 h at 4 °C followed by centrifugation at 3,000 x g for 2 min at 4 °C. For this step, the beads were pre-conditioned in PBS overnight and washed 2 x in PBS with centrifugation at 3,000 x g for 2 min at 4 °C in between. Next, the beads were precipitated by centrifugation at 3,000 x g for 2 min at 4 °C, the supernatant (unbound fraction) was discarded and the precipitate was washed four times in 1 mL lysis buffer, with centrifugation in between each wash. After the final wash step for protein elution the beads were resuspended in 15  $\mu$ L of lysis buffer and 15  $\mu$ L of 2x concentrated Laemmli sample buffer to a total of 30  $\mu$ L and heated at 95 °C for 5 min. 25% of the samples were loaded onto a 10 % SDS-PAGE gel and separated. The gel was stained with Gelcode Blue 1.5 h after fixing it with 50:10:40 MetOH:AceticAcid:MilliQ-H<sub>2</sub>O (30 min). The gel was stored in 5 % Acetic acid at 4 °C until processing for in-gel digestion. After optimization of the amount of protein lysate required, U2OS cell lysates expressing CK2-WT (24 h of induction and 16 h of treatment with 1 mg/ml of tetracycline) treated or not for 16 h with 20 mM NAM were subjected to CTTN immunoprecipitation. The lysates were pooled together to achieve 2.4 mg of total protein and the resulting sample was immunoprecipitated with 38  $\mu$ g of H-5. The immunoprecipitation was performed as described above except that the bound CTTN was eluted in 60  $\mu$ L of buffer.

**In-gel digestion.** Sample picking and processing was performed at the Functional Proteomics Facility (London Regional Proteomics Centre). As a note, in-gel protein digestion was performed both with chymotrypsin and trypsin to increase peptide coverage.

**CK2 manipulation datasets.** Several phosphoproteomic data sets relevant to CK2 manipulation were reused from the literature and from our group. **1)** Global screening of *in vitro* CK2 substrates by an integrated phosphoproteomic workflow by Bian *et al.* 2013 [49] (see their Supp. Table 1). Control reaction: light dimethyl label; CK2 reaction: heavy dimethyl label. CK2 substrates: fold change  $\geq 1.5$ . **2)** SILAC (Stable Isotope Labeling by/with Amino acids in Cell culture) phosphoproteomic study by Rusin *et al.* 2017 [51]. Light cells: HeLa cells arrested in mitosis and treated with 10  $\mu\text{M}$  MG132 for 30 min then treated with DMSO for 45 min. Heavy cells: same as light cells but treated with 5  $\mu\text{M}$  CX-4945 for 45 min instead of DMSO [51]. Site localization was selected for phosphoRS score  $\geq 0.75$ . Differentially modulated sites were selected with a log<sub>2</sub> fold change cut-off  $\leq -1$ . **3)** SILAC time course phosphoproteomic study by Rabalski [52]. Light: HeLa cells treated for 1, 2, 3, and 4 h with DMSO, respectively. Heavy: HeLa cells treated for 1, 2, 3, and 4 h with 20  $\mu\text{M}$  CX-4945, respectively. Site localization was selected Localization prob  $\geq 0.75$ . Differentially modulated sites were selected with a log<sub>2</sub> fold change cut-off  $\leq -1$ . **4)** SILAC phosphoproteomic study by Cruise (<https://ir.lib.uwo.ca/etd/5164>). Light: U2OS cells over-expressing CSNK2A1-WT treated for 4 h with DMSO. Medium: U2OS cells over-expressing CSNK2A1-WT treated for 4 h with 30  $\mu\text{M}$  CX-4945. Heavy: U2OS cells over-expressing CSNK2A1-TM (V66A/H160D/I174A) treated for 4 h with 30  $\mu\text{M}$  CX-4945. Site localization was selected Localization prob  $\geq 0.75$ . Differentially modulated sites were selected from with a log<sub>2</sub> Medium/Light fold change cut-off  $\leq -1$ .

The data sets 1, 3 and 4 were generated with MaxQuant thus all the phosphopeptides of the same multiplicity (mono-, di-, and tri-phosphorylated) that provide quantitative information on a given site were aggregated to produce phosphosite quantification data. For the dataset 2, only phosphopeptide quantification is available. **5)**

**Acetylome dataset.** A large-scale SILAC acetylome study by Scholz *et al.* 2015 [24] (see their [Supp. Table 2](#)) profiling the acetylation sites increased in HeLa cells treated with: 19 different lysine deacetylase inhibitors (targeting all 18 human lysine deacetylases), Heavy condition, or vehicle, Light condition. The upregulated sites upon lysine deacetylase inhibitor treatment were selected with a fold change  $\geq 1.5$ .

**Algorithm for the functional prioritization of protein datasets to CK2.** The functional link between CK2 and different sets of proteins (e.g., the acetyl-proteins reported by Scholz *et al.* 2015 or human proteins matching the K motif) was determined by summarizing available biological information from different sources into a score between 0-1, low-high (see Appendix C). The summarized information is as follows: **1) Known CK2 substrates** (Yes: 1, No: 0): kinase-substrate relationships retrieved from IPTMnet and PSP databases using the Bioconductor's package `iptmnet` v0.1.8 [88] and manual download, respectively. **2) Known CK2 interactors** (Yes: 1, No: 0): direct interaction evidence retrieved from IMEx consortium [94] using the Bioconductor's package `PSICQUIC` v1.24.0 [75]. **3) Mention in the CK2 literature** (Yes: 1, No: 0): genes mentioned in papers retrieved by querying PubMed with "'Casein Kinase II'[Mesh]' as input using the Bioconductor's package `reutils` v0.2.3 [95]. **4) Neighborhood of CK2** (Yes: 1, No: 0): proteins that "interact" with CK2 in STRING v10 database using Bioconductor's package `STRINGdb` v1.26.0 [96]; `STRING score_threshold = 0.7`. **5) Semantic similarity between GO biological processes annotations** of CK2 and a given protein (GO BP, `measure=Wang`, `combine=BMA`) (0-1) calculated using Bioconductor's package `GoSemSim` v2.12.1 [97]. Wang method measures the semantic similarity of two GO terms based on both the locations of these terms in the GO graph and their relations with their ancestor terms. The combined Best-Match Average (BMA) strategy calculates the average of all maximum similarities on each row and column between GO terms associated with two genes [97]. **6) CK2 motif hit**, using the ScanProsite pattern `[ST][DE]X[DE]` or the regex `[ST][DE].[DE]` (Yes: 1, No: 0) **and semantic similarity for GO cellular components**; the values are averaged to control for proteins containing the CK2 motif but that don't colocalize with CK2 (0-1). For the retrieval of protein sequences for scanning against the CK2 motif the Bioconductor's package `UniProt.ws` v2.26.0 [98] and/or R's package `curl` v4.3 [99] were used. **7) Co-regulation with CK2**:

annotations retrieved from ProteomeHD [100], the cutoff was selected to retrieve approximately the top 150 co-regulated proteins for each of the CK2 subunits (Yes: 1, No: 0). **8) Differentially modulated by CK2 manipulation** (Yes: 1, No: 0): proteins containing sites differentially modulated by short treatment with the CK2 inhibitor CX-4945 in mitotic HeLa cells identified by Rusin *et al.* 2017 [51].

A “similarity” score to CK2 was calculated for each protein of interest by dividing the sum of the values observed across the different resources by the total number of resources i.e., 8 (see Appendix C). Two different scores were calculated: “score” (0-1), all resources values are averaged and “score\_sub”, all resources values are averaged except for the CK2 targets which is summed to the average score allowing for direct identification of known CK2 substrates as outliers. To prioritize the list of proteins these were ranked based on each of the similarity scores using the R base v3.6.3 function rank and the ties.method=”random”. The latter was chosen to indicate that for certain proteins there is no information to prioritize them above others while maintaining the original number of proteins to rank. The scores were represented as boxplots using R’s package ggstatsplot v0.4.0 [101] and the outlier.coef (Tukey’s method) was chosen in accordance with the data set of interest to identify a potential list of candidates (see Fig. 5 for an example). The scores were also clustered using one-dimensional k-means using R base v3.6.3 package kmeans with a k parameter = 5, selected by the elbow method i.e., plotting k (1-20) vs Total within-cluster sum of squares. The clustering allowed for the identification of five clusters and three main sets: 1) known CK2 substrates, 2) proteins functionally linked to CK2 that are not CK2 substrates, 3-4) proteins non-related to CK2 or with missing information on the resources used (see Fig. 5 for an example).

**Motif discovery.** The set of 8,103 unique 13 mer sequences centered at the acetylation site identified by Scholz *et al.* 2015 was analyzed for motif discovery using MoMo v5.0.5 [54] command-line tool from MEME suite. The parameters used were motif-x algorithm, width=13, min-occurrences=15, control peptides from shuffled input peptides, adjusted P value > 0.01. The motifs discovered with MoMo v5.0.5 were searched against the human proteome (UniProt reviewed human sequences (canonical and isoforms) downloaded in April 2019) using the FIMO v5.0.5 [58] command-line tool from MEME suite. The

FIMO hits ( $q$ -value  $< 0.05$ ) were scanned for  $K$  motif occurrences (see Motif scan above).

**Motif scan.** The linear patterns or motifs of interest were represented as regular expressions and searched using R base v3.6.3 functions `gregexpr` and `regexpr` against unique sequence windows provided, for example, in the MaxQuant [102] output for site identification or unique sequences retrieved from UniProt using R's package `curl` v4.3 or Bioconductor's package `UniProt.ws` v2.26.0. For proteome-wide scans several tools were used. Motif scan against the human proteome was performed using ScanProsite and the `patmatdb` tool of EMBOSS v6.6.0 suite using the reviewed human proteome annotated in UniProt (December 2019). To determine the number of matches by chance of the motif a random database was generated using the `fasta-shuffle-letters` command line tool of MEME v5.1.0 suite [56];  $k$ -mer=1/2/20, shuffle the sequences so that the frequencies of all words of length 1/2/20 are preserved. Motif scan against the human proteome was also performed using the FIMO v5.0.5 and MoMo v5.0.5 motifs as input (see Motif discovery below).

**Probability score of finding hits matching a given pattern.** The probability of finding a given pattern by chance in any given protein sequence was determined as indicated in the database ELM [59] (see ELM help: probability filter). For each position of a given pattern specifying one or more amino acids the corresponding amino acid probabilities (derived from UniProt and using an IUPred cutoff of  $\geq 0.4$ ) are summed. Then the sum of the amino acid probabilities calculated for each position is multiplied. This score was used to predict degenerate pattern/motif instances.

**Motif boundary and cellular context.** The  $K$  motif incidence in the human proteome was studied using the SLiMSearch4 framework [63], which allows the proteome-wide discovery of functional modules with an expected number of instances returned ( $E$ )  $< 10,000$  and localized to intrinsically disordered regions, IUPred disordered score  $\geq 0.4$  [63]. The results obtained for each  $K$  motif were downloaded and processed in R (see Appendix C). The SLiMSearch4 scores were visualized using R's package `ggstatsplot` v0.4.0. Additional information about domain (UniProt), motif (UniProt), secondary



structure (PDB), interaction with acetyllysine writers and erasers (PSICQUIC services), and mutations (COSMIC curated [66]) was also retrieved and overlapped to the positions matching the K motif (see Appendix C). This information is more specific than SLiMSearch4 domain analysis because it doesn't include domain information corresponding to the flank sequences and is not limited to the K motifs hits outputted by SLiMSearch4 since the information was retrieved for all hits in the proteome retrieved by ScanProsite. The consensus matrix for PDB structure was generated using the Bioconductor's package Biostrings v2.42.1 [103]. The heatmap was generated using the Bioconductor's package CrispRVariants v1.8.0 [104].

**SILAC cell culture.** U2OS cells were adapted to light (Arg0/Lys0), medium (Arg6/Lys4) and heavy (Arg10/Lys8) SILAC-dropout DMEM (Wisent) lacking L-arginine and L-lysine, supplemented with 10 % dialyzed-FBS (Wisent), 100 U/ml Penicillin, 0.1 mg/mL Streptomycin, and 400 mg/mL Pro, and the corresponding isotopically-labeled amino acid (Silantes) (Table 2.27). Total incorporation of U2OS labeled cells was determined to be > 95% achieved after 5 passages; cells continued to be passed in SILAC media for future experiments. SILAC media was filtered-sterilized using 0.2 µm filters (Nalgene).

**Table 2.27.** Isotopically-labelled amino acids.

Label	Formula	Concentration
L-arginine (Arg-0)	non-labelled	83.9 mg/mL
L-lysine (Lys-0)	non-labelled	60.04 mg/mL
L-arginine (Arg-6)	<sup>13</sup> C6	86.2 mg/L (0.398 mM)
L-lysine (Lys-4)	4,4,5,5-D4	61.16 mg/L (0.274 mM)
L-arginine (Arg-10)	<sup>13</sup> C6, <sup>15</sup> N4	87.7 mg/L (0.397 mM)
L-lysine (Lys-8)	<sup>13</sup> C6, <sup>15</sup> N2	52.4 mg/L (0.274 mM)

**Cell treatment for mass spectrometry.** For mass spectrometry analysis, label-free U2OS and HeLa cells or isotopically-labeled U2OS cells (light, medium, and/or heavy) were grown up to 80-90% confluency and treated with DMSO (light) or a variety of inhibitors for 6-16 h (label-free, medium, heavy). For CX-4945 inhibition followed by

protein digestion with Asp-N, light and medium isotopically-labeled U2OS cells were treated with DMSO and 30  $\mu$ M CX-4945 for 6 h, respectively. One phosphoproteome sample was run in triplicate in the LC-MS. For the study of modulation of CK2-dependent signaling by deacetylase inhibitors, labelled-U2OS cells were treated with 2.5  $\mu$ M TSA and 20 mM NAM in 6 and 16 h experiments. 6 h triple-SILAC experiment: NAM vs control condition (M/L condition) and TSA vs control condition (H/L condition). 16 h experiments, triple-SILAC experiment: NAM vs control condition (M/L condition) and TSA vs control condition (H/L condition); double-SILAC experiment: NAM (M/L condition) or TSA (M/L condition). At least three phosphoproteomes and total proteomes from independent biological replicates were measured for each of these experiments. For determining the acetylation part of the phosphoproteome in a label-free experiment both unlabelled HeLa and U2OS cells were treated either with 20 mM NAM, 2.5  $\mu$ M TSA, or 1 mM Bx.

**Cell lysis for mass spectrometry.** The culture plates were washed in PBS and harvested in 300  $\mu$ L-1 mL Guanidinium chloride lysis buffer (6 M Guanidinium chloride (GdmCl), 100 mM Tris pH 8.5, 10 mM tris(2-carboxyethyl)phosphine (TCEP), 40 mM 2-chloroacetamide (CAA)). The cells were scraped from the culture plates using GdmCl lysis buffer, the cell lysates were heated to 95  $^{\circ}$ C for 5 min, then kept in ice for 15 min, sonicated with a probe sonifier output 3 for 4 x 30 sec, and heated to 95  $^{\circ}$ C for 5 min again. Samples were spun at 3,500 x g for 30 min. The cell lysates were quantified using a NanoDrop and frozen at -80  $^{\circ}$ C until further use.

**Sample preparation for mass spectrometry.** Sample preparation for phosphoproteome (and total proteome) was performed using the EasyPhos platform as described by [Humphrey SJ et al. 2015 \[105\]](#) except for protein precipitation. For SILAC experiments in U2OS the samples belonging to each label were mixed 1:1 before protein precipitation. For label-free experiments, no mixing was required. Whenever possible the procedures were carried using low binding 1, 2, and 5-mL tubes.

**Protein precipitation.** Lysates for mass spectrometry were diluted < 2 M GdmCl. Proteins were precipitated using methanol-chloroform extraction. For each 100  $\mu$ L of

sample, the precipitation was performed as follows: add 400  $\mu$ L of methanol, vortex well, add 100  $\mu$ L of chloroform, vortex well, add 300  $\mu$ L of Mass Spectrometry grade water, and vortex well (sample should look cloudy). Next, centrifuge the samples at 14,000 x g for 2 min and remove the top aqueous layer (protein is visible in the interface). Add a total of 400  $\mu$ L of methanol per sample, vortex well, and centrifuge the samples at 14,000 x g for 3 min. After centrifugation, as much methanol as possible was removed without letting the pellet dry completely.

**Protein digestion.** Protein precipitates were resuspended in ABC buffer (50 mM ammonium bicarbonate (ABC), 10 % 2,2,2-Trifluoroethanol) to a maximum of 1.25 mg/mL. Digestion was initiated by the addition of 1:100 Trypsin/Lys-C Mix (Promega, V5073) resuspended in the ABC buffer for 30 min. After 30 min 1:100 Trypsin (Promega, V511A) resuspended in the ABC buffer was added to the samples. Samples were incubated in a shaker for 18 hours at 37°C, 2,000 rpm.

**Phosphopeptide enrichment.** Phosphopeptide enrichment was performed as described in [Humphrey SJ \*et al.\* 2015 \[105\]](#) using Empore SPE Disks-SDB (styrene divinylbenzene) 47 mm (66886-U) and 3M™ Empore™ SDB-RPS Extraction Disks (14-386-8).

**Phosphopeptide dephosphorylation.** The eluate from phosphopeptide enrichment was dephosphorylated in a 50  $\mu$ L reaction as indicated by NEB protocol: 40  $\mu$ L of the sample (~ 40  $\mu$ g of peptides) in water, 5  $\mu$ L of 10X of protein metallo-phosphatase buffer (500 mM Tris-HCl, 1 M NaCl), 5  $\mu$ L of MnCl<sub>2</sub>, and 1  $\mu$ L of 1:20 Lambda phosphatase dilution (1:1,000 final dilution, purified in our lab). The dephosphorylation reactions were carried out at 37 °C for 4 h.

**Acetyl-peptide enrichment.** Eluates from the phosphopeptide enrichment step were used for acetyl-peptide enrichment. The anti-acetyllysine antibody beads (Acetyl Lysine Antibody, Agarose 5mg, ICP0388-5MG) were washed in immunoprecipitation buffer (50 mM MOPS pH 7.2, 10 mM Sodium phosphate, 50 mM Sodium chloride) 3 x 5 min, centrifugation in between at 2,000 x g for 1 min. Next, a bead slurry was prepared with the immunoprecipitation buffer at the appropriate concentration. The phosphopeptides

were diluted in the immunoprecipitation buffer and 20  $\mu\text{L}$  of the bead slurry was added to a total of 500  $\mu\text{L}$ . Samples were incubated for 12 hrs at 4  $^{\circ}\text{C}$  on a rotation wheel. Next, the beads were washed on a rotation wheel 4 times for 2 min with 1 mL immunoprecipitation buffer, centrifugation in between at 1,000 x g, followed by two washes with 1 mL distilled water, centrifugation in between at 1,000 x g. All the steps were performed at 4  $^{\circ}\text{C}$ . Residual water was removed, and the phospho-acetyl peptides bound to the beads were eluted by adding 50  $\mu\text{L}$  of 0.1% TFA in water, mixing with the pipette, and incubating for 5 min. The elution was repeated 3 x with centrifugation in between at 2,000 x g for 2 min, the eluate was collected in a clean low retention vial using a DNA gel loading tip. The final eluate (150  $\mu\text{L}$ ) was filtered using a Microcon-10kDa Centrifugal Filter Unit with Ultracel-10 membrane (MRCPR010) to eliminate any polymer contamination from the beads. The filters were pre-conditioned by adding 50  $\mu\text{L}$  of 30% ACN and centrifugation at 15,000 x g for 15'. After filtering the eluate at 15,000 x g for 30 min, the filter was washed with 50  $\mu\text{L}$  of 30 % ACN by centrifugation at 15,000 x g for 15 min. The filtered samples (250  $\mu\text{L}$ ) were concentrated in a Speed Vac vacuum centrifuge at 45  $^{\circ}\text{C}$  under vacuum until  $\sim$  5  $\mu\text{L}$  remained after which  $\sim$  7  $\mu\text{L}$  of 0.1 % FA solution was added to ensure the acidification of the solution.

**Total protein clean-up.** For total protein analysis clean-up, C18 stage-tips were assembled in the laboratory using pipette tips and 3 cores of 2 layers of C18 (Empore SPE Disks C18 47mm diameter, 66883-U) for each tip assembled. Samples were diluted to 0.1  $\mu\text{g}/\mu\text{L}$  (20  $\mu\text{g}$  in 200  $\mu\text{L}$ ) with Buffer A' (2% ACN, 0.1% TFA). The pH was checked to be  $<$  3.0 and adjusted if necessary, with 5% or 10% TFA. The tips were activated with Buffer B' (80% ACN, 0.1% TFA) 3 x 200  $\mu\text{L}$ , centrifuged at 3,000 x g, 2 min. Next the tips were equilibrated with Buffer A' (2% ACN, 0.1% TFA) 3 x 200  $\mu\text{L}$ , centrifuged at 3,000 x g, 2 min. After equilibration 20  $\mu\text{g}$  of the sample in 200  $\mu\text{L}$  was loaded and centrifuged at 750 x g, 5 min. The flow-through was reloaded and centrifuges again. The bound samples were washed with Buffer A (2% ACN, 0.2% FA) 3 x 200  $\mu\text{L}$ , centrifuged at 3,000 x g, 2 min. After washing the samples were eluted with Buffer B (80% ACN, 0.2% FA), 100  $\mu\text{L}$ , centrifuged at 2,000 x g, 3 min. The elution step was repeated with 50  $\mu\text{L}$  of Buffer B. The eluates were concentrated in a Speed Vac vacuum centrifuge at 45  $^{\circ}\text{C}$  under vacuum until 10-20  $\mu\text{L}$  remained after which 1-2  $\mu\text{L}$  of 1% FA

solution was added to ensure the acidification of the solution. <sup>[L</sup><sub>SEP]</sub>The final peptide concentration was determined by measuring A280 using a Nanodrop. For an LC-MS/MS analysis the samples were diluted to ~ 0.1 g/L (a final load volume of 3-7  $\mu$ L should contain 0.5-0.9  $\mu$ g) and ~ 15  $\mu$ L were transferred into glass LC-MS vials. Alternatively, clean-up for total protein was performed using 100  $\mu$ l bed Pierce™ C18 Tips (87784) as indicated by the manufacturer.

**LC-MS run.** Phosphoproteome and proteome samples were analyzed using an Orbitrap Elite Hybrid Ion Trap-Orbitrap mass spectrometer (Thermo Scientific) or a Q Exactive Plus mass spectrometer (Thermo Scientific).

Orbitrap Elite Hybrid Ion Trap-Orbitrap mass spectrometer was connected to a NanoFlex (Thermo Electron Corp., Waltham, MA) nanospray ionization source with a source voltage of 2.4 KV. Mass spectrometer was operated by Thermo XCalibur software (version 2.7.0) in a data-dependent mode using an FT/IT/CID Top 10 scheme for phosphoproteome and Top 20 scheme for proteome samples. MS1 scans were performed at 120,000 resolution scanning from 400 – 1450 m/z with Automatic Gain Control (AGC) set to  $5 \times 10^2$  ions and an isolation window of 1.0 m/z. The top 20 most abundant ions were selected for MS2. A lock mass ion was enabled for 445.120025 m/z. Peptide ions were fragmented using a CID (Collision-Induced Dissociation) with Normalized Collision Energy set to 35%, an activation time of 10 ms, and a Q value of 0.25. Data-dependent MS2 scans were acquired in the linear ion trap using rapid scan selection, with a dynamic exclusion window of 60 seconds.

Q Exactive Plus mass spectrometer was connected to a NanoFlex (Thermo Electron Corp., Waltham, MA) nanospray ionization source with a source voltage of 2.4 kV. The mass spectrometer was controlled by Xcalibur software (Thermo, v. 4.0). Discovery data were acquired in a data-dependent mode (DDA) using an FT/FT/HCD Top 10 scheme for phosphoproteome and Top 20 scheme for proteome samples. MS1 scans were performed at 70,000 resolution scanning from 375 – 1500 m/z with Automatic Gain Control (AGC) set to  $3 \times 10^6$  ions and a maximum IT of 250. A lock mass ion was enabled for 445.120025 m/z. Peptide ions were fragmented using an HCD (Higher-Energy Collision Dissociation)

with Normalized Collision Energy set to 25%. Data-dependent MS<sup>2</sup> scans were acquired at 35,000 resolution, with an Automatic Gain Control (AGC) set to 2e5 ions and a maximum IT of 108. Dynamic exclusion was 30 seconds.

Both mass spectrometers were connected to an M-class nanoAquity UHPLC system (Waters). Samples were first injected onto a Symmetry C18 trapping column (5 mm, 20 mm x 180 mm, 100 Å) at a flow rate of 5 µL/min in 99 % mobile phase A (0.1 % FA (v/v), 1 % mobile phase B (0.1 % FA (v/v) in ACN) for 5 minutes. Peptides were then separated on an ACQUITY UPLC M-Class Peptide BEH C18 Column, 130 Å, 1.7 mm, 75 mm × 250 mm, operating at a flow rate of 300 nL/min at 35°C. Phosphoproteome peptides were separated using non-linear gradients: 130 min (initial condition was 1 % B, which increased to 7.5 % B over 1 min, to 25 % B at 60 min, to 32.5 % B at 74 min, to 60 % B at 80 min, to 98 % B at 85 min and held for 5 min, then decreased to 2 % B over 5 min and held at 2 % for 5 minutes, increased to 98 % B over 5 minutes and held at 98 % for 5 min, and finally decreased to 2 % B over 5 minutes and held at 2 % for 15 min) or 160 min (initial condition was 1 % B, which increased to 4%B over 1 min, to 25 % B at 120 min, to 50 % B at 130 min, to 95 % B at 135 min and held at 95% for 5 min, then decreased to 5%B over 5 min and held at 5% for 15 minutes). Proteome peptides were separated using a 290 min non-linear gradient: initial condition was 5 % B, which increased to 7.5 % B over 1 min, to 25 % B at 180 min, to 32.5 % B at 220 min, to 60 % B at 240 min, to 98 % B at 245 min and held for 5 min, then decreased to 2 % B over 5 min and held at 2 % for 5 minutes, increased to 98%B over 5 minutes and held at 98 % for 5 min, and finally decreased to 5 % B over 5 minutes and held at 5 % for 15 min.

**Computational proteomics.** For SILAC experiments protein and phosphopeptide identification was performed using the open software MaxQuant v1.6 [102]. For label-free experiments, both PEAKS v.9 [106] and SearchGUI v3.3.17 [107] were used. Unless specified the default, parameters were used for the tools. Protein and peptide searches were performed using the human proteome containing reviewed canonical and isoforms sequences only (April 2019).

**MaxQuant analysis.** Trypsin/P was set as the protease, a maximum of two missed cleavages was allowed unless specified, and specific search was selected. For CK2 inhibition SILAC experiment, Asp-N was selected as the protease and a maximum of three missed cleavages was set. Cysteine carbamidomethylation was set up as fixed modification for all searches. N-terminal protein acetylation, Met oxidation, and Asn/Gln deamidation were used as variable modifications for all searches. A maximum number of 4 modifications was accepted per peptide. Protein quantification was performed using only the unmodified peptides or peptides containing carbamidomethyl cysteine and/or oxidized methionine. Protein and peptide identification were controlled at 1 % FDR. To increase identification coverage Match between runs was selected. Phosphorylation and acetylation site localization probability was filtered by sites with localization probabilities  $\geq 0.75$ . The significantly modulated proteins and modified sites (phosphosites and acetylation sites) were found using a one-sample t-test; the obtained P values were FDR adjusted 1 %. These and other common processing steps for MaxQuant output were performed in R (see Appendix C) following a typical Perseus [108] analysis workflow.

**SearchGUI analysis.** For double enrichment experiments with and without dephosphorylation, SearchGUI 3.3.17 (October 29, 2019) was used in Windows for peptide identification and a maximum of four missed cleavages was allowed. All identification search engines available from within SearchGUI were used. Same variable modifications and proteome database was set as in the MaxQuant search. The identification results were visualized and exported using PeptideShaker v1.16.42 (July 10, 2019) [109].

**PEAKS analysis.** For double enrichment experiments with dephosphorylation PEAKS v.9 was also used for peptide identification, a maximum of four missed cleavages was allowed. The other parameters were set as above.

## 2.8 Discussion

CK2 target sites occur in the vicinity of other modified residues including lysine acetylation sites. Since lysine is regarded as a negative determinant of CK2-dependent phosphorylation and lysine acetylation removes the positive charge on this residue the

breath of our work addresses if these residues are tolerated in the context of the CK2 motif and the extent of the occurrence in the human proteome. To achieve this, we performed kinase assay, immunoblotting, and phosphoproteomic experiments along with the processing and analysis of PTM and phosphoproteomic data available in the databases and the literature. As a result, we identified *in vitro* a link between CK2-dependent phosphorylation and lysine acetylation at the substrate level and determined *in vivo* candidates of this regulatory association.

### 2.8.1 PTM interplay at CK2 target sites

As indicated by our global analysis of CK2 networks, the neighboring residues of CK2 target sites (e.g., +/-7 amino acids) are modified by other constituents of regulatory networks and are likely involved in interplay. By studying the cis-acting interplay predicted for the CK2 targets annotated in the PhosphoSitePlus database (PSP), interplaying lysine acetylation sites were identified based on evolutionary history and closeness in the protein 3D structure. Although the number of tentative interplaying acetylation sites observed was not comparable to phosphorylation, likely due to the prevalence of phosphorylation data, the *in silico* evidence pointed to the co-evolution of these two PTM types. An interplay between lysine acetylation and phosphorylation sites has been previously reported in the H3 histone tail where phosphorylation of Ser10 residue is enhanced by the acetylation of either Lys9 or Lys14 residues [110].

### 2.8.2 CK2 specificity against peptides containing lysine or acetyllysine as determinants

Considering that lysine acetylation changes the local properties of the modified sequence by removing a positive charge and that CK2 is an acidophilic kinase, *in solution* peptides resembling the CK2 peptide substrate DSD but containing either lysine or acetyllysine at positions +1, +2, or +3 from the phosphorylation site were subjected to phosphorylation by CK2 *in vitro*. Peptide phosphorylation data indicates that CK2 phosphorylation occurs in the presence of a lysine residue at position +2 (24.19 %) and, more importantly, that acetylation at this position promotes phosphorylation to 71.13 % of DSD signal. Phosphorylation by CK2 does not occur in the presence of lysine or acetyllysine at



position +1 and +3, with acetyllysine at position +3 being the most tolerated (~ 6.26 % rescue) among them. These results are in accordance with the working model for CK2 substrate specificity where positions +1 to +3 are deemed the most relevant for phosphorylation [111,112]. They are also in accordance with a model of the phosphoacceptor substrate binding region bound to the peptide substrate RRRADDSDDDD and with CK2 mutants defective in determinants recognition [113]. In the model and mutational analysis, the Asp determinant at position +2 is only directly implicated with the Lys49 residue in the phosphoacceptor binding region, whereas the Asp determinants at positions +1 and +3 are implicated directly/indirectly with a network of positive residues including Lys198/Arg191/Arg195 and Lys74-77/Arg80/Lys83, respectively [113]. The latter could suggest a higher tolerance to the chemical nature of the residue at position +2 due to a decreased necessity for accommodating positive charges in the phosphoacceptor binding site if position +2 on the peptide substrate were to be compromised by the substitution of the negative charge. In addition, a web logo analysis of 433 known CK2 substrates annotated in the database phosphoELM also pointed to higher variability of determinants at position + 2 with 53 % of the residues being different than Asp or Glu, compared to 43 % and 25 % for positions +1 and +3, respectively [114]. In this regard, lysine was demonstrated to be tolerated for CK2-dependent phosphorylation at position +2. Furthermore, for the first time acetyllysine was identified as promoting phosphorylation to a higher extent than lysine further reinforcing the relationship between acetylation and CK2.

To examine the phosphorylation of endogenous sites containing lysine/acetyllysine at position +2 by CK2, two peptides conforming to Ser33 of CTTN and Thr24 of SET-iso 2 sites were phosphorylated *in vitro*. Although the unmodified lysine peptides were not phosphorylated a modest increase was observed for the acetylated endogenous sequences compared to DSD. The lower phosphorylation observed for these sequences could be due to the fact that the control peptide assessed was designed to mimic an ideal CK2 peptide substrate, whereas endogenous sequences are obviously more complex and other mechanisms are at play in favoring phosphorylation in the cellular context [115]. We also performed immunoprecipitation of candidates to perform LC-MS analysis for identification of modifications on these proteins. However, we were unable to achieve

immunoprecipitation of sufficient amounts of endogenous protein to enable identification of PTMs. Thus, the analysis was moved forward to the identification of candidates by mining the information in PTM databases like PhosphoSitePlus, iPTMnet, and PTMcode v2 and in the literature.

### 2.8.3 Identification of K motif candidates

Several linear patterns were derived from the peptide phosphorylation data, referred to as K motifs variants that conformed to the minimal or extended CK2 consensus motif in the context of Lys at position +1, +2, or +3. The known CK2 substrates found to contain Lys at position +2 are Ser429 of TFAP2A (transcriptional regulator), Ser53 of RNPS1 (mRNA splicing), Ser236 of MAPRE2 (microtubule polymerization), and Ser123 of APEX1 (DNA repair). The number of hits found among the curated CK2 substrates in PSP was significantly different between the three positions in the K motifs with more hits found for +2 as might be expected from our studies with peptides. Similar results were obtained from a global study for the identification of CK2 substrates by using *in vitro* CK2 assay followed by phosphoproteomics [49]. In addition, it was observed that the K motif hits retrieved for position +1 and +3 contain four or more negatively charged residues, with a total among variants of 10 *in vitro* phosphorylated hits containing four or more negative residues towards the C-terminal of the phosphorylatable site. This finding indicates that a highly negative environment is preferred by CK2 in the context of Lys. Moving away from *in vitro* analysis, phosphoproteomic studies where CK2 activity was manipulated in U2OS and HeLa cells by inhibition with the ATP-competitive inhibitor CX-4945 were also analyzed. As before, similar results were obtained when analyzing the matches among the significantly inhibited *in vivo* sites with most of the site matching position +2. Interestingly, inhibition of a CK2 site containing Arg in the position +2 was also observed *in vivo*. This observation brings up the question of how translatable is the information obtained for Lys as a determinant at position +2 is to Arg. Overall, a role of lysine as a positive determinant in the appropriate negative environment was observed both *in vitro* and *in vivo* in sites responding to CK2 activity manipulation. Likely many other potential sites are yet still to be identified for which the adoption of a phosphoproteomic strategy where the phosphopeptides are generated using Asp-N may

contribute to report complementary information to that obtained with conventional trypsin digestion.

#### 2.8.4 CK2 and lysine acetylation networks

The connection between CK2 and lysine acetylation networks was also explored *in silico* to help understand the potential regulatory role of this modification as determinant of CK2-dependent phosphorylation. As a result, hundreds of proteins known to be acetylated in the cell identified by Scholz *et al.* 2015 [24], including 124 CK2 substrates, ranked high in an algorithm that we developed to predict functional association to CK2 based on protein interaction, GO similarity, CK2 motif scan, literature mentions, CK2 activity manipulation, etc. Among these proteins, we found the histone deacetylase HDAC1, a CK2 substrate directly linked to the maintenance of acetylation in the cells [35], and COPS2, a protein encoded by a gene whose expression is regulated by the lysine deacetylase SIRT7 [116]. COPS2 functions as a component of the COP9 signalosome complex, a regulator of the ubiquitin conjugation pathway with associated CK2 activity [117]. Taken together, these findings demonstrate the existence of direct and indirect links between CK2-dependent signaling and the acetylation networks in the cell. Furthermore, quantitative information gathered on the acetyl-proteins functionally linked to CK2 pointed to the lysine deacetylase inhibitors Nicotinamide (NAM), Trichostatin A (TSA), and Bufexamac (Bx) as chemical tools important to study these links. Moreover, motif discovery analysis revealed that sites known to be acetylated in the cell occur in negative sequences (adjusted P value < 0.01) indicating a representation of acidic acetylation motifs in the cell. Interestingly, from several PTM sources it was confirmed that many sites matching the K motif linear patterns in the human proteome are targets of acetylation and phosphorylation in the cell, with no distinction among variants. Instances of phosphorylation and acetylation targeting the same K motif hits were observed, including the CK2 target sites Ser123 of APEX1 and Ser236 of MAPRE2. However, although no experimental information is available to establish the co-occurrence of these modifications at these sites, a small number of these were predicted to interplay *in cis*.

The kinase-substrate relationships retrieved indicate that kinases other than CK2 including the basophilic kinases AKT1, PRKACA (PKA), and PRKCA (PKC) phosphorylate the K motif hits. This was somewhat expected as the presence of upstream motifs provide the determinants required for these and other kinases; also, the Lys residue in the K motifs serves as a positive determinant of phosphorylation by PRKACA and PRKCA kinases. Furthermore, protein-protein interaction information corresponding to the proteome-wide K motif hits revealed 411 gene-lysine acetyltransferase (KAT) and 396 gene-lysine deacetylase (KDAC) tentative underlying enzyme-substrate associations, with > 190 unique proteins hits-KAT-KDAC interaction pairs. The wealth of writers and erasers acting upon the K motif hits is also expanded by the fact that the Lys matching the K motif were also found to be targets of methylation, sumoylation, and ubiquitination. This brings new regulatory players into consideration and makes us ponder on the intricacy of the regulation and signal integration that could be occurring at the sequences conforming to the K motifs in cell.

Overall, a list of K motif candidate sites for CK2 phosphorylation was generated from the integrated information which can serve as a base for future validation experiments. Among these, three relevant activities that regulate Lys residues modifications in the cell were identified, i.e., the lysine acetyltransferase KAT6A, the histone methyltransferase PRDM2, and the deubiquitinase OTUD7B. The site Ser1104 (sKDDEEDE) of the oncoprotein KAT6A (MYST3) is a CK2 candidate target site. This acetyltransferase is capable of acetylating H3 and H4 residues *in vitro* and H3 as part of the MOZ/MORF complex (source UniProtKB). Furthermore, KAT6A controls the activation of TP53 by PML via acetylation of 'Lys-120' and 'Lys-382' residues of TP53 (source UniProtKB). The site Ser1256 (SKEEEE) of PRDM2 is also a candidate for CK2 phosphorylation. PRDM2 methylates Lys9 of H3 (source UniProtKB), a repressive mark [118]. Finally, it is important to add that the number of hits found in the human proteome scan equaled the number of hits found by chance in a shuffled version of the human proteome. This is because the K motif variants are short linear motifs, however, the probability of finding them by chance in a protein is still considerably lower than the accepted CK2 consensus sequence [ST]XXE. According to a study of false occurrences of functional motifs in the human proteome by Allegra *et al.* [119], false motif occurrences do not tend to compete

with true positive hits since there are other boundary conditions required by motifs for their functionality. For instance, the discrimination of true positive hits by interactors of similar sites [119]. This differentiation can also be observed, at least in part, in phosphoproteomic experiments where not all proteins matching the CK2 consensus or known CK2 substrates are differentially modulated upon manipulation of CK2 activity [51].

### 2.8.5 Boundary and cellular conditions of the K motif hits

The analysis of the boundary and cellular conditions of the K motif hits in the human proteome showed that hundreds of hits for all variants are localized to disordered regions and that several hits have a propensity to fold upon binding. The latter is an important property of functional regulatory motifs that provides conformational flexibility and affinity control [64]. A significant number of hits were also found localized to protein domains, this should not raise a warning as known modification sites also localize to functional domains; the latter can be due to the fact that annotated domains may describe sequences with accessible loops or represent conserved disordered regions [63].

Conservation of the residues matching the motif vs that of the flanks is another important aspect of the K motif hits. An assessment of the proportion of conserved hit sequences compared to their direct flanks (5 amino acids up/downstream) showed strong conservation for ~20 of the hits for K motif position +2. The list includes Ser161 of BAZ1B, a neighbor of Ser160 which was found inhibited by treatment with CX-4945; phosphorylation at both sites was confirmed in the proteome database GPMdb. Several hits were also found present at the same position as the query species motif in more than 10 species. The conservation of K motif hits matches the view that functional motifs in rapidly evolving intrinsically disordered regions are likely to be more conserved than the neighboring regions [63].

Another interesting finding regarding the K motif hits is that they are targets of SNPs and mutations. Notably, several instances where the Lys residue in the K motif was replaced by Glu, or vice versa, were observed. These changes were recorded for all the three variants of the K motif, but they could be playing a bigger role at positions +1 and +3 where Lys is not tolerated and Glu is a positive determinant. For example, the Lys1994 to

Asp substitution in the protein ATM detected in hematopoietic and lymphoid tissue occurs in a sequence that matches the K motif variant +1, which could potentially impact the phosphorylation of the known phosphosite Ser1993 and create a target sequence for CK2. The mutational analysis highlights how proteoform dynamics could change the primary sequence at K motifs. A functional association between CK2 and ATM pathways has been discussed elsewhere [120,121].

### 2.8.6 Integration of phosphorylation and lysine acetylation networks

The evidence gathered both experimentally and *in silico* for the K motif hits paves the way for investigating the co-occurrence of phosphorylation-acetylation *in vivo* and at the K motif variants. To achieve this, U2OS and HeLa cells were treated with three different lysine deacetylase inhibitors known to target acetylation sites matching the K motif. After assessing the treatment for different periods of time, increased lysine acetylation in both cell lines was observed along with the modulation of the CK2-dependent phosphoproteome at 6 and 16 h of treatment with both 2.5  $\mu$ M trichostatin A (pan-HDAC inhibitor) and 20mM nicotinamide (pan-SIRT inhibitor). The modulation occurs at phosphosites matching the CK2 consensus or by potentially targeting acetylation sites in their vicinity. After observation of a direct/indirect impact on the CK2-dependent phosphoproteome, a double enrichment strategy was applied exploring the acetylated fraction of the phosphoproteome in U2OS and HeLa cells. To estimate the effectiveness of the double enrichment, the  $ms^2$  spectra identified were searched for the acetyllysine diagnostic fragment ion at 126.0913  $m/z$  [76,77] at 10 ppm. After comparing the retention time of the spectra containing the diagnostic fragment between the double enriched samples and their dephosphorylated counterparts a displacement of the  $m/z$  hits from early to later time points was observed. This is consistent with a decrease in polarity due to the removal of the phosphorylation group suggesting that the corresponding acetylated peptides indeed carried phosphorylation and identifying them as phospho-acetyl peptides. This is an exciting discovery using a high-throughput method considering the assumptions often made in the literature regarding different PTMs types as mutually exclusive. However, as mentioned before the presence *in vivo* of phosphorylation-acetylation pairs and their interplay *in cis* has been previously described in the histone

code [122,123]. To assess the validity of the double enrichment approach, GO biological process classification was performed and compared to the processes represented in a large-scale study of the acetylome by Scholz *et al.* 2015. Overlapping processes were observed including gene expression, RNA metabolism, chromosome organization, cytoskeleton organization, apoptosis, and mitosis. Notably, acetylation of HIF1A upon treatment with the deacetylase inhibitor Bx was also observed as described in Scholz *et al.* 2015. However, different modification sites were uncovered in the protein this time, namely, acetylation of Lys251 in concert with phosphorylation of either Ser247 or Ser244 residues. These sites localize to the second PAS domain of HIF1A, i.e., PAS-B, which is necessary for heterodimerization [78]. Interestingly, Ser247 has been previously identified as a target of CK1; phosphorylation by CK1 impacts HIF1A binding to ARNT affecting HIF1A induced gene expression [83]. Additionally, several known phosphorylation and acetylation sites were identified for histones and other proteins, a direct validation of our findings. Interestingly, many of the co-occurring histone phosphorylation and acetylation sites identified localize to the C-terminal of H1 histone variants. This region is a long unstructured domain that functions in maintaining the binding of H1 to the nucleosome and in chromatin condensation [124]. Functional classification and enrichment analysis of this set of proteins indicate that they function on biological processes linked to lysine acetylation, such as histone modification, protein acetylation/deacetylation, and regulation of chromatin organization. Several kinase-substrate relationships and one acetyltransferase-substrate relationship were found highlighting potential enzymes that could be used in follow-up experiments to understand phosphorylation-acetylation interplay. Among the kinases found are: CDK2, NEK9, AURKB, and CHEK2, which function in cell cycle regulation; STK4, pro-apoptosis; IKBKE, inflammation and DNA damage-induced cell death; and PRKCA, positive and negative regulation of cell proliferation, apoptosis, differentiation, migration, and adhesion, etc. The acetyltransferase observed was EP300 corresponding to an autocatalytic site [79]. Importantly, several regulatory functions associated with the phosphorylation sites in phosphorylation-acetylation pairs were observed regarding regulation of molecular association, induction/inhibition of activity, change of protein conformation, alteration of receptor desensitization, receptor internalization, and protein

stabilization. Given the functional relevance of these phosphorylation sites, we could speculate that the co-occurring acetylation sites could also display regulatory functions. In this regard, including all proteins with co-occurring phosphorylation and acetylation sites, three main relevant groups of writers, erasers, and/or readers were observed: methylation, acetylation, and phosphorylation. Indicating that co-occurring phosphorylation and acetylation sites target essential regulators and mediators of cell signaling.

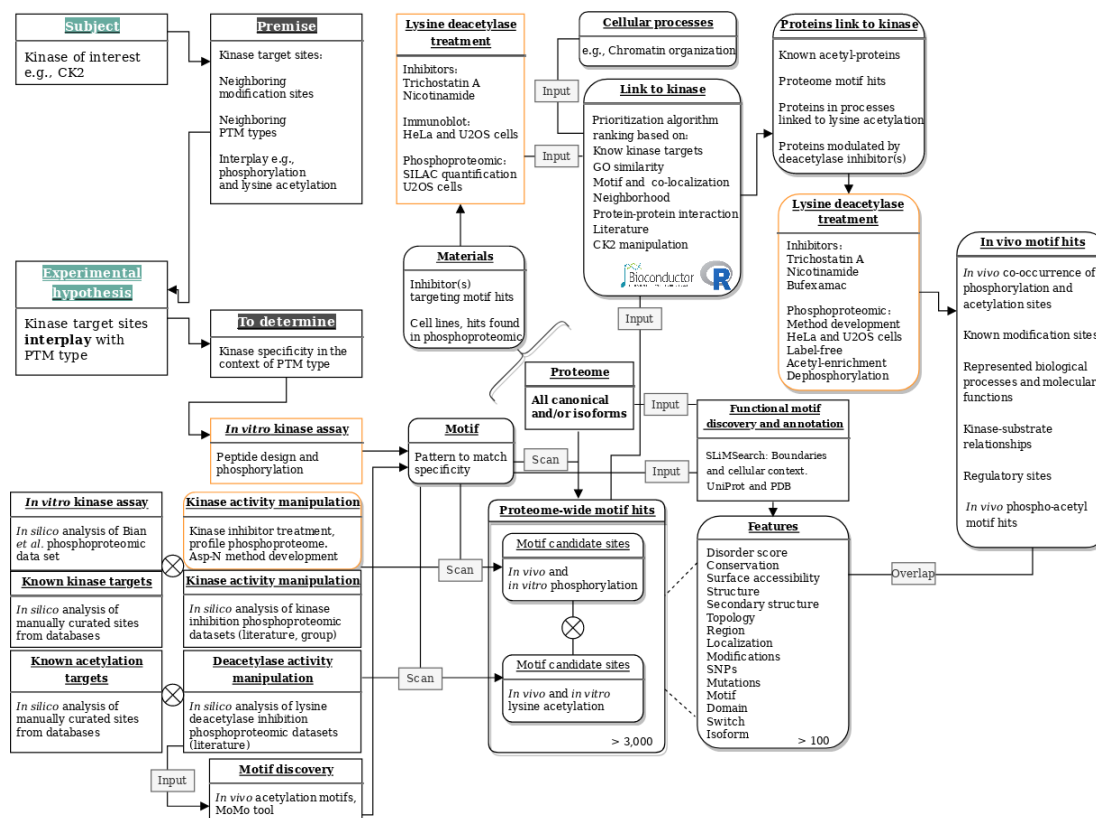
Finally, and more importantly, several hits for CK2 consensus sites were observed for phosphopeptides co-occurring with acetylation and more in particular for the K motif +2 variant. Of interest is the PPFIBP1 protein which functions in the disassembly of focal adhesions. In this case, both the Ser and Lys residues matching the K motif have been previously found phosphorylated and acetylated, respectively. Furthermore, this protein was found acetylated in the Bx treatment which inhibits deacetylation by HDAC6 an enzyme with an important role in microtubule-dependent cell motility by deacetylating tubulin. It has been shown that HDAC6 inhibition results in a decrease in microtubule dynamics which in turn results in decreased focal adhesion turnover [125]. Among other candidates for studying the interplay between CK2-dependent phosphorylation and lysine acetylation are DDX18, PPP1R10 (PNUTS), HDGFL2, NOP56, SLIRP, BDP1, ATRX, CHD5, POLR3A, CORIN, ALPL, and DNMT1 (Supp. Table 19).

Overall, our results further extend understanding of the complexity of kinase specificity and more specifically on CK2 specificity and the role of lysine acetylation. We provide *in silico*, *in vitro*, and *in vivo* evidence supporting acetylation of lysine as a positive determinant of CK2-dependent phosphorylation at position +2. We also provide experimental evidence on the presence of phosphorylation sites, some of which are known to be regulatory, and neighboring acetylation. This was accompanied by the identification of several kinases including CK2 that target such phosphorylation sites and are potential candidates in the search for the interplay between phosphorylation and acetylation. Thus, our studies lay the groundwork for further low and high-throughput exploratory and validation experiments.



## 2.9 Concluding remarks and implications

Here we combined proteomic, bioinformatic, and molecular biology strategies for understanding the interplay between phosphorylation and acetylation networks in the cell (Fig. 2.17). As such, we provide a framework for studying experimentally and *in silico* how phosphorylation by protein kinases can be regulated by the addition of a nearby modification through the exploration of the impact of lysine acetylation on CK2-dependent phosphorylation. Applying this framework brought us one step closer to deciphering the phosphoproteome and its dynamic integration with regulatory PTM networks. This was possible by generating our own context-dependent datasets and by reusing data available in the literature and databases and thus analyzing it outside the original context in which they were generated. It was also possible using the R programming language extended by Bioconductor packages. The development of a data-driven application that summarizes all the main analytical steps that allows the reanalysis and reuse of the data is recommended (see Chapter 4). In addition, exploring the functional association of CK2 and biological processes regulated by acetylation such as chromatin organization is recommended to provide groundwork on the relevance of such connections (see Chapter 3).



**Figure 2.17.** Established framework. Round boxes: main results; orange border: main experiments.

## 2.10 References

1. Nuñez de Villavicencio-Diaz, T.; Rabalski, A.J.; Litchfield, D.W. Protein Kinase CK2: Intricate Relationships within Regulatory Cellular Networks. *Pharmaceuticals (Basel)*. **2017**, *10*, 27, doi:10.3390/ph10010027.
2. Aebersold, R.; Agar, J.N.; Amster, I.J.; Baker, M.S.; Bertozzi, C.R.; Boja, E.S.; Costello, C.E.; Cravatt, B.F.; Fenselau, C.; Garcia, B.A.; et al. How many human proteoforms are there? *Nat. Chem. Biol.* **2018**, *14*, 206–214.
3. EA, P.; EV, P.; EV, I.; MA, P.; AT, K.; VG, Z.; AV, L.; AI, A. The Size of the Human Proteome: The Width and Depth. *Int. J. Anal. Chem.* **2016**, *2016*, doi:10.1155/2016/7436849.
4. Kong, A.T.; Leprevost, F. V.; Avtonomov, D.M.; Mellacheruvu, D.; Nesvizhskii, A.I. MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **2017**, *14*, 513–520, doi:10.1038/nmeth.4256.
5. Aderem, A. Systems Biology: Its Practice and Challenges. *Cell* **2005**, *121*, doi:10.1016/j.cell.2005.04.020.
6. Bhalla, U.S.; Iyengar, R. Emergent properties of networks of biological signaling pathways. *Science* **1999**, *283*, 381–7.
7. Woodsmith, J.; Kamburov, A.; Stelzl, U. Dual Coordination of Post Translational

- Modifications in Human Protein Networks. *PLoS Comput. Biol.* **2013**, *9*, doi:10.1371/journal.pcbi.1002933.
8. Minguez, P.; Parca, L.; Diella, F.; Mende, D.R.; Kumar, R.; Helmer-Citterich, M.; Gavin, A.C.; Van Noort, V.; Bork, P. Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.* **2012**, *8*, doi:10.1038/msb.2012.31.
  9. Beltrao, P.; Albanèse, V.; Kenner, L.R.; Swaney, D.L.; Burlingame, A.; Villén, J.; Lim, W.A.; Fraser, J.S.; Frydman, J.; Krogan, N.J. Systematic functional prioritization of protein posttranslational modifications. *Cell* **2012**, *150*, 413–425, doi:10.1016/j.cell.2012.05.036.
  10. Fischle, W. Talk is cheap - Cross-talk in establishment, maintenance, and readout of chromatin modifications. *Genes Dev.* 2008, *22*, 3375–3382.
  11. Vaissière, T.; Herceg, Z. Histone code in the cross-talk during DNA damage signaling. *Cell Res.* **2010**, *20*, 113–5, doi:10.1038/cr.2010.14.
  12. Bannister, A.J.; Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* 2011, *21*, 381–395.
  13. Musselman, C.A.; Lalonde, M.-E.; Côté, J.; Kutateladze, T.G. Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1218–27, doi:10.1038/nsmb.2436.
  14. Huang, Y.; Xu, B.; Zhou, X.; Li, Y.; Lu, M.; Jiang, R.; Li, T. Systematic characterization and prediction of post-translational modification cross-talk. *Mol. Cell. Proteomics* **2015**, *14*, 761–770, doi:10.1074/mcp.M114.037994.
  15. Meek, D.W.; Anderson, C.W. Posttranslational modification of p53: cooperative integrators of function. *Cold Spring Harb. Perspect. Biol.* **2009**, *1*, a000950, doi:10.1101/cshperspect.a000950.
  16. Kasthuber, E.R.; Lowe, S.W. Putting p53 in Context. *Cell* 2017, *170*, 1062–1078.
  17. Rust, H.L.; Thompson, P.R. Kinase consensus sequences: a breeding ground for crosstalk. *ACS Chem. Biol.* **2011**, *6*, 881–92, doi:10.1021/cb200171d.
  18. St-Denis, N.; Gabriel, M.; Turowec, J.P.; Gloor, G.B.; Li, S.S.C.; Gingras, A.C.; Litchfield, D.W. Systematic investigation of hierarchical phosphorylation by protein kinase CK2. *J. Proteomics* **2015**, *118*, 49–62, doi:10.1016/j.jprot.2014.10.020.
  19. Duncan, J.S.; Turowec, J.P.; Duncan, K.E.; Vilks, G.; Wu, C.; Lüscher, B.; Li, S.S.-C.; Gloor, G.B.; Litchfield, D.W. A peptide-based target screen implicates the protein kinase CK2 in the global regulation of caspase signaling. *Sci. Signal.* **2011**, *4*, ra30, doi:10.1126/scisignal.2001682.
  20. Turowec, J.P.; Duncan, J.S.; Gloor, G.B.; Litchfield, D.W. Regulation of caspase pathways by protein kinase CK2: identification of proteins with overlapping CK2 and caspase consensus motifs. *Mol. Cell. Biochem.* **2011**, *356*, 159–67, doi:10.1007/s11010-011-0972-5.
  21. Turowec, J.P.; Vilks, G.; Gabriel, M.; Litchfield, D.W. Characterizing the convergence of protein kinase CK2 and caspase-3 reveals isoform-specific phosphorylation of caspase-3 by CK2 $\alpha$ : implications for pathological roles of CK2 in promoting cancer cell survival. *Oncotarget* **2013**, *4*, 560–71, doi:10.18632/oncotarget.948.

22. Patel, J.; Pathak, R.R.; Mujtaba, S. The biology of lysine acetylation integrates transcriptional programming and metabolism. *Nutr. Metab.* 2011, *8*.
23. Zhao, Y.; Jensen, O.N. Modification-specific proteomics: Strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 2009, *9*, 4632–4641.
24. Schölz, C.; Weinert, B.T.; Wagner, S.A.; Beli, P.; Miyake, Y.; Qi, J.; Jensen, L.J.; Streicher, W.; McCarthy, A.R.; Westwood, N.J.; et al. Acetylation site specificities of lysine deacetylase inhibitors in human cells. *Nat. Biotechnol.* **2015**, *33*, 415–425, doi:10.1038/nbt.3130.
25. Suraweera, A.; O’Byrne, K.J.; Richard, D.J. Combination therapy with histone deacetylase inhibitors (HDACi) for the treatment of cancer: Achieving the full therapeutic potential of HDACi. *Front. Oncol.* 2018, *8*.
26. Riley, N.M.; Coon, J.J. Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. *Anal. Chem.* **2016**, *88*, 74–94, doi:10.1021/acs.analchem.5b04123.
27. Kim, S.C.; Sprung, R.; Chen, Y.; Xu, Y.; Ball, H.; Pei, J.; Cheng, T.; Kho, Y.; Xiao, H.; Xiao, L.; et al. Substrate and Functional Diversity of Lysine Acetylation Revealed by a Proteomics Survey. *Mol. Cell* **2006**, *23*, 607–618, doi:10.1016/j.molcel.2006.06.026.
28. Wiśniewski, J.R.; Zougman, A.; Krüger, S.; Ziolkowski, P.; Pudελko, M.; Bebenek, M.; Mann, M. Constitutive and dynamic phosphorylation and acetylation sites on NUCKS, a hypermodified nuclear protein, studied by quantitative proteomics. *Proteins* **2008**, *73*, 710–8, doi:10.1002/prot.22104.
29. Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M.L.; Rehman, M.; Walther, T.C.; Olsen, J. V.; Mann, M. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science (80-. )*. **2009**, *325*, 834–840, doi:10.1126/science.1175371.
30. Torres, M.P.; Dewhurst, H.; Sundararaman, N. Proteome-wide Structural Analysis of PTM Hotspots Reveals Regulatory Elements Predicted to Impact Biological Function and Disease. *Mol. Cell. Proteomics* **2016**, *15*, 3513–3528, doi:10.1074/mcp.M116.062331.
31. Solntsev, S.K.; Shortreed, M.R.; Frey, B.L.; Smith, L.M. Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *J. Proteome Res.* **2018**, *17*, 1844–1851, doi:10.1021/acs.jproteome.7b00873.
32. Grimes, M.; Hall, B.; Foltz, L.; Levy, T.; Rikova, K.; Gaiser, J.; Cook, W.; Smirnova, E.; Wheeler, T.; Clark, N.R.; et al. Integration of protein phosphorylation, acetylation, and methylation data sets to outline lung cancer signaling networks. *Sci. Signal.* **2018**, *11*, doi:10.1126/scisignal.aqa1087.
33. Kang, H.; Jung, J.-W.; Kim, M.K.; Chung, J.H. CK2 is the regulator of SIRT1 substrate-binding affinity, deacetylase activity and cellular response to DNA-damage. *PLoS One* **2009**, *4*, e6611, doi:10.1371/journal.pone.0006611.
34. Eom, G.H.; Cho, Y.K.; Ko, J.H.; Shin, S.; Choe, N.; Kim, Y.; Joung, H.; Kim, H.S.; Nam, K. II; Kee, H.J.; et al. Casein kinase-2 $\alpha$ 1 induces hypertrophic response by phosphorylation of histone deacetylase 2 S394 and its activation in the heart. *Circulation* **2011**, *123*, 2392–2403, doi:10.1161/CIRCULATIONAHA.110.003665.

35. Khan, D.H.; He, S.; Yu, J.; Winter, S.; Cao, W.; Seiser, C.; Davie, J.R. Protein kinase CK2 regulates the dimerization of histone deacetylase 1 (HDAC1) and HDAC2 during mitosis. *J. Biol. Chem.* **2013**, *288*, 16518–28, doi:10.1074/jbc.M112.440446.
36. Kim, H.S.; Chang, Y.G.; Bae, H.J.; Eun, J.W.; Shen, Q.; Park, S.J.; Shin, W.C.; Lee, E.K.; Park, S.; Ahn, Y.M.; et al. Oncogenic potential of CK2 $\alpha$  and its regulatory role in EGF-induced HDAC2 expression in human liver cancer. *FEBS J.* **2014**, *281*, 851–61, doi:10.1111/febs.12652.
37. Jia, Z.-M.; Ai, X.; Teng, J.-F.; Wang, Y.-P.; Wang, B.-J.; Zhang, X. p21 and CK2 interaction-mediated HDAC2 phosphorylation modulates KLF4 acetylation to regulate bladder cancer cell proliferation. *Tumour Biol.* **2016**, *37*, 8293–304, doi:10.1007/s13277-015-4618-1.
38. Chen, H.; Sharp, B.M. Content-rich biological network constructed by mining PubMed abstracts. *BMC Bioinformatics* **2004**, *5*, 147, doi:10.1186/1471-2105-5-147.
39. Doncheva, N.T.; Morris, J.H.; Gorodkin, J.; Jensen, L.J. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J. Proteome Res.* **2019**, *18*, 623–632, doi:10.1021/acs.jproteome.8b00702.
40. Ham, H.J.; Park, J.W.; Bae, Y.S. Defect of SIRT1-FoxO3a axis is associated with the production of reactive oxygen species during protein kinase CK2 downregulation-mediated cellular senescence and nematode aging. *BMB Rep.* **2019**, *52*, 265–270, doi:10.5483/BMBRep.2019.52.4.156.
41. Gouot, E.; Bhat, W.; Rufiange, A.; Fournier, E.; Paquet, E.; Nourani, A. Casein kinase 2 mediated phosphorylation of Spt6 modulates histone dynamics and regulates spurious transcription. *Nucleic Acids Res.* **2018**, *46*, 7612–7630, doi:10.1093/nar/gky515.
42. Wang, H.; Song, C.; Ding, Y.; Pan, X.; Ge, Z.; Tan, B.H.; Gowda, C.; Sachdev, M.; Muthusami, S.; Ouyang, H.; et al. Transcriptional regulation of JARID1B/KDM5B histone demethylase by ikaros, histone deacetylase 1 (HDAC1), and casein kinase 2 (CK2) in B-cell acute lymphoblastic leukemia. *J. Biol. Chem.* **2016**, *291*, 4004–4018, doi:10.1074/jbc.M115.679332.
43. Adenuga, D.; Rahman, I. Protein kinase CK2-mediated phosphorylation of HDAC2 regulates co-repressor formation, deacetylase activity and acetylation of HDAC2 by cigarette smoke and aldehydes. *Arch. Biochem. Biophys.* **2010**, *498*, 62–73, doi:10.1016/j.abb.2010.04.002.
44. Watabe, M.; Nakaki, T. Protein kinase CK2 regulates the formation and clearance of aggresomes in response to stress. *J. Cell Sci.* **2011**, *124*, 1519–1532, doi:10.1242/jcs.081778.
45. Eom, G.H.; Nam, Y.S.; Oh, J.G.; Choe, N.; Min, H.K.; Yoo, E.K.; Kang, G.; Nguyen, V.H.; Min, J.J.; Kim, J.K.; et al. Regulation of acetylation of histone deacetylase 2 by p300/CBP-associated factor/histone deacetylase 5 in the development of cardiac hypertrophy. *Circ. Res.* **2014**, *114*, 1133–1143, doi:10.1161/CIRCRESAHA.114.303429.
46. Hornbeck, P. V.; Kornhauser, J.M.; Tkachev, S.; Zhang, B.; Skrzypek, E.; Murray, B.; Latham, V.; Sullivan, M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-

- translational modifications in man and mouse. *Nucleic Acids Res.* **2012**, *40*, D261-70, doi:10.1093/nar/gkr1122.
47. Minguez, P.; Letunic, I.; Parca, L.; Garcia-Alonso, L.; Dopazo, J.; Huerta-Cepas, J.; Bork, P. PTMcode v2: A resource for functional associations of post-translational modifications within and between proteins. *Nucleic Acids Res.* **2015**, *43*, D494–D502, doi:10.1093/nar/gku1081.
  48. Markwell, S.M.; Ammer, A.G.; Interval, E.T.; Allen, J.L.; Papenberg, B.W.; Hames, R.A.; Castaño, J.E.; Schafer, D.A.; Weed, S.A. Cortactin phosphorylation by casein kinase 2 regulates actin-related protein 2/3 complex activity, invadopodia function, and tumor cell invasion. *Mol. Cancer Res.* **2019**, *17*, 987–1001, doi:10.1158/1541-7786.MCR-18-0391.
  49. Bian, Y.; Ye, M.; Wang, C.; Cheng, K.; Song, C.; Dong, M.; Pan, Y.; Qin, H.; Zou, H. Global screening of CK2 kinase substrates by an integrated phosphoproteomics workflow. *Sci. Rep.* **2013**, *3*, 3460, doi:10.1038/srep03460.
  50. Siddiqui-Jain, A.; Drygin, D.; Streiner, N.; Chua, P.; Pierre, F.; O'Brien, S.E.; Bliesath, J.; Omori, M.; Huser, N.; Ho, C.; et al. CX-4945, an orally bioavailable selective inhibitor of protein kinase CK2, inhibits prosurvival and angiogenic signaling and exhibits antitumor efficacy. *Cancer Res.* **2010**, *70*, 10288–10298, doi:10.1158/0008-5472.CAN-10-1893.
  51. Rusin, S.F.; Adamo, M.E.; Kettenbach, A.N. Identification of candidate casein kinase 2 substrates in mitosis by quantitative phosphoproteomics. *Front. Cell Dev. Biol.* **2017**, *5*, doi:10.3389/fcell.2017.00097.
  52. Rabalski, A.J. Quantitative Proteomic Characterization of CX-4945, a Clinical Stage Inhibitor of Protein Kinase CK2, 2017.
  53. Giansanti, P.; Tsiatsiani, L.; Low, T.Y.; Heck, A.J.R. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protoc.* **2016**, *11*, 993–1006, doi:10.1038/nprot.2016.057.
  54. Cheng, A.; Grant, C.E.; Noble, W.S.; Bailey, T.L.; Hancock, J. MoMo: Discovery of statistically significant post-translational modification motifs. *Bioinformatics* **2019**, *35*, 2774–2782, doi:10.1093/bioinformatics/bty1058.
  55. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, doi:10.1093/nar/gkp335.
  56. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43*, W39–W49, doi:10.1093/nar/gkv416.
  57. Huang, H.; Arighi, C.N.; Ross, K.E.; Ren, J.; Li, G.; Chen, S.-C.; Wang, Q.; Cowart, J.; Vijay-Shanker, K.; Wu, C.H. iPTMnet: an integrated resource for protein post-translational modification network discovery. *Nucleic Acids Res.* **2018**, *46*, D542–D550, doi:10.1093/nar/gkx1104.
  58. Grant, C.E.; Bailey, T.L.; Noble, W.S. FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **2011**, *27*, 1017–1018, doi:10.1093/bioinformatics/btr064.
  59. Gouw, M.; Sámano-Sánchez, H.; Van Roey, K.; Diella, F.; Gibson, T.J.; Dinkel, H. Exploring Short Linear Motifs Using the ELM Database and Tools. *Curr. Protoc. Bioinforma.* **2017**, *58*, 8.22.1–8.22.35, doi:10.1002/cpbi.26.
  60. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515, doi:10.1093/nar/gky1049.

61. Gattiker, A.; Gasteiger, E.; Bairoch, A. ScanProsite: a reference implementation of a PROSITE scanning tool. *Appl. Bioinformatics* **2002**, *1*, 107–108.
62. Rice, P.; Longden, L.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* 2000, *16*, 276–277.
63. Krystkowiak, I.; Davey, N.E. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res.* **2017**, *45*, W464–W469, doi:10.1093/nar/gkx238.
64. Mészáros, B.; Simon, I.; Dosztányi, Z. Prediction of protein binding regions in disordered proteins. *PLoS Comput. Biol.* **2009**, *5*, doi:10.1371/journal.pcbi.1000376.
65. Berman, H.M.; Battistuz, T.; Bhat, T.N.; Bluhm, W.F.; Bourne, P.E.; Burkhardt, K.; Feng, Z.; Gilliland, G.L.; Iype, L.; Jain, S.; et al. The Protein Data Bank. *Acta Crystallogr. D. Biol. Crystallogr.* **2002**, *58*, 899–907.
66. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **2019**, *47*, D941–D947, doi:10.1093/nar/gky1015.
67. Maréchal, A.; Zou, L. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, doi:10.1101/cshperspect.a012716.
68. Akimov, V.; Barrio-Hernandez, I.; Hansen, S.V.F.; Hallenborg, P.; Pedersen, A.K.; Bekker-Jensen, D.B.; Puglia, M.; Christensen, S.D.K.; Vanselow, J.T.; Nielsen, M.M.; et al. Ubisite approach for comprehensive mapping of lysine and n-terminal ubiquitination sites. *Nat. Struct. Mol. Biol.* **2018**, *25*, doi:10.1038/s41594-018-0084-y.
69. Echeverria, I.; Liu, Y.; Gabelli, S.B.; Amzel, L.M. Oncogenic mutations weaken the interactions that stabilize the p110 $\alpha$ -p85 $\alpha$  heterodimer in phosphatidylinositol 3-kinase  $\alpha$ . *FEBS J.* **2015**, *282*, 3528–3542, doi:10.1111/febs.13365.
70. Fenyő, D.; Beavis, R.C. The GPMDB REST interface. *Bioinformatics* **2015**, *31*, 2056–8, doi:10.1093/bioinformatics/btv107.
71. Salizzato, V.; Borgo, C.; Cesaro, L.; Pinna, L.A.; Donella-Deana, A. Inhibition of protein kinase CK2 by CX-5011 counteracts imatinib-resistance preventing rpS6 phosphorylation in chronic myeloid leukaemia cells: New combined therapeutic strategies. *Oncotarget* **2016**, *7*, 18204–18218, doi:10.18632/oncotarget.7569.
72. Sondka, Z.; Bamford, S.; Cole, C.G.; Ward, S.A.; Dunham, I.; Forbes, S.A. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* 2018, *18*, 696–705.
73. Fabregat, A.; Jupe, S.; Matthews, L.; Sidiropoulos, K.; Gillespie, M.; Garapati, P.; Haw, R.; Jassal, B.; Korninger, F.; May, B.; et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **2018**, *46*, D649–D655, doi:10.1093/nar/gkx1132.
74. Aranda, B.; Blankenburg, H.; Kerrien, S.; Brinkman, F.S.L.; Ceol, A.; Chautard, E.; Dana, J.M.; De Las Rivas, J.; Dumousseau, M.; Galeota, E.; et al. PSICQUIC and PSISCORE: Accessing and scoring molecular interactions. *Nat. Methods* 2011, *8*, 528–529.
75. Shannon, P. PSICQUIC: Proteomics Standard Initiative Common QUery InterfaCe. 2019.

76. Henriksen, P.; Wagner, S.A.; Weinert, B.T.; Sharma, S.; Bačinskaja, G.; Rehman, M.; Juffer, A.H.; Walther, T.C.; Lisby, M.; Choudhary, C. Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **2012**, *11*, 1510–1522, doi:10.1074/mcp.M112.017251.
77. Nakayasu, E.S.; Wu, S.; Sydor, M.A.; Shukla, A.K.; Weitz, K.K.; Moore, R.J.; Hixson, K.K.; Kim, J.-S.; Petyuk, V.A.; Monroe, M.E.; et al. A method to determine lysine acetylation stoichiometries. *Int. J. Proteomics* **2014**, *2014*, 730725, doi:10.1155/2014/730725.
78. Dengler, V.L.; Galbraith, M.D.; Espinosa, J.M. Transcriptional regulation by hypoxia inducible factors. *Crit. Rev. Biochem. Mol. Biol.* *2014*, *49*, 1–15.
79. Black, J.C.; Mosley, A.; Kitada, T.; Washburn, M.; Carey, M. The SIRT2 Deacetylase Regulates Autoacetylation of p300. *Mol. Cell* **2008**, *32*, 449–455, doi:10.1016/j.molcel.2008.09.018.
80. Hansen, K.; Farkas, T.; Lukas, J.; Holm, K.; Rönstrand, L.; Bartek, J. Phosphorylation-dependent and -independent functions of p130 cooperate to evoke a sustained G1 block. *EMBO J.* **2001**, *20*, 422–32, doi:10.1093/emboj/20.3.422.
81. Farkas, T.; Hansen, K.; Holm, K.; Lukas, J.; Bartek, J. Distinct phosphorylation events regulate p130- and p107-mediated repression of E2F-4. *J. Biol. Chem.* **2002**, *277*, 26741–52, doi:10.1074/jbc.M200381200.
82. Fung, S.Y.S.; Kitagawa, M.; Liao, P.-J.; Wong, J.; Lee, S.H. Opposing Activities of Aurora B Kinase and B56-PP2A Phosphatase on MKlp2 Determine Abscission Timing. *Curr. Biol.* **2017**, *27*, 78–86, doi:10.1016/j.cub.2016.10.042.
83. Kalousi, A.; Mylonis, I.; Politou, A.S.; Chachami, G.; Paraskeva, E.; Simos, G. Casein kinase 1 regulates human hypoxia-inducible factor HIF-1. *J. Cell Sci.* **2010**, *123*, 2976–86, doi:10.1242/jcs.068122.
84. Lorenz, K.; Lohse, M.J.; Quitterer, U. Protein kinase C switches the Raf kinase inhibitor from Raf-1 to GRK-2. *Nature* **2003**, *426*, 574–9, doi:10.1038/nature02158.
85. Huang, J.; Mahavadi, S.; Sriwai, W.; Grider, J.R.; Murthy, K.S. Cross-regulation of VPAC(2) receptor desensitization by M(3) receptors via PKC-mediated phosphorylation of RKIP and inhibition of GRK2. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2007**, *292*, G867-74, doi:10.1152/ajpgi.00326.2006.
86. Li, J.; Taylor, I.A.; Lloyd, J.; Clapperton, J.A.; Howell, S.; MacMillan, D.; Smerdon, S.J. Chk2 oligomerization studied by phosphopeptide ligation: implications for regulation and phosphodependent interactions. *J. Biol. Chem.* **2008**, *283*, 36019–30, doi:10.1074/jbc.M804075200.
87. Zhang, H.; Wu, Z.; Lu, J.Y.; Huang, B.; Zhou, H.; Xie, W.; Wang, J.; Shen, X. DEAD-Box Helicase 18 Counteracts PRC2 to Safeguard Ribosomal DNA in Pluripotency Regulation. *Cell Rep.* **2020**, *30*, 81-97.e7, doi:10.1016/j.celrep.2019.12.021.
88. Gavali, S.; Cowart, J.; Chen, C.; Ross, K.E.; Arighi, C.; Wu, C.H. RESTful API for iPTMnet: a resource for protein post-translational modification network discovery. *Database (Oxford)*. **2020**, *2020*, doi:10.1093/database/baz157.
89. Munoz-Torres, M.; Carbon, S. Get GO! retrieving GO data using AmiGO, QuickGO, API, files, and tools. In *Methods in Molecular Biology*; Humana Press



- Inc., 2017; Vol. 1446, pp. 149–160.
90. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **2012**, *16*, 284–287, doi:10.1089/omi.2011.0118.
  91. Carlson, M. org.Hs.eg.db: Genome wide annotation for Human. 2019.
  92. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13, doi:10.1093/nar/gkn923.
  93. Turowec, J.P.; Duncan, J.S.; French, A.C.; Gyenis, L.; St. Denis, N.A.; Vilk, G.; Litchfield, D.W. Protein kinase CK2 is a constitutively active enzyme that promotes cell survival: Strategies to identify CK2 substrates and manipulate its activity in mammalian cells. In *Methods in Enzymology*; Academic Press Inc., 2010; Vol. 484, pp. 471–493.
  94. Orchard, S.; Kerrien, S.; Abbani, S.; Aranda, B.; Bhate, J.; Bidwell, S.; Bridge, A.; Briganti, L.; Brinkman, F.S.L.; Cesareni, G.; et al. Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. *Nat. Methods* **2012**, *9*, 345–350, doi:10.1038/nmeth.1931.
  95. Schöfl, G. reutils: Talk to the NCBI EUtils. 2016.
  96. Franceschini, A.; Szklarczyk, D.; Frankild, S.; Kuhn, M.; Simonovic, M.; Roth, A.; Lin, J.; Minguez, P.; Bork, P.; Von Mering, C.; et al. STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **2013**, *41*, doi:10.1093/nar/gks1094.
  97. Yu, G.; Li, F.; Qin, Y.; Bo, X.; Wu, Y.; Wang, S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* **2010**, *26*, 976–8, doi:10.1093/bioinformatics/btq064.
  98. Carlson, M.; Bioconductor Package Maintainer UniProt.ws: R Interface to UniProt Web Services. 2019.
  99. Ooms, J. curl: A Modern and Flexible Web Client for R. 2019.
  100. Kustatscher, G.; Grabowski, P.; Schrader, T.A.; Passmore, J.B.; Schrader, M.; Rappsilber, J. Co-regulation map of the human proteome enables identification of protein functions. *Nat. Biotechnol.* **2019**, *37*, 1361–1371, doi:10.1038/s41587-019-0298-5.
  101. I, P. “ggplot2” Based Plots with Statistical Details 2018.
  102. Tyanova, S.; Temu, T.; Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **2016**, *11*, 2301–2319, doi:10.1038/nprot.2016.136.
  103. H. Pagès; Aboyoun, P.; Gentleman, R.; DebRoy, S. Biostrings: Efficient manipulation of biological strings. 2019.
  104. Lindsay, H.; Burger, A.; Biyong, B.; Felker, A.; Hess, C.; Zaugg, J.; Chiavacci, E.; Anders, C.; Jinek, M.; Mosimann, C.; et al. CrispRVariants charts the mutation spectrum of genome engineering experiments. *Nat. Biotechnol.* **2016**, *34*, 701–2, doi:10.1038/nbt.3628.
  105. Humphrey, S.J.; Azimifar, S.B.; Mann, M. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nat. Biotechnol.* **2015**, *33*, 990–995, doi:10.1038/nbt.3327.
  106. Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.;

- Lajoie, G.A.; Ma, B. PEAKS DB: De novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics* **2012**, *11*, doi:10.1074/mcp.M111.010587.
107. Barsnes, H.; Vaudel, M. SearchGUI: A Highly Adaptable Common Interface for Proteomics Search and de Novo Engines. *J. Proteome Res.* **2018**, *17*, 2552–2555, doi:10.1021/acs.jproteome.8b00175.
108. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M.Y.; Geiger, T.; Mann, M.; Cox, J. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **2016**, *13*, 731–740.
109. Vaudel, M.; Burkhardt, J.M.; Zahedi, R.P.; Oveland, E.; Berven, F.S.; Sickmann, A.; Martens, L.; Barsnes, H. PeptideShaker enables reanalysis of MS-derived proteomics data sets: To the editor. *Nat. Biotechnol.* **2015**, *33*, 22–24.
110. Prigent, C.; Dimitrov, S. Phosphorylation of serine 10 in histone H3, what for? *J. Cell Sci.* **2003**, *116*, 3677–3685, doi:10.1242/jcs.00735.
111. Meggio, F.; Marin, O.; Pinna, L.A. Substrate specificity of protein kinase CK2. *Cell. Mol. Biol. Res.* **1994**, *40*, 401–409.
112. Meggio, F.; Pinna, L.A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **2003**, *17*, 349–68, doi:10.1096/fj.02-0473rev.
113. Sarno, S.; Vaglio, P.; Cesaro, L.; Marin, O.; Pinna, L.A. A multifunctional network of basic residues confers unique properties to protein kinase CK2. *Mol. Cell. Biochem.* **1999**, *191*, 13–9.
114. Salvi, M.; Sarno, S.; Cesaro, L.; Nakamura, H.; Pinna, L.A. Extraordinary pleiotropy of protein kinase CK2 revealed by weblogo phosphoproteome analysis. *Biochim. Biophys. Acta* **2009**, *1793*, 847–59, doi:10.1016/j.bbamcr.2009.01.013.
115. Ubersax, J.A.; Ferrell, J.E. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 530–541.
116. Hałasa, M.; Wawruszak, A.; Przybyszewska, A.; Jaruga, A.; Guz, M.; Kałafut, J.; Stepulak, A.; Cybulski, M. H3K18Ac as a Marker of Cancer Progression and Potential Target of Anti-Cancer Therapy. *Cells* **2019**, *8*, 485, doi:10.3390/cells8050485.
117. Uhle, S.; Medalia, O.; Waldron, R.; Dumdey, R.; Henklein, P.; Bech-Otschir, D.; Huang, X.; Berse, M.; Sperling, J.; Schade, R.; et al. Protein kinase CK2 and protein kinase D are associated with the COP9 signalosome. *EMBO J.* **2003**, *22*, 1302–12, doi:10.1093/emboj/cdg127.
118. Saccani, S.; Natoli, G. Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. *Genes Dev.* **2002**, *16*, 2219–2224, doi:10.1101/gad.232502.
119. Via, A.; Gherardini, P.F.; Ferraro, E.; Ausiello, G.; Tomba, G.S.; Helmer-Citterich, M. False occurrences of functional motifs in protein sequences highlight evolutionary constraints. *BMC Bioinformatics* **2007**, *8*, doi:10.1186/1471-2105-8-68.
120. Guerra, B.; Iwabuchi, K.; Issinger, O.-G. Protein kinase CK2 is required for the recruitment of 53BP1 to sites of DNA double-strand break induced by radiomimetic drugs. *Cancer Lett.* **2014**, *345*, 115–23, doi:10.1016/j.canlet.2013.11.008.
121. Lee, J.-H.; Mand, M.R.; Kao, C.-H.; Zhou, Y.; Ryu, S.W.; Richards, A.L.; Coon,

- J.J.; Paull, T.T. ATM directs DNA damage responses and proteostasis via genetically separable pathways. *Sci. Signal.* **2018**, *11*, doi:10.1126/scisignal.aan5598.
122. Lo, W.S.; Trievel, R.C.; Rojas, J.R.; Duggan, L.; Hsu, J.Y.; Allis, C.D.; Marmorstein, R.; Berger, S.L. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Mol. Cell* **2000**, *5*, 917–926, doi:10.1016/S1097-2765(00)80257-9.
123. Cheung, P.; Tanner, K.G.; Cheung, W.L.; Sassone-Corsi, P.; Denu, J.M.; Allis, C.D. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol. Cell* **2000**, *5*, 905–915, doi:10.1016/S1097-2765(00)80256-7.
124. Li, Y.; Li, Z.; Dong, L.; Tang, M.; Zhang, P.; Zhang, C.; Cao, Z.; Zhu, Q.; Chen, Y.; Wang, H.; et al. Histone H1 acetylation at lysine 85 regulates chromatin condensation and genome stability upon DNA damage. *Nucleic Acids Res.* **2018**, *46*, 7716–7730, doi:10.1093/nar/gky568.
125. Tran, A.D.A.; Marmo, T.P.; Salam, A.A.; Che, S.; Finkelstein, E.; Kabarriti, R.; Xenias, H.S.; Mazitschek, R.; Hubbert, C.; Kawaguchi, Y.; et al. HDAC6 deacetylation of tubulin modulates dynamics of cellular adhesions. *J. Cell Sci.* **2007**, *120*, 1469–1479, doi:10.1242/jcs.03431.

## 2.11 Supplemental Materials

**Supp. Table 2.1.** Chilibot and PubMed query: “ck2” AND “acetylation” with network representation and GO biological process enrichment.

Sheet	Name	Content
1	Table 1	Chilibot query: “ck2” AND “acetylation”. PubMed query: “ck2” AND “acetylation”; STRING representation 0.7 confidence, top 100. GO biological process enrichment of PubMed query: “ck2” AND “acetylation”; STRING representation 0.7 confidence, top 100.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.2.** Known CK2 target sites annotated in PhosphoSitePlus (PSP) and iPTMnet databases matching the K motif variants and their function (UniProt).

Sheet	Name	Content
1	Table 2	Known CK2 target sites annotated in PhosphoSitePlus (PSP) and iPTMnet databases matching the K motif variants and their function (UniProt). P: position; -p: phosphorylation.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.3.** K motif hits from *in silico* analysis of *in vitro* high-throughput CK2 kinase assay.

Sheet	Name	Content
1	Table 3	Sites identified by Bian <i>et al.</i> 2013 [49] that match the K motif variants before and after filtering of phosphorylation vs control ratio > 1.5. K motif hits (not including K motif, minimal hits). Workflow evaluation. Jurkat cells. HeLa cells.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.4.** Differentially inhibited phosphosites identified in three different phosphoproteomic studies of CX-4945 inhibition that match the K motif variants.

Sheet	Name	Content
1	Table 4	Percent of phosphosites identified in three different phosphoproteomics studies that match the K motif variants. All identified sites. Differentially inhibited phosphosites identified in three different phosphoproteomics studies of CX-4945 inhibition that match the K motif variants. HeLa cells 20 $\mu$ M CX-4945 for 1-4 h. U2OS cells, 30 $\mu$ M CX-4945 for 4 h. HeLa mitosis, 5 $\mu$ M CX-4945 for 45 min.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.5.** Motif discovery on acetylation sites identified in Scholz *et al.* 2015 using MoMo v5.0.5.

Sheet	Name	Content
1	Table 5	Motif discovery on acetylation sites identified in Scholz <i>et al.</i> 2015 using MoMo v5.0.5.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.6.** K motif hits annotated phosphorylated and acetylated in PSP database (September 2019); does not imply co-occurrence of phosphorylation and acetylation (September 2019).

Sheet	Name	Content
1	Table 6	K motif hits annotated phosphorylated and acetylated in PSP database (September 2019); does not imply co-occurrence of phosphorylation and acetylation. K motif extended hits targeted by acetylation and phosphorylation (co-occurrence not implied). K motif hits targeted by acetylation and phosphorylation (co-occurrence not implied).
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.7.** K motif hits annotated phosphorylated in PSP database (September 2019).

Sheet	Name	Content
1	Table7	K motif hits annotated phosphorylated in PSP database (September 2019). K motif extended hits targeted phosphorylation. K motif hits targeted by phosphorylation.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.8.** K motif hits annotated phosphorylated in the iPTMnet database (September 2019) reported acetylated in iPTMnet and/or the Scholz *et al.* 2015 data and corresponding PTM interplay from PTMcode v2 (October 2019).

Sheet	Name	Content
1	Table 8	K motif hits annotated phosphorylated in the iPTMnet database (September 2019) reported acetylated in iPTMnet and/or the Scholz <i>et al.</i> 2015 data and corresponding PTM interplay from PTMcode v2. Phosphorylation (iPTMnet) and acetylation (iPTMnet). Phosphorylation (iPTMnet) and acetylation (Scholz <i>et al.</i> 2015). PTM interplay retrieved from PTMcode v2 (October 2019).
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.9.** Kinase-substrate relationships corresponding to K motif hits annotated phosphorylated in the PSP and iPTMnet databases (September 2019).

Sheet	Name	Content
1	Table 9	Kinase-substrate relationships corresponding to K motif hits annotated phosphorylated in the iPTMnet database (September 2019).

		Kinase-substrate relationships corresponding to K motif hits annotated phosphorylated in PSP database (September 2019).
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.10.** K motif representation on the human proteome (September 2019) with FIMO v5.0.5 by using the MoMo v5.0.5 motifs discovered in the acetylation sites identified in Scholz *et al.* 2015 as input; unique sequences. UniProtKB human sequences reviewed and isoforms April 26, 2019.

Sheet	Name	Content
1	Table 10	K motif representation on the human proteome with FIMO v5.0.5 by using the MoMo v5.0.5 motifs discovered in the acetylation sites identified in Scholz <i>et al.</i> 2015 as input; unique sequences. UniProtKB human sequences reviewed and isoforms April 26, 2019. All K motif hits. K motif extended.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.11.** K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off  $\geq 0.4$ ).

Sheet	Name	Content
1	Table 11	K motif representation on the human proteome using SLiMSearch4; unique sites on canonical sequences (IUPred disordered cut-off $\geq 0.4$ ). All K motif hits.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.12.** UniProtKB domain and COSMIC mutations (February 2020) functional features of K motif hits.

Sheet	Name	Content
1	Table 12	Domain and COSMIC mutations functional features of K motif hits. Functional domains that contain K motif hits (UniProtKB). COSMIC mutations, February 2020. K to E mutations.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.13.** Functional features of sequences conforming to K motif hits: Regulatory Sites, CGS, KDAC-KAT, and Sites in Pathways.

Sheet	Name	Content
1	Table 13	Functional features of sequences conforming to K motif hits: Regulatory Sites, CGS, KDAC-KAT, Sites in Pathways. Regulatory role for K motif hits retrieved from PSP (September 2019). KDAC and KAT interactors of K motif hits. K motif hits in Cancer Gene Census (February 2020). PTM in Reactome pathways.
Appendix A		Supplemental Material – Chapter 2

**Supp. Table 2.14.** Summary of CK2 candidate sites matching the K motif.

Sheet	Name	Content
1	Table 14	Summary of CK2 candidate sites matching the K motif. Known CK2 substrates curated from PSP and iPTMnet databases. Human <i>in vitro</i> K motif hits for p+2 and K motif extended p+1 and p+3 variants. Differentially inhibited phosphosites identified in three different phosphoproteomics studies of CX-4945 inhibition that match the K motif variants. PSP K motif extended hits targeted by acetylation and phosphorylation (co-occurrence not implied). PSP K motif extended hits targeted phosphorylation. All candidates. GO molecular function, level 5: activity, histone, and chromatin.

Appendix A	Supplemental Material – Chapter 2
------------	-----------------------------------

**Supp. Table 2.15.** CK2 candidate sites modulated by lysine deacetylase treatment in U2OS cells.

Sheet	Name	Content
1	Table 15	CK2 candidate sites modulated by lysine deacetylase treatment in U2OS cells.
Appendix A	Supplemental Material – Chapter 2	
Appendix B	MS identification and quantification data	

**Supp. Table 2.16.** Peptides with acetylation and phosphorylation sites co-occurrences.

Sheet	Name	Content
1	Table 16	Peptides with acetylation and phosphorylation sites co-occurrences.
Appendix A	Supplemental Material – Chapter 2	
Appendix B	MS identification data	

**Supp. Table 2.17.** GO level 5 classification of proteins with co-occurring phosphorylation and acetylation sites.

Sheet	Name	Content
1	Table 17	GO level 5 classification of proteins with co-occurring phosphorylation and acetylation sites. GO biological process, level 5, represented in the phosphorylated and acetylated proteins. GO cellular compartments, level 5, represented in the phosphorylated and acetylated proteins.
Appendix A	Supplemental Material – Chapter 2	

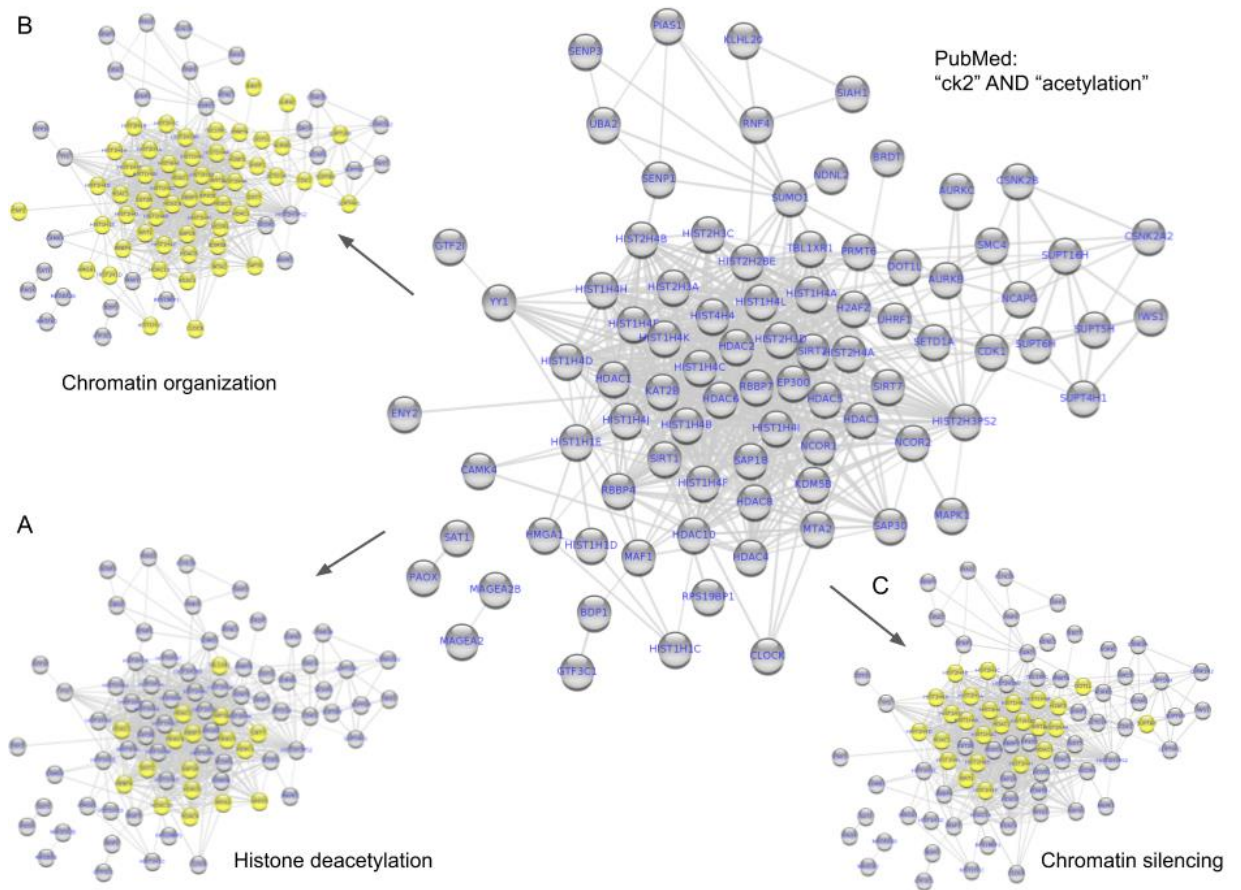
**Supp. Table 2.18.** Known sites (iPTMnet and PSP), known kinase-substrate relationships (iPTMnet), and known regulatory sites (PSP) corresponding to the identified sites in acetylated and phosphorylated peptides (September and December 2019).

Sheet	Name	Content
1	Table 18	Known sites (iPTMnet and PSP), known kinase-substrate relationships (iPTMnet), and known regulatory sites (PSP) corresponding to the identified sites in acetylated and phosphorylated peptides (September and December 2019). Known kinase-substrate relationships corresponding to the identified sites iPTMnet, December 2019. Regulatory sites annotate in PSP corresponding to the identified sites, December 2019. Known sites iPTMnet. Known sites PSP.
Appendix A	Supplemental Material – Chapter 2	

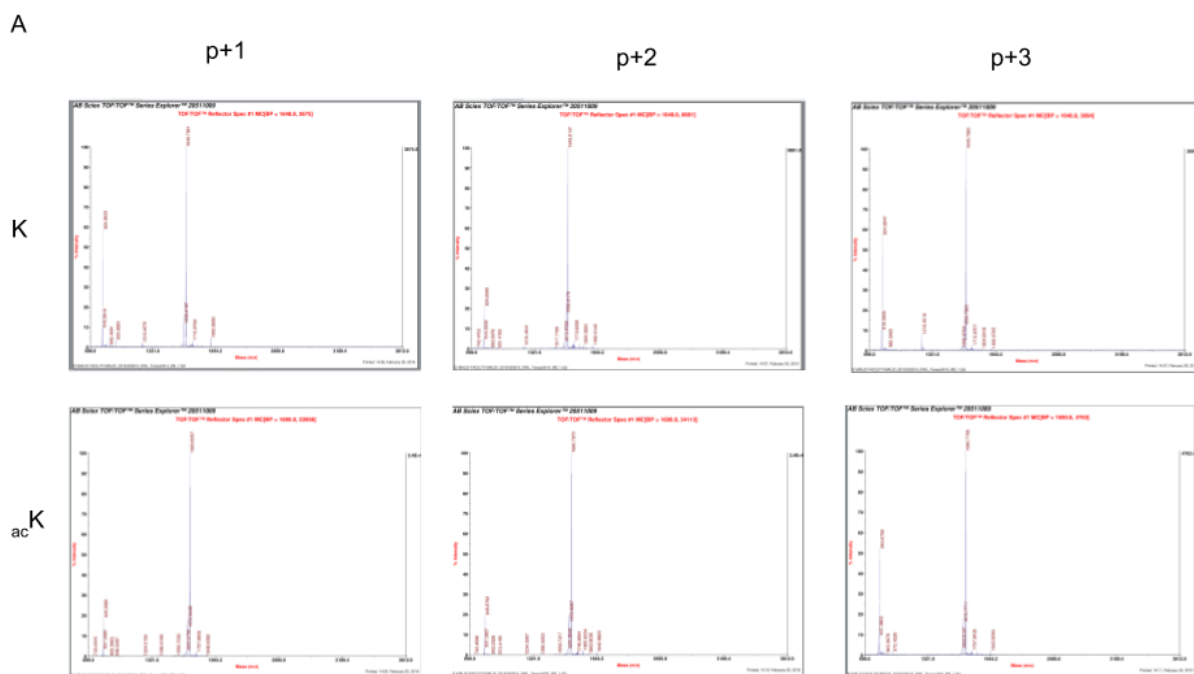
**Supp. Table 2.19.** CK2 consensus and K motif hits found *in vivo* and their corresponding protein function and site modification information.

Sheet	Name	Content
1	Table 19	CK2 consensus and K motif hits found <i>in vivo</i> and their corresponding protein function and site modification information. K motif hits acetylated and phosphorylated and corresponding functional information. CK2 consensus hits phosphorylated located nearby acetylation sites. UniProtKB function. Functional link to CK2 for [ST][DE].[DE] and extended hits. Known modification sites for [ST][DE].[DE] and extended hits.
Appendix A	Supplemental Material – Chapter 2	

## Figures



**Supp. Figure 2.1.** Functional association STRING network corresponding to the PubMed query: “ck2” AND “acetylation” represented with Cytoscape v3.8.0; confidence cut-off score: 0.7, and a maximum number of proteins: 100. A) Histone deacetylation subnetwork, 17 proteins (FDR < 0.001), B) Chromatin organization subnetwork, 59 proteins (FDR < 0.001), C) Chromatin silencing subnetwork, 25 proteins (FDR < 0.001).

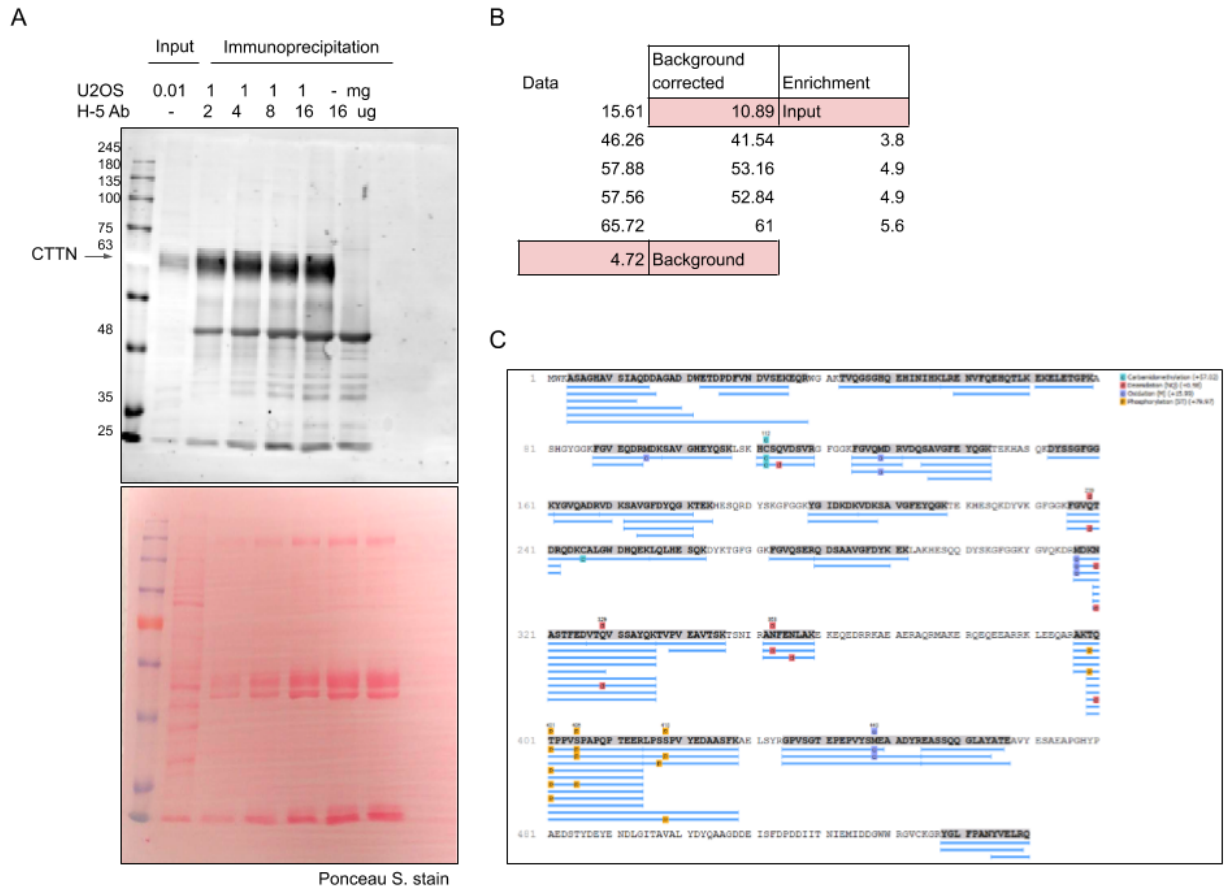


B

Peptide	K position and modification	Monoisotopic mass	Peaks
E10	p+1, K	1649.65	Mass: 1648.7961. Peaks: 1630.8197 (16%), $\Delta=18$ Dehydration (-H <sub>2</sub> O)
E11	p+2, K	1649.65	Mass: 1648.8147. Peaks: 1630.8179 (22%) $\Delta=18$ Dehydration (-H <sub>2</sub> O)
E12	p+3, K	1649.65	Mass: 1648.7863. Peaks: 1630.7920 (16%) $\Delta=18$ Dehydration (-H <sub>2</sub> O)
F1	p+1, <sub>ac</sub> K	1691.65	Mass: 1690.8057. Peaks: 1672.8446 (18%) $\Delta=18$ Dehydration (-H <sub>2</sub> O)
F2	p+2, <sub>ac</sub> K	1691.65	Mass: 1690.7970. Peaks: 1672.8267 (20%) $\Delta=18$ Dehydration (-H <sub>2</sub> O)
F3	p+3, <sub>ac</sub> K	1691.65	Mass: 1690.7706. Peaks: 1672.7711 (18%) $\Delta=18$ Dehydration (-H <sub>2</sub> O)
D12	p+2, K (endogenous)	1690.79	Mass: 1689.9139. Peaks: 1672.9467 (11%), $\Delta=-17$ N6-(L-isoglutamyl)-L-lysine (loss ammonia); 1618.8661 (30%), $\Delta=-71$ Ala?
E1	p+2, <sub>ac</sub> K (endogenous)	1732.79	Mass: 1731.9475. Peaks: 1714.9934 (10%), $\Delta=-17$ N6-(L-isoglutamyl)-L-lysine (loss ammonia); 1660.9017 (15%), $\Delta=-71$ Ala?; 1360.7061, $\Delta=-371.2414$ ?; 1275.6417 (39%), $\Delta=-456.3058$ ?
E2	p+2, K (endogenous)	1701.86	Mass: 1700.9750.
E3	p+2, <sub>ac</sub> K (endogenous)	1743.86	Mass: 1742.9775. Peaks: 1995.0814 (100%), $\Delta=252.1$ Fmoc-Arg(Pbf)-OH, 1895.9968 (10%), $\Delta=153$ Nps protects amino, hydroxy and thiol groups

**Supp. Figure 2.2.** MALDI quality analysis of *in-solution* peptides, > 1000 m/z with > 3% intensity. 1 Dalton difference amide C-terminal. P/p: position, ac: acetylation.

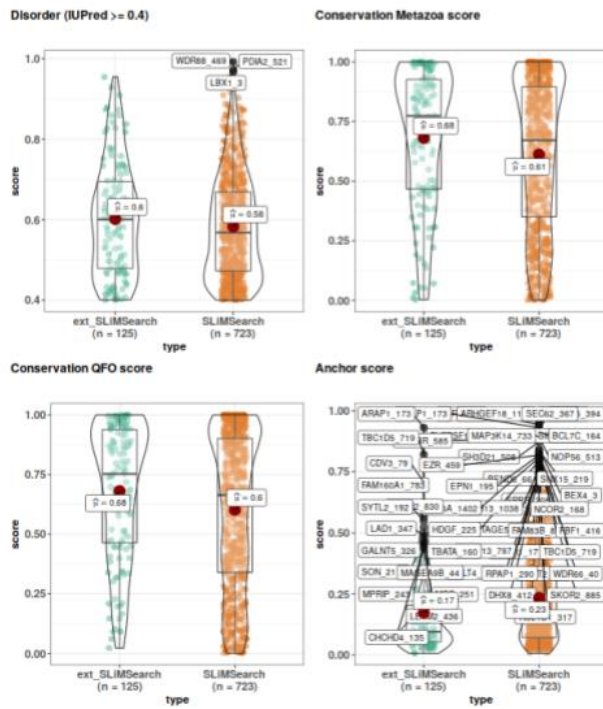




**Supp. Figure 2.3.** CTTN modification status in U2OS cells. A) Antibody optimization using U2OS lysate. B) Quantification of antibody optimization. C) PEAKS coverage and CTTN modifications after in-gel digestion and LC-MS/MS analysis.

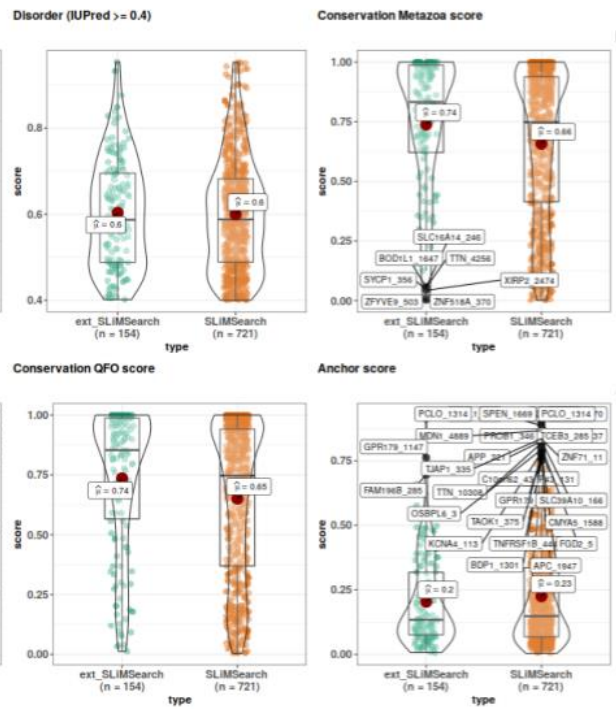
A

p+1



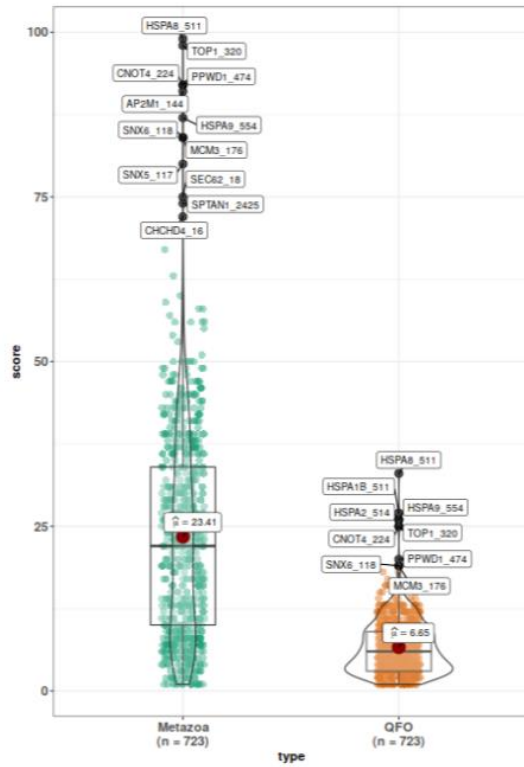
B

p+3

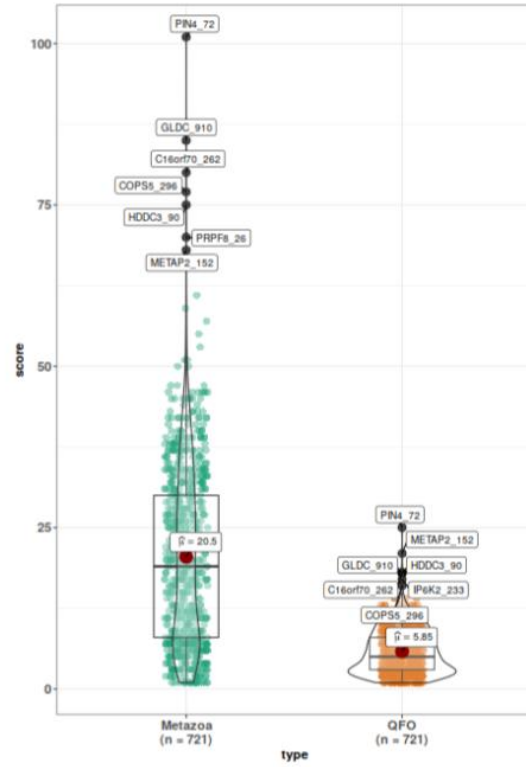


**Supp. Figure 2.4.** SLiMSearch4 scores for K motif variants A) +1 and B) +3; K motif extended (ext\_SLiMSearch) and all hits (SLiMSearch). Left-Right: Disorder score (IUPred). Conservation in Metazoa. Conservation in Quest for Orthologs (QFO), Anchor score (disorder-order transition upon binding). Outliers coefficient (Tukey's method) = 1.5.

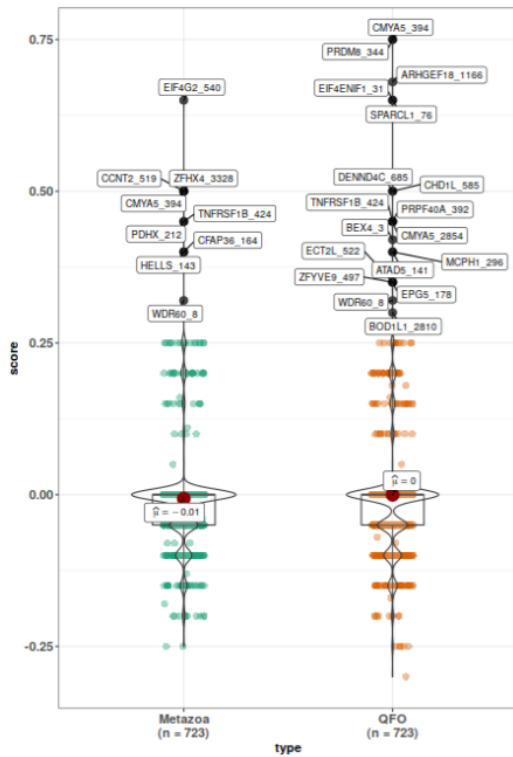
**A** Conserved counter (QFO, Metazoa)



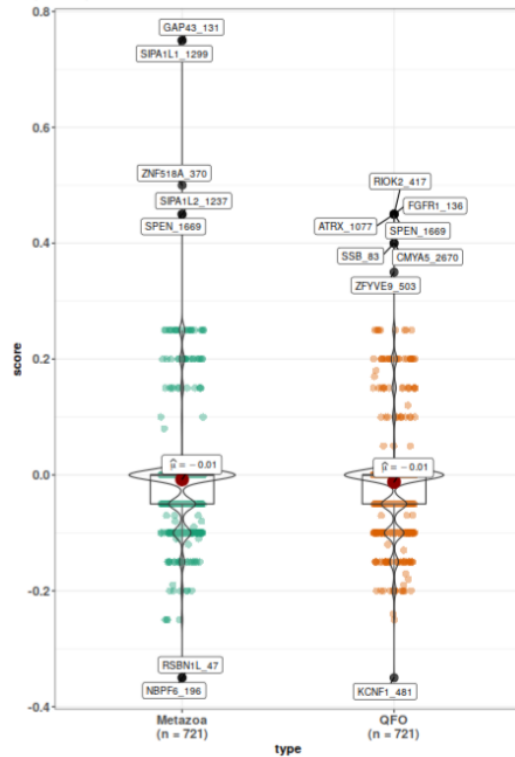
**B** Conserved counter (QFO, Metazoa)



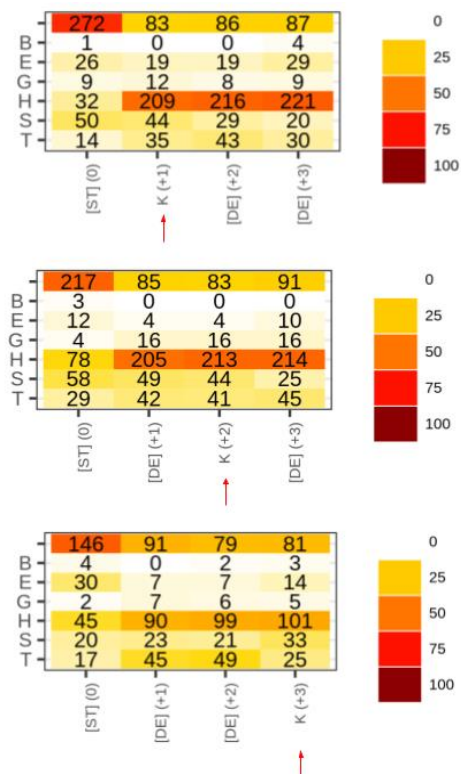
Conservation, motif vs flank (QFO, Metazoa)



Conservation, motif vs flank (QFO, Metazoa)



**Supp. Figure 2.5.** K motif variants A) +1 and B) +3; K motif extended (ext\_SLiMSearch) and all hits (SLiMSearch). Top-Bottom: Conserved counter (Outliers coefficient = 1.5), Conservation motif vs flank (Outliers coefficient (Tukey's method) = 5).



**Supp. Figure 2.6.** PDB secondary structure information for K motif hits found in the human proteome scan. H = alpha helix; B = residue in isolated beta-bridge; E = extended strand, participates in beta ladder; G = 3-helix (3/10 helix); I = 5 helix (pi helix); T = hydrogen bonded turn; S = bend; "space" = unordered.

## Chapter 3

### 3 Protein kinase CK2: Regulatory Role in Cellular Signaling and its Impact on Chromatin Organization

An intricate ensemble of biological networks orchestrates cellular events. These networks, which include metabolic, cell signaling, protein-protein interaction, genetic interaction, and gene/transcriptional regulatory networks are precisely regulated in cells and undergo significant rewiring to promote cancer progression [1,2]. High-throughput technologies and system analyses continue to expand our understanding of these networks and how they are modified in cancer. Interestingly, many of the networks/subnetworks converge, demonstrating the emergent properties and robustness of biological systems [3]. For instance, attention has recently been drawn to how distinct signaling pathways that are deregulated in cancer target histones through post-translational modifications (histone code), which could ultimately change the cell's epigenetic marks; the latter refers to chemical modifications of histone proteins and DNA that dictate gene expression without affecting the gene sequence [2]. Epigenetic changes are known to be key players in cancer by mediating, among other effects, the silencing of tumor suppressor genes [4,5]. However, when considering the hereditary aspect of epigenetics it is unclear how histone modifications could transmit the information of the chromatin structure from mother to daughter cells, how widespread are the mechanisms involved, and if other mediators are required for the transmission of information [6]. Nevertheless, by changing the modification status of histones the deregulated signaling pathways in cancer directly impact processes like transcription and DNA repair through chromatin (DNA and histones) remodeling i.e., local changes in the structure of the chromatin [7].

Histones represent major structural components within the nucleus, where they coordinate the assembly of DNA into chromatin. In the cell, histone proteins are modified by the addition of post-translational modifications (PTMs) such as phosphorylation, acetylation, methylation, and ubiquitination [5]. The DNA can also be modified through methylation. The addition, removal, and recognition of these chemical modifications of histones and DNA are mediated by several catalytic or binding activities often referred to

as writers (e.g., acetyltransferases and methyltransferases), erasers (e.g., deacetylases and demethylases), and readers (e.g., ADD, Chromodomain, Bromodomain, 14-3-3, MDB, and BIR), respectively [8,9]. Similar to histones, the histone-modifying enzymes, and more generally the chromatin-modifying enzymes (which include DNA-modifying catalytic activities in addition to histone-modifying activities) can also be the targets of aberrant signaling pathways in cancer [10], adding another layer of complexity. The combined action of these enzymes mediates chromatin remodeling, characterized by dynamic structural changes to the eukaryotic chromatin which are critical for both transcriptional regulation and chromosome segregation. Consequently, a comprehensive understanding of the regulatory relationships that modulate chromatin structure and dynamics is a prerequisite for devising novel and effective therapies in cancer. Towards this goal, systematic phosphoproteomic and acetylome studies profiling the effect of kinase and lysine deacetylase inhibitors can provide us with proteome-wide information regarding the modification status of chromatin-modifying enzymes and their interaction partners. A detailed review of how signaling pathways modulate histone phosphorylation upon upstream kinase activation can be found elsewhere [2]. In contrast, the information pertaining to chromatin-modifying enzymes and their regulation through various signaling pathways and post-translational modifications spans a functionally more diverse set of proteins [6,11,12] and has not yet been comprehensively summarized.

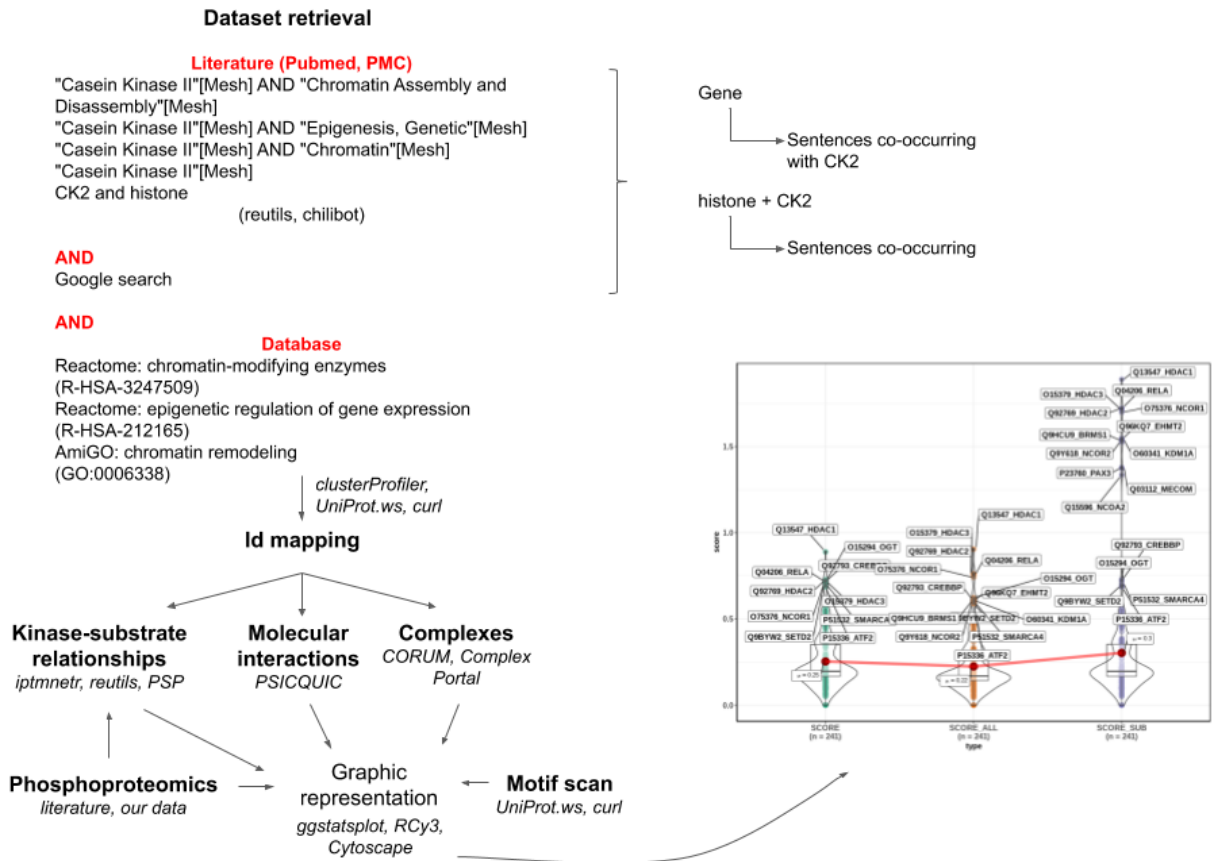
One protein kinase that has been linked to chromatin remodeling and transcriptional regulation is protein kinase CK2 [13,14], a constitutively expressed and ubiquitous protein Ser/Thr kinase that targets acidophilic motifs (pS/T-[DE]-X-[DE]) found in many nuclear and cytoplasmic proteins [15]. CK2 also displays Tyr kinase activity evidenced by phosphorylation of H2A in yeast and mammals [16]. CK2 typically forms a tetrameric holoenzyme composed of two catalytic (encoded by CSNK2A1 and CSNK2A2 genes, respectively) and two regulatory (encoded by CSNK2B) subunits [17]. The upregulation of CK2 catalytic subunits has been described in cancer where it displays pro-survival and antiapoptotic roles [18]. Both low and high-throughput experiments have corroborated the extensive network of phosphoproteins modulated by CK2 which includes histone-modifying enzymes such as lysine deacetylases and demethylases [19,20]. In addition, a regulatory link was established between CK2-dependent phosphorylation and lysine

acetylation. To extend these observations, we sought to illustrate and summarize the connections of CK2-dependent signaling with chromatin remodeling (i.e., dynamic structural changes to eukaryotic chromatin occurring throughout the cell division cycle; definition as in GO:0006338). Since inhibitors of both CK2 and chromatin-modifying enzymes are being assessed in the clinic [21,22], of interest are relationships between CK2 and these proteins that may be useful from the perspective of therapeutic potential or translational research.

### 3.1 Exploring the link between CK2 and chromatin organization

In order to approach this challenge, the functional link between CK2 and proteins involved in chromatin organization and epigenetics, processes tightly regulated by the balance of acetylation-deacetylation of histone/non-histone proteins in the cell, was investigated. For this, several bioinformatic tools and/or their output were applied to implement a workflow in R that integrates information retrieved from the literature and databases (Fig. 3.1). First, an automated literature search [23] was performed for finding the co-occurrence of relevant terms with CK2 using Pubmed and PMC (provides full-text information) (Supp. Table 3.1). Complementarily, the text mining tool Chilibot [24] was used to extract relationships between two specific terms e.g., “ck2” and “histone” (Supp. Table 3.1). Given that Google is a powerful search engine it was also included as a tool for literature search. Next, lists of proteins involved in chromatin organization and epigenetics were retrieved from AmiGO [25] and Reactome [26] databases. If necessary, the protein mentions extracted from the databases were mapped [27] to the common database identifiers UniProt accession numbers and identifiers, HUGO gene symbols, and Entrez GeneID. As a result, lists of proteins involved in chromatin organization and epigenetics were generated. These lists were searched for CK2 substrates, CK2 interactors, literature co-occurrence with CK2, members of relevant protein complexes, presence of CK2 consensus motif, and proteins changed in phosphoproteome studies as a result of CK2 manipulation (Fig. 3.1, see Chapter 2). Since database information may not be regularly updated depending on the source, the information retrieved from the literature in the first step was essential to expand the analysis and not to overlook relevant

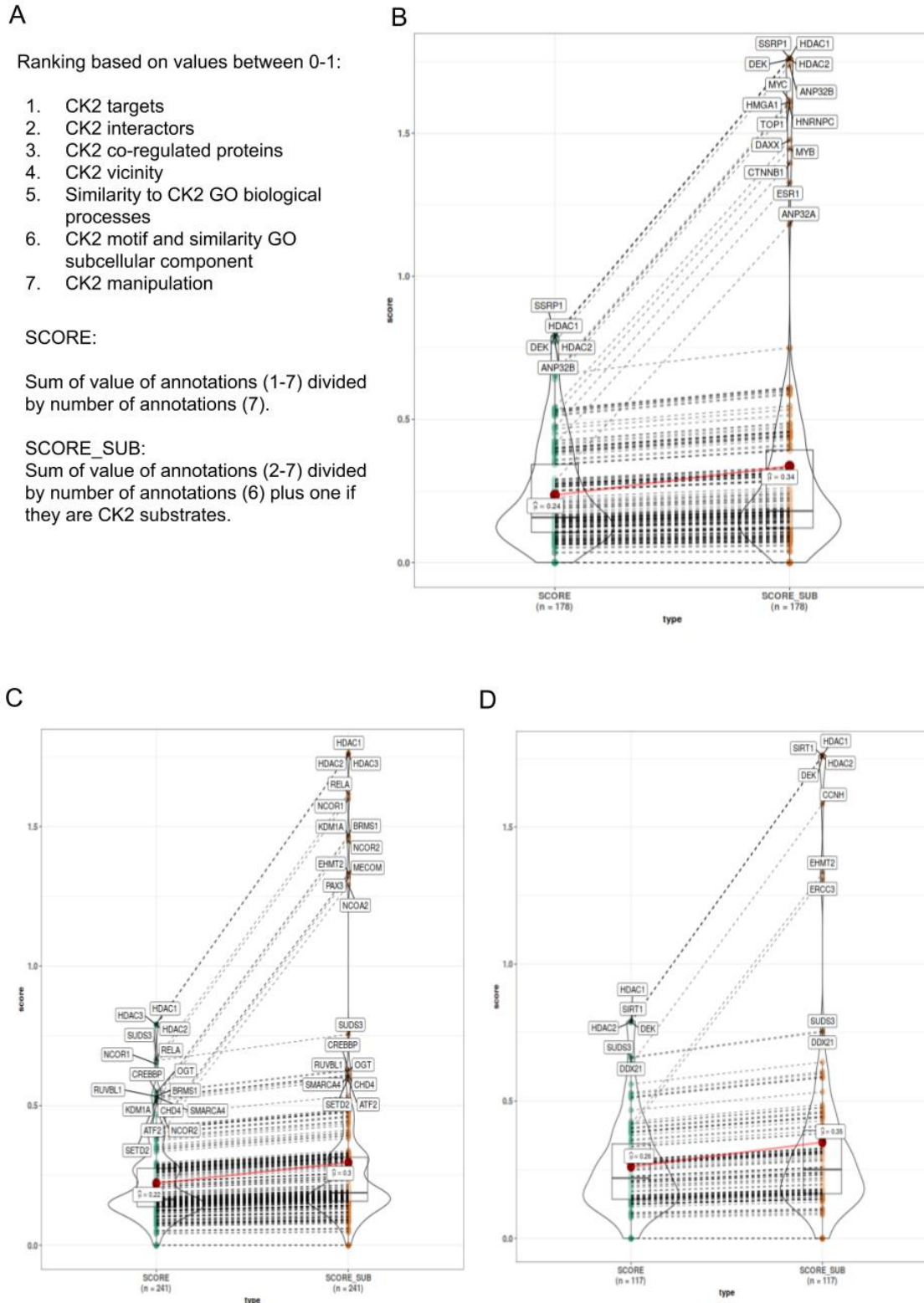
proteins and their association with CK2. If appropriate, the results were represented as networks such as the involvement of CK2 in the regulation of chromatin-remodeling complexes.



**Figure 3.1.** Bioinformatics workflow implemented for finding chromatin organization proteins functionally linked to CK2. Left: examples of what to expect from literature and ranking analysis.

All the proteins in the chromatin organization data sets (Supp. Table 3.2) were scored and ranked using the functional link to CK2 using the algorithm described in Chapter 2 (Fig. 3.2A). High scoring proteins that are known CK2 substrates involved in chromatin organization were observed (Fig. 3.2B-D, Supp. Table 3.3) along with a few high scoring chromatin organization proteins that are not known CK2 substrates but are strongly linked to CK2 (Fig. 3.2B-D, Supp. Table 3.3). Overall, several proteins functionally linked to CK2 were identified using the designed workflow, serving as a motivation to further summarize and investigate these findings.





**Figure 3.2.** Chromatin organization proteins functionally linked to CK2. A) Annotations included and score calculation B) AmiGO chromatin remodeling. C) Reactome

chromatin-modifying enzymes. D) Reactome epigenetics. Outlier coefficient (Tukey's method) = 1.5.

### 3.2 CK2-dependent signaling and histones

Histone proteins form the nucleosomes along with DNA. The nucleosome is the basic organizational unit of genomic DNA in the cell, which assembles into a higher protein-DNA structure known as chromatin. Many events regulate the organization of nucleosomes (i.e., assembly, arrangement, and disassembly) and thus the access to the DNA, including the modification of histones and chromatin-associated proteins, the incorporation of histone variants, histone replacement, histone folding, and ATP-dependent nucleosome remodeling. Such events confer dynamism to the chromatin throughout the cell cycle allowing and/or resulting from transcription regulation, DNA replication, DNA repair, cell cycle progression, etc. [28–30].

Histone modification by phosphorylation profoundly affects histone-DNA interaction by masking the positive charge of histones and it also provides a recognition site for chromatin-associated proteins [31]. Consequently, there has been considerable focus on investigating the interplay between kinases and histones and its impact on chromatin organization, i.e., chromatin specification, formation, and maintenance. To gain an appreciation of the extent of this regulation, the phosphorylation of histones is mediated by at least 20 different kinases including CK2 (Supp. Table 3.1), PRKCD, JNK, and AURKB [2]. Collectively, phosphorylation of histones has been linked to transcription, DNA damage response, apoptotic chromatin condensation, and mitosis [2]; all these processes require a certain degree of accessibility to the DNA.

In particular, the CK2-dependent phosphorylation of histones plays a role in DNA damage, transcriptional elongation, and in the regulation of transcription [31]. CK2 targets Ser1 residue of H4 in yeast in response to double-strand break DNA damage as part of the DNA-repair histone code playing a role in non-homologous end joining; phosphorylation of Ser1 by CK2 occurs independently of the cell cycle [32]. In yeast, phosphorylation of Ser1 negatively impacts H4K5, H4K8, and H4K12 acetylation by the acetyltransferase NuA4 [33]. The proposed mechanism is that the deacetylase complex Sin3/Rpd3 mediates deacetylation of H4 tails while the associated CK2 activity

phosphorylates Ser1 blocking further acetylation [33]. The impact of decreased H4 tail acetylation is not well understood but it could constitute a stable epigenetic mark based on a study in early murine development [34] and that along with phosphorylation of H2A at Ser129 promotes genome stability upon DNA damage [33]. CK2 also phosphorylates Tyr57 residue of H2A; the modification of this residue is thought to regulate the activity of the SAGA complex and thus transcriptional elongation [35]. The SAGA complex localizes to the basal transcriptional machinery recruited to gene loci, where it mediates histone acetylation of promoters and facilitates transcriptional elongation by deubiquitinating and acetylating histones [36]. In yeast and humans, both Tyr57Phe H2A mutant (non-phospho mimic) and CK2 inhibition were shown to enhance SAGA's H2B deubiquitinase activity [35]. In cancer, the enzymatic activities of the SAGA complex have been linked to tumorigenesis and poor prognosis [36]. Thus CK2-dependent phosphorylation of H2A, a direct target of the complex, points to a potential strategy to modulate SAGA's role in cancer. Additionally, a recent study showed *in vitro* evidence of phosphorylation of Thr11 residue of H3 by CK2 in response to nutritional stress in yeast [37]. As summarized here, beyond phosphorylation, CK2 has an overall impact in the histone code (Table 3.1), highlighting a crucial regulatory role of this kinase on chromatin dynamics.

**Table 3.1.** Histone modification functionally linked to CK2.

Histone	Modification	CK2	Proposed functional relevance
H4 (yeast)	Ser1-p	Phosphorylation by	Role in DNA damage repair. Possibly inhibitory of acetylation at H4K5, H4K8, and H4K12 [33].
H2A (yeast, human)	Tyr57-p	Phosphorylation by	Regulation of transcriptional elongation. Impact on H2B ubiquitination levels [35].
H3 (yeast)	Thr11-p	Phosphorylation by	Response to nutritional stress in yeast [37].
H3 (human)	Lys9-me3	Downregulation of	Senescence and DNA damage response [38-40].
H3 (yeast)	Lys56-ac	Downregulation of	H3/H4 turnover [43].

Furthermore, CK2 regulates histone methylation and acetylation with an impact on gene expression and histone recognition, folding, and turnover (Table 3.1). A link between CK2 and H3K9me has been described in cellular senescence and DNA damage response. Specifically, the histone methyltransferase EHMT2 (G9a) which catalyzes H3K9me1 and

H3K9me2 was recently found to be a target of CK2 at Ser211 in the DNA damage response [38]. On the other hand, CK2 downregulation in MCF-7 and HCT116 cells resulted in increased H3K9me3 due to an increased expression of the H3K9 trimethyltransferase SUV39H1 (KMT1A) and a decrease in EHMT1 and EHMT2 (GLP) [39,40]. H3K9me3 is a repressive histone-mark, localized to constitutive heterochromatin and tissue-specific genes, that prevents transcription factor binding [39].

Correspondingly, the CK2-driven increase of H3K9me3 was shown to promote the binding of CBX3 (HP1 $\gamma$ ) to H3K9me3, a marker of senescence-associated heterochromatin foci, resulting in the repression of cell cycle progression genes [40]. Similarly, CK2 phosphorylation of Ser11/14 in human and mouse cell lines enhanced CBX5 (HP1 $\alpha$ ) binding to nucleosomes containing H3K9me3 [41,42]. CBX5 is a component of heterochromatin that functions in repression. In yeast, CK2 depletion was shown to promote H3K56ac in transcribed genes (permissive mark) product of H3/H4 turnover, resulting in antisense and cryptic transcription [43]. Interestingly, the histone chaperone Spt6 required for H3/H4 turnover was shown to be a target of CK2. Mutants mimicking Spt6 N-terminal phosphorylation by CK2 decreased H3/H4 exchange favoring proper chromatin refolding after transcription [43].

Finally, CK2 could potentially impact histone acetylation through the phosphorylation of ANP32A (PP32) a component of the INHAT (inhibitor of histone acetyltransferases) complex [44]. The INHAT complex binds and protects histones tails from the acetyltransferases CREBBP, EP300, and PCAF [44,45]. INHAT is composed of two highly acidic proteins SET (isoform 1: TAFIA, isoform 2: TAF1A), and ANP32A [46], which contain low-complexity acidic regions (LCAR). Multiple acetylation and single phosphorylation of H3 tails have been shown to inhibit the binding of the INHAT complex [45]. ANP32A is a moonlighting protein, which functions as a regulator of transcription, apoptosis, PP2A activity, etc. [47]. CK2 phosphorylates the C-terminal residues Ser158 and Ser204 of ANP32A located at a region necessary for its tumor suppressor function and at the LCAR, respectively. In the context of transcriptional regulation, the deletion of the LCAR region did not have an impact on the binding to histones, but deletion of both LCAR and the tumor suppression region resulted in impaired binding [44]. In correspondence, CK2-dependent phosphorylation of Ser158

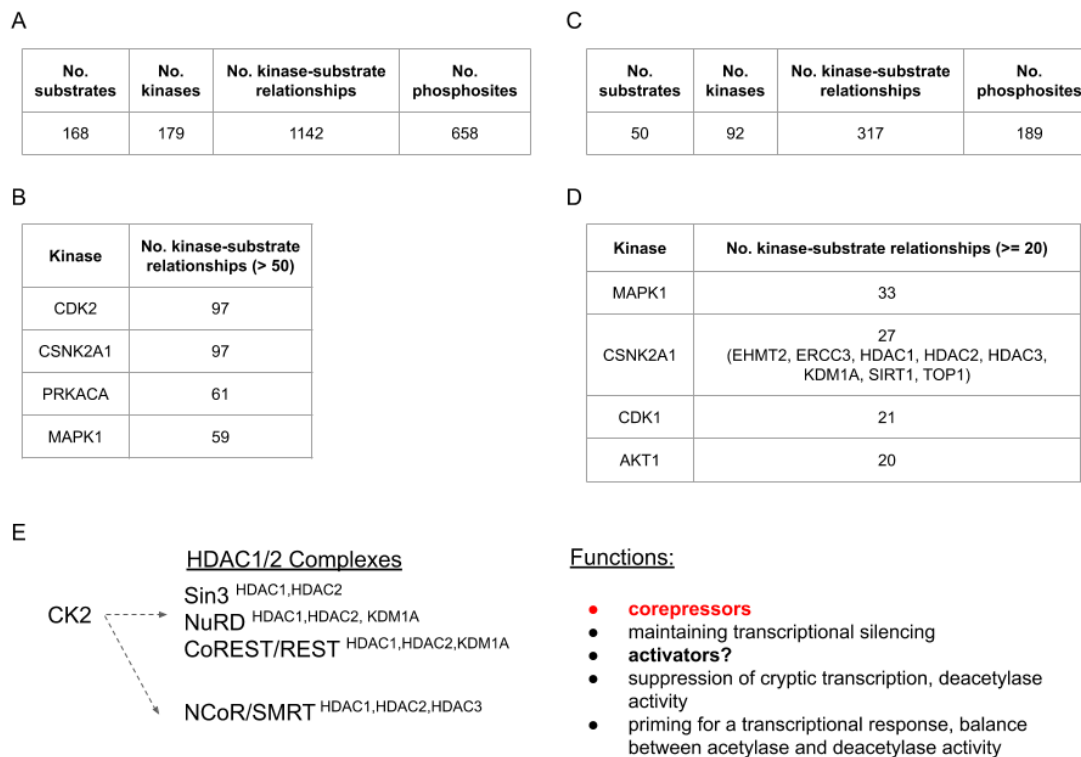
alone or in combination with the phosphorylation of Ser204 could increase the negative charge of the LCAR and its adjacent region enhancing INHAT activity; however, the implications of phosphorylation by CK2 are unknown.

### 3.3 CK2-dependent signaling and chromatin-modifying enzymes

The covalent modification of chromatin is dependent on a diverse group of molecular functions referred to as chromatin-modifying enzymes. This group contains “writer” and “eraser” proteins, which catalyze the addition and removal of the modifications, respectively. Examples of writers are kinases (as previously discussed for histones), acetyltransferases, and methyltransferases whereas their opposing erasers are phosphatases, deacetylases, and demethylases, respectively. These enzymes [6] are also subjected to covalent modifications themselves, acting as proxies for signal integration and storage in the chromatin [2]. The latter connects diverse stimuli such as metabolic changes, hypoxia, growth factors, etc., with changes to chromatin organization. Changes introduced to the chromatin can be sensed by a group of proteins called “readers” and/or acted upon by other chromatin-modifying proteins or associated molecular complexes [10]. For example, phosphorylation of the DNA methyltransferase DNMT1 by AKT increases its resistance to proteasome-dependent degradation, which could in part explain the aberrant DNA methylation observed in cancers with abnormal levels of AKT pathway activation [48]. In addition to writers and erasers, the chromatin-modifying group also includes proteins that mediate ATP-dependent nucleosome remodeling or “remodelers”. Modification of these remodelers can regulate their function in protein-DNA interaction disruption, nucleosome sliding, and deposition of histone variants [10]. For instance, the SAGA complex, mentioned before for its role in histone acetylation, can also acetylate the remodeler complex RSC in yeast [49].

To further investigate the extent to which proteins involved in chromatin organization (Reactome: R-HSA-3247509 and R-HSA-212165, AmiGO: GO:0006338) are modified and the signaling pathways involved the corresponding kinase-substrate relationships and the regulatory roles of such modifications were retrieved from the databases iPTMnet and PhosphoSitePlus (Supp Fig. 3.1). As a result, it was found that the chromatin

organization proteins display several PTM types, with more information available on phosphorylation (> 600 sites) (Fig. 3.3A). In this regard, around 160 chromatin organization proteins were found phosphorylated by approximately 170 kinases involving more than 1000 kinase-substrate relationships (Fig. 3.3A). Among the top kinases for which information is available, CK2 appeared with more than 90 phosphosites reported (Fig. 3.3B).



**Figure 3.3.** Merged kinase-substrate relationships found in PhosphoSitePlus (PSP) and iPTMnet databases. A) Summary of available information. B) Top kinases with information available. C) Merged kinase-substrate relationships found for proteins with EC numbers. D) Top kinases with information available evidencing from C. E) CK2 and HDAC-containing complexes; complex functions.

Next, information on chromatin organization proteins that display enzymatic activity was retrieved by selecting those with EC numbers. As a result, 50 enzymes were found to be modified by 92 kinases (Fig. 3.3C). Among the top kinases (sites  $\geq 20$ ) for which information is available CK2 could be found (Fig. 3.3D). CK2 was shown to phosphorylate not only the lysine deacetylases HDAC1, HDAC2, HDAC3, and SIRT1, but also the lysine demethylase KDM1A, and the lysine methyltransferase EHMT2 (Fig.

3.3C). These activities are critical in modulating chromatin organization by erasing/writing acetyl or methyl groups from histones and non-histone proteins. Histone acetylation affects nucleosome assembly in concert with histone chaperones and makes chromatin more permissive by disrupting histone-DNA binding, deacetylation and methylation oppose this effect [50]. Maintaining the balance between acetylation-deacetylation and methylation-demethylation is essential for transcription regulation as it dictates the accessibility to the DNA. Correspondingly, loss of histone acetylation and methylation have been identified as common hallmarks in cancer [5].

The observation of CK2-dependent phosphorylation of several chromatin-modifying enzymes reinforces the view of a global impact of CK2 signaling on chromatin dynamics under normal and abnormal conditions (Fig. 3.3D). However, additional information is still necessary to fully comprehend the impact of the CK2-dependent phosphorylation events in chromatin organization. Briefly, CK2 phosphorylates HDAC1, HDAC2, and HDAC3 at several sites, which increases their deacetylase activity and regulates their binding to members of transcriptional repressor complexes [51–54] (see next subsection), potentially modulating a plethora of biological processes and pathways that require HDAC activity. In fact, CK2 can have important regulatory roles in HDAC1/2 containing complexes that function as corepressors or activators depending on the signaling context and are essential in maintaining transcriptional silencing, suppressing cryptic transcription, and priming for a transcriptional response [55] (Fig. 3.3E). SIRT1 phosphorylation by CK2 modulates its enzymatic activity [56,57]. Phosphorylation of EHMT2 by CK2 has only been studied in DNA repair thus the role of phosphorylation in histone methylation is unknown [38] (see the previous section). ERCC3 (XPB) is an ATP-dependent DNA helicase that functions in general transcription and DNA repair; phosphorylation of Ser94, Ser751, and Ser892 by CK2 impacts its role in nucleotide excision repair by modulating DNA damage recognition [58,59]. Finally, CK2 also phosphorylates Ser10 and Ser506 residues of TOP1, a chromatin remodeling protein that releases the supercoiling and torsional tension of DNA. Phosphorylation by CK2 may enhance TOP1–DNA binding and relaxation of plasmid supercoils, but this has been tested *in vitro* only [60].

### 3.3.1 CK2-dependent phosphorylation of HDAC lysine deacetylases

CK2 regulates the activity of several lysine deacetylases, main regulators of transcription by mediating the deacetylation of both histones and non-histones proteins [61].

Acetylation of histones is thought to induce transcriptional activation by allowing the access of transcription factors to the chromatin [62]; conversely, deacetylation, in general, is regarded as an event that induces transcriptional repression of gene expression [63]. The HDAC superfamily of lysine deacetylases is composed of four classes (class I-IV) distributed across two structurally distinct protein families: the HDAC (histone deacetylase) and SIRT (SIR2 regulator) families. The classes are defined by their structural and sequence homology to yeast proteins as follows: class I (HDAC1, HDAC2, HDAC3, and HDAC8) shares homology with Rpd3-like proteins, class II (IIa: HDAC4, HDAC5, HDAC7, and HDAC9; IIb: HDAC6, HDAC10) with Hda1-like proteins, and class III (SIRT1-7) with SIR2-like proteins; class IV (HDAC11) shares homology with classes I and II [61]. The discussion on the regulatory role of CK2 on the HDAC family follows; the discussion of SIRT family regulation is covered later.

CK2 has been shown to phosphorylate HDAC1, HDAC2, HDAC3, and HDAC6 proteins. The functional implications of CK2 phosphorylation of HDAC1-3 have been well documented [19,51–54,64–68] whereas phosphorylation of HDAC6 was only described recently [69]. The CK2-dependent phosphorylation of HDAC1, HDAC2, and HDAC3 is of interest since these form corepressor complexes with ligand-free nuclear receptors and with chromatin remodelers [63,70]. In the early 2000s, mass spectrometry analysis of HDAC1 allowed the identification of Ser421 and Ser423 residues as main phosphorylation sites conforming to the CK2 motif. Furthermore, Ser to Ala substitutions of these residues led to a decrease in the deacetylase activity and impaired complex formation of HDAC with transcriptional regulators [65]. Around the same time, the phosphorylation of the C-terminal domain of HDAC2 by CK2 at Ser394, Ser422, and Ser424 sites was described with Ala substitution showing the same detrimental effects on the enzymatic activity and complex formation [66]. The phosphorylated form of HDAC2 was found to have low abundance in estrogen receptor-positive human breast cancer T5 cells and was bound to chromatin in a complex with HDAC1, and the transcriptional



regulator SP3 [67]. Regulation of HDAC1/2 by CK2 also extends to mitosis where phosphorylation promotes the dissociation of complexes containing both HDAC1 and HDAC2, possibly to allow the formation of mitotic-specific complexes [52]. Moreover, under hypoxia, phosphorylation of HDAC1/2 by CK2 was found to promote HDAC1/2-mediated inactivation of the tumor suppressor gene VHL and subsequent HIF1A protein stabilization in HeLa cells [19]. Since CK2 directly activates HDAC1/2, exploring the effect of CK2 inhibitors, alone or in combination with other drugs, on HDAC1/2 mediated transcriptional repression will help to further reveal the functional context for the association between CK2 association and the HDAC family. In this regard, a recent experiment combining the clinical CK2 inhibitor CX-4945 and the proteasome inhibitor bortezomib showed synergistic effects in B-cell acute lymphoblastic leukemia by inducing the association between the RELA subunit of NF $\kappa$ B and HDAC1 and the subsequent repression of the anti-apoptotic genes BCL-xL and XIAP [71].

Early studies also showed HDAC3 phosphorylation by CK2 at the Ser424 site not found in HDAC1 and HDAC2 [66]. Phosphorylation of HDAC3 by CK2 is also critical for its deacetylase activity and may be counteracted in cells by the action of a complex containing Ser/Thr phosphatase-4 activity [68]. Recently, a mitotic complex containing CK2, HDAC3 phosphorylated at Ser424, and the linker histone H1.3 was identified in HeLa cells suggesting a role of CK2-dependent phosphorylation in chromatin compaction and microtubule stability [72]. CK2 also regulates HDAC3 function by an indirect mechanism, which involves CK2-dependent phosphorylation of PDCD5 at Ser119. PDCD5 is a novel protein found downregulated in cancer patients where overexpression facilitates genotoxic stress-induced apoptosis mediated by TP53 [73]. Phosphorylation of PDCD5 by CK2 enhances its stability and facilitates its nuclear import. In the nucleus, PDCD5 promotes the dissociation of HDAC3 from TP53 in response to genotoxic stress which leads to HDAC3 proteasomal degradation and the stability of TP53 by increased acetylation [73]. Thus, CK2-dependent phosphorylation not only regulates HDAC3 deacetylase activity directly but also its stability.

HDACs have been implicated in cancer both as members of HDAC-containing transcriptional repressor complexes and as deacetylases of non-histone proteins [74,75].

For instance, differential expression patterns of HDACs have been reported in gastric cancer for HDAC1 and HDAC2, and in colon and breast cancer for HDAC3 and HDAC6 [63]. Hence, it is not surprising that HDAC inhibitors are either FDA approved (e.g., SAHA) or being tested in the clinics as chemotherapeutic agents [76]. Given, that CK2 can enhance the activity and/or the stability of these deacetylases combined therapies of HDAC and CK2 inhibitors could be beneficial for cancer treatment. In fact, a recent study described the synthesis and evaluation of dual inhibitors derived from the SAHA and TBB compounds, which are inhibitors of HDAC1 and CK2, respectively. The study showed enhanced pro-apoptotic effect for the dual compounds compared to their parental counterparts [77]. These dual compounds constitute a starting point and further experiments are needed to understand if synergistic/additive effects can be achieved when combining HDAC and CK2 inhibitors.

### 3.4 CK2-dependent signaling and regulation of gene expression

Regulation of gene expression is necessary to respond and adapt to stimuli and stressors in cells. Such regulation involves coordinated events by chromatin-modifying enzymes to open or restrict the accessibility of transcription factors and the transcriptional machinery to the DNA. In the cell, three large HDAC1 and HDAC2 containing transcriptional repressor complexes can be found: SIN3 (SIN3A/B), NuRD (nucleosome remodeling and histone deacetylase), and CoREST (corepressor of RE1-silencing transcription factor) complexes (subunits of each complex listed in Supp. Table 3.4). SIN3 and NuRD share HDAC1, HDAC2, RBBP4, and RBBP7 proteins suggesting the existence of an HDAC complex that matures into these repressor complexes; CoREST and NuRD contain KDM1A as an auxiliary subunit (Figure 3.3E). These complexes regulate the balance of histone deacetylation/acetylation and act as scaffolds for context-specific recruitment of other chromatin-modifying enzymes facilitating chromatin compaction and thus repressing transcription [78]. Recent studies also indicate the role of these complexes in transcriptional activation [55]. This may be a functional dichotomy but may serve a role in transcription priming. A balance between acetylation and deacetylation is essential in priming transcription by maintaining promoters in a more responsive state that can be

switched between repression and activation (see [55] for a comprehensive review). Since CK2-dependent phosphorylation directly enhances the deacetylase activity of these complexes a potential role of CK2 in transcription priming may be considered, particularly shifting the balance towards a repressive state. However, CK2 has also been shown to phosphorylate and/or interact with other core and auxiliary members of these complexes, suggesting a more intricate and context-dependent role for CK2-dependent regulation in maintaining such balance. The multiple functions and auxiliary subunits of these complexes have been discussed elsewhere [79,80] here we focus on presenting evidence that supports a regulatory role by CK2.

The SIN3 complex functions as a scaffold that recruits lysine deacetylases, methylases, demethylases, methylated-DNA binding proteins, and transcription factors to regulate gene expression; a role in DNA repair, transcription elongation stalling, and proteins stability has also been proposed [79]. CK2 has been shown to phosphorylate BRMS1 [81] an auxiliary subunit of SIN3. BRMS1 is a nuclear transcriptional repressor related to SUDS3 adapter protein (a core component of SIN3 complex), that facilitates the recruitment of SIN3 to promoter regions and down-regulates the expression of anti-apoptotic NF $\kappa$ B responsive genes [82]. BRMS1 has been shown to inhibit metastasis in breast cancer and melanoma with the region coding for BRMS1 often lost in late-stage breast cancer [83]. CK2 phosphorylation of BRMS1 at Ser30 residue promotes BRMS1 nuclear export by YWHAE and its ubiquitin-dependent proteasomal degradation in response to TNF treatment. In correspondence, Ser30Ala mutants display a reduced level of metastases in orthotopic xenograft lung cancer models [81]. CK2 also phosphorylates components of the NF $\kappa$ B signaling pathway extending its regulatory role both upstream and downstream of BRMS1 function in metastasis [84].

The NuRD complex is formed by the shared HDAC complex and MTA1/2/3 (chromatin and deacetylase binding), MBD2/3 (methylated-DNA binding), CHD3/4 (DNA helicase, nucleosome remodeling), and GATAD2A/B (protein bridging and DNA binding) proteins [78]. A literature search showed that CK2 could be phosphorylating GATAD2A and GATAD2B in mitotically arrested HeLa cells as decreased phosphorylation at Ser107 and Ser129, respectively, was observed upon treatment the CK2 inhibitor CX-

4945 [85]. Interestingly, the phosphorylation of GATAD2B at Ser129 was somewhat rescued by the expression of an inhibitor tolerant CK2 mutant variant in U2OS cells treated with CX-4945 (<https://ir.lib.uwo.ca/etd/5164>).

As a member of the CoREST complex KDM1A has been linked to breast cancer metastasis, negative regulation of differentiation in neuroblastoma, and acute myeloid leukemia [86]. Overexpression of KDM1A has been described in glioblastoma, bladder, and breast cancer, among others [87]. Recently, KDM1A inhibition in glioblastoma stem cells was shown to induce apoptosis and differentiation [88]. In addition, KDM1A demethylates both TP53 and DNMT1 (DNA cytosine-5-methyltransferase). Demethylation of TP53 affects interaction with its repressor positively regulating apoptosis while demethylation of DNMT1 (DNA methyltransferase) stabilizes the proteins contributing to the maintenance of the DNA methylation pattern [86,89,90]. Phosphorylation of KDM1A by CK2 occurs at several sites (Ser131, Ser137, and Ser166) on the N-terminal region [91] and modulates KDM1A function in DNA damage [90]. Phosphorylation at Ser131 and Ser137 promotes KDM1A interaction with the E3 ubiquitin ligase RNF168 and its recruitment to DNA damage sites where KDM1A can demethylate H3 and RNF168 ubiquitinates H2A. KDM1A phosphorylation also promotes RNF168 interaction with TP53BP1 and its subsequent ubiquitination; TP53BP1 ubiquitination is essential for the non-homologous end joining pathway. Furthermore, phosphorylation of Ser131 and Ser137 has been linked to cell proliferation and survival after treatment with bleomycin or etoposide [90].

CK2 also has the potential to modulate the corepressor complexes NCOR and SMRT [92] (Supp. Table 4). These complexes function by repressing unliganded nuclear receptor signals [93]. The NCOR complex contains NCOR1, GPS2, CORO2A, and TBL1XR1 while SMRT complex contains NCOR2 (SMRT) instead of NCOR1 and lacks CORO2A. Two more general complexes have also been described: the NCOR1 and the NCOR2 complexes, which mediate gene silencing by several transcriptional factors. The NCOR1 complex is formed by HDAC3, NCOR1, SMARCA4, SMARCB1, TRIM28, SF3B3, SF3A1, SRCAP, and SMARCC1/2 [94]. The NCOR2 complex is composed of HDAC1/2/3, NCOR1/2, SAP30, and SIN3A thus sharing proteins with the SIN3 complex

[94]. CK2 phosphorylates both NCOR1 and NCOR2 at Ser2436 and Ser2522 residues, respectively. NCOR1 promotes histone deacetylation and repression of the chromatin and phosphorylation of the C-terminal by CK2 stabilizes NCOR1 from ubiquitin-dependent proteasomal degradation [95]. Phosphorylated NCOR1 binds to the AP1 binding site of the metastasis-related gene CXCL10, which along with phosphorylated HDAC3 and JUN promotes transcriptional repression and prevents the recruitment of coactivators [95]. Another study showed that the balance between NCOR1 complex and the H3K4 methyltransferase KMT2D coactivator complex at Notch target genes changes upon mutation of Ser2436 site of NCOR1 and CK2 inhibition [96]. Similarly, NCOR2 promotes chromatin condensation mediated by nuclear receptors. Phosphorylation *in vitro* using CK2 stabilized to some extent NCOR2 interaction with the thyroid hormone receptor [97]. NCOR2 phosphorylation by CK2 is sensed by the corepressor SPEN (SHARP) [98].

### 3.4.1 CK2 regulatory role in PRC-mediated gene repression

In addition to the complexes described above, HDAC1 and HDAC2 have been linked to the Polycomb repressive complex PRC1.6 [99], however, the functions of PRC1.6 are unknown. In general, PRC1 complexes mono-ubiquitinate H2A providing a marker of transcriptional repression and promoting ubiquitination-independent chromatin compaction [100] (Supp. Table 3.4). The PRC1 complex nomenclature groups a set of complexes that contain RNF2 (also known as RING1B or RING2) and PCGF1-6 as core subunits and several other accessory subunits. Based on the composition of PCGF proteins six groups of PRC1 have been proposed [99]. Interestingly, CK2 has been shown to phosphorylate Ser168 residue of RNF2, the core E3 ubiquitin ligase of PRC1 complexes, inhibiting its enzymatic activity. However, phosphorylation of this residue has only been reported for an AUTS2-PRC1 like complex (PRC1.5) in neuronal diseases. In the context of the PRC1.5 complex (PCGF5, RNF2, CK2 holoenzyme, RYBP, and AUTS2) the protein AUTS2 is needed to maintain the repression of neuronal developmental genes [101]. Phosphorylation of RNF2 by CK2 is believed to be a mechanism that changes the repressive role of this PRC1.5 complex variant into a transcriptional activator [101]. Beyond PRC1.5 and PRC1.6 complexes, CK2 could also

be a regulator of PRC1.4 complex, which contains BMI1 (PCGF4) a CK2 substrate, and the core PCGF protein. BMI1 is heavily phosphorylated in cancer showing context-specific phosphorylation patterns but the role of CK2-dependent phosphorylation in the context of the PRC1.4 complex is unknown. CK2 phosphorylates BMI1 at Ser110, which increases protein stability and may contribute to exacerbate BMI1 role in conferring clonal self-renewal property to ovarian cancer cells [102]. What is more, CK2 phosphorylates CBX2, a core member of PRC1.4 and PRC1.2 complexes, at a poly-Ser region, residues 464–469, next to an acidic region [103]. CBX proteins are necessary for the recruitment of PRC1 to the chromatin by binding to methylated lysine residues; CBX2 preferentially binds H3K27me3 but it also contains a DNA binding motif [99]. Phosphorylation by CK2 increases CBX2 affinity for H3K27me3-containing nucleosomes and decreases its DNA-binding activity. The deletion of either the poly-Ser domain, where phosphorylation by CK2 occurs, or the acidic domain negatively impacts the repression of PRC1 target genes [103]. All the evidence shown here point to a key regulatory role of PRC1-mediated repression by CK2.

### 3.4.2 CK2 regulatory role in DNA methylation

DNA methylation is catalyzed by DNMTs (DNA methyltransferases) and is counteracted by TET (Ten-eleven Translocation) proteins. DNA methylation is directly linked to the modification of histones to maintain proper gene expression after DNA replication. DNA methylation has been associated with a repressive state of chromatin and transcription inhibition [104]. A literature search revealed that CK2 phosphorylates DNMT3A, a *de novo* DNMT that modifies unmethylated CpG dinucleotides. CK2 targets two conserved residues, Ser386 and Ser389, located in the vicinity of the PWWP (Pro-Trp-Trp-Pro) motif and negatively regulates its ability to methylate the DNA [105]. The PWWP domain is required for binding the H3K36me3 mark and thus for the localization of DNMT3A to heterochromatin [104]. CK2-dependent phosphorylation promotes DNMT3A localization to heterochromatin and has an overall impact on the methylation patterns with a most notable decrease in methylation of the SINE (short interspersed nuclear element) CpGs [104]. A study of fetal globin gene silencing described the importance of the H4 methylation mark H4R3me2s generated by the arginine

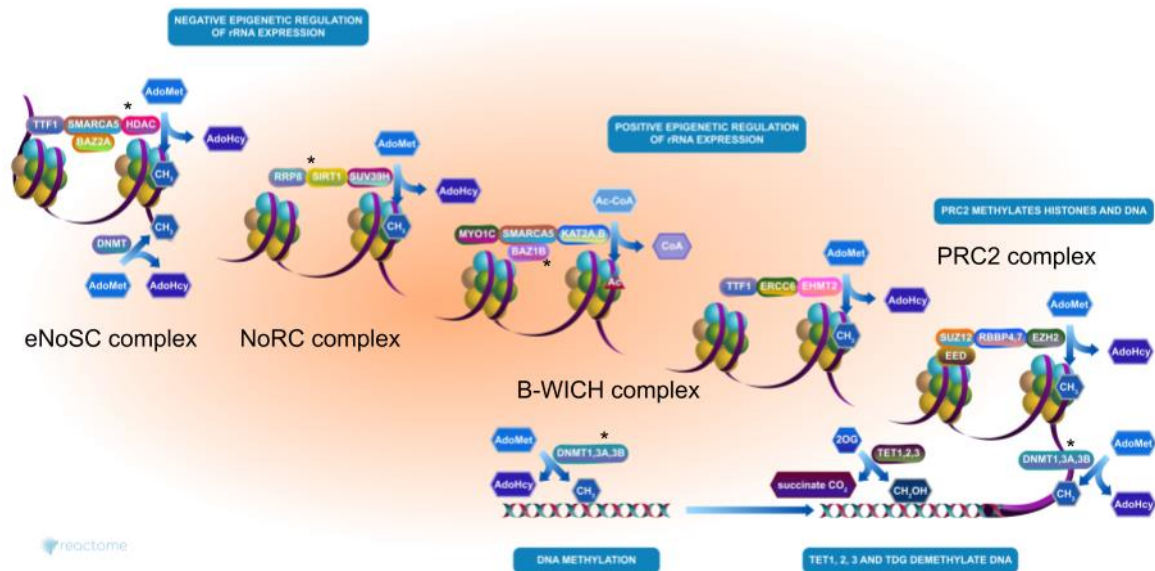
methyltransferase PRMT5 as a factor required for the recruitment of DNMT3A and for the assembly and recruitment of multiprotein repressor complex containing CK2 and HDAC1, among other proteins [106]. The latter suggests a context-dependent role of CK2 in transcription repression as a direct regulator of DNMT3A or a component of HDAC-containing repressor complexes as documented before.

The text mining analysis also indicated that CK2 phosphorylates the epigenetic regulator UHRF1 (ICBP90), potentially at the Ser165 and Ser354 residues [107]. Researchers propose that CK2-dependent phosphorylation regulates UHRF1 transcriptional activity and thus its antiapoptotic properties [107]. Interestingly, UHRF1 is a multidomain protein that promotes the propagation of DNA methylation during DNA replication by recruiting DNMT1 to hemimethylated DNA sites at replication forks [108]. In this regard, the Ser165 and Ser354 sites, the suggested targets of CK2, are in the Tudor-like 1 and the Histone H3R2me0 binding domains, respectively. Both these regions of the protein are involved in binding of H3R2me0 (un-methylated) thus CK2-dependent phosphorylation could be essential for regulating the epigenetic role of UHRF1. What is more, UHRF1 is overexpressed in several cancers and its downregulation blocks the cell cycle in G1 [109]. In correspondence, additional information could expand on CK2 as a regulator of UHRF1 function in normal cells and cancer and provide further evidence on the regulatory role of CK2 on the DNA methylation landscape.

### 3.4.3 CK2 regulatory role in ribosomal RNA expression

The expression of ribosomal RNA (rRNA) is epigenetically regulated by the eNoSC (energy-dependent nucleolar silencing complex) and the NoRC (nucleolar remodeling complex) complexes (Fig. 3.4, Supp. Table 3.4). The lysine deacetylase SIRT1 is a CK2 substrate and functions as a metabolic sensor in the cell. As a core component of the eNoSC complex, along with HDAC1 and HDAC2, also known CK2 substrates, and the DNA methyltransferases DNMT1 and DNMT3B, it promotes a repressive state of the chromatin at the rRNA genes [110–112]. In this context SIRT1 mediates H3 deacetylation, which can then be subsequently modified to H3K9me2 by the methyltransferase SUV39H1, thus inhibiting the transcription of rRNA [113] (Fig. 3.4). The concerted action of SIRT1 and SUV39H1 is also important for the formation of

facultative heterochromatin and SIRT1 has been shown to regulate the activity of SUV39H1 by acetylation [114]. SIRT1 also deacetylates the subunit TFA1B of SL1 complex (RNA polymerase I core factor complex), which recruits polymerase I to the ribosomal DNA (rDNA) promoters thus inhibiting transcription of the rRNA genes [115].



**Figure 3.4.** Epigenetic regulation of gene expression in *Homo sapiens* (Reactome identifier: R-HSA-212165). \* Examples of CK2 substrates in the pathway.

CK2 phosphorylates SIRT1 at Ser162 (corresponding to S154 in mice), Ser659, Ser661, and S683 residues [56,57] but the regulatory role of these phosphorylation events is not well understood. In mice, phosphorylation of SIRT1 by CK2 was shown to increase the deacetylation rate and to protect cells against DNA damage-induced apoptosis by promoting the deacetylation and binding of TP53 [57]. Since CK2-dependent phosphorylation was found to increase SIRT1 deacetylase activity it could be of interest to understand the impact of phosphorylation in the context of rRNA expression. The NoRC complex binds to HDAC1 and HDAC2 to facilitate the deacetylation of H4 which allows for DNA methylation and silencing [116]. In both cases, CK2 could be modulating the deacetylase activity at the rRNA genes thus enhancing rDNA repression mediated by both these pathways. CK2 also phosphorylates TAF1C (TAFI110) a component of the SL1 complex and UBTF, which recognizes the rRNA gene promoter and activates transcription. CK2 phosphorylation of TAF1C prevents SL1 from binding



to rDNA impeding the formation of the preinitiation complex (PIC) [117,118]. Conversely, phosphorylation of the UBTF C-terminal stabilizes and activates the protein promoting transcription of rDNA [117,118].

Besides its potential role in the negative regulation of rRNA expression, CK2 has also been shown to phosphorylate protein DEK [119]. DEK is a component of the B-WICH complex, a multi-subunit complex that recruits histone acetyltransferases positively regulating the expression of rRNA [120,121] (Fig. 3.4). This protein is a pleiotropic oncoprotein [122] that was first identified as the fusion protein DEK-CAN (NUP214) in AML (acute myeloid leukemia) [123] and is the target of autoantibodies in several diseases including lupus and arthritis [119]. The function of DEK is not well understood but CK2 has been identified as the main kinase that phosphorylates it. Phosphorylation by CK2 occurs at Ser32 and at the C-terminal residues Ser243, Ser244, Ser251, Ser301, Ser303, Ser306, and Ser307 [119]. The knockdown of DEK leads to genomic instability due to a global loss of chromatin integrity manifested in the loss of heterochromatin (increase in histone acetylation and a decrease in histone methylation) [124]. DEK is regarded as an architectural protein and phosphorylation of CK2 regulates the binding to DNA. Thus, through modulation of DEK binding affinity CK2 could be an important regulator of DEK functions in rRNA expression but also in the maintenance of a global chromatin structure [119].

The B-WICH complex is an extension of the WICH complex, a member of the ISWI (Imitation SWItch) family of chromatin remodeling complexes composed of seven different complexes containing either the helicase SMARCA5 or SMARCA1 [125]. Specifically, the WICH complex is composed of SMARCA5 and BAZ1B (WSTF) proteins [126]. During replication, BAZ1B interacts with PCNA promoting an open chromatin structure along with SMARCA5 [127]. Interestingly, BAZ1B, an atypical tyrosine-protein kinase and a core member of WICH, that recruits the complex to the site of DNA replication was described to be a substrate of CK2 at Thr945 and Ser947 residues in a phosphoproteomic study of mitotically arrested HeLa cells treated with CX-4945 [85]. The inhibition by CX-4945 of BAZ1B at the Ser1468 residue which conforms to the CK2 motif was also observed when expressing an exogenous inhibitor tolerant

CSNK2A1 mutant variant (<https://ir.lib.uwo.ca/etd/5164>) in U2OS cells. The phosphorylation of Ser1468 by CK2 was also confirmed in an *in vitro* dephosphorylation-phosphorylation assay followed by phosphoproteomic analysis and peptide synthesis and *in vitro* CK2 phosphorylation assay [128]. The tentative CK2 target sites on BAZ1B are not located in critical domains, however, without further studies is impossible to define the relevance of CK2-dependent phosphorylation of BAZ1B and the subsequent impact on the function of the WICH complex in chromatin remodeling.

#### 3.4.4 CK2 regulation of transcription factors

CK2 phosphorylates the transcriptional corepressor DAXX [129] which interacts with the CK2 substrates: DEK, HDAC1, and HDAC2. As a transcriptional corepressor DAXX binds sumoylated (small ubiquitin-like modification) transcription factors (e.g., AR, SMAD4, and CBP) through its SUMO-interacting motif (SIM) negatively regulating basal and activated transcription [130]. The presence of a SIM motif on DAXX also mediates its sumoylation [130,131]. Interestingly, CK2 phosphorylates DAXX at Ser737 and Ser739 within the SIM motif ( $I^{733}IVLSDSD^{740}$ ) and changes its preference to SUMO1 compared with SUMO2 and SUMO3 [131]. Evidence indicates that phosphorylation of the SIM motif of DAXX sensitizes cells to apoptosis by repressing the expression of anti-apoptotic genes [131]. DAXX function is also regulated by its binding to PML nuclear bodies through interaction with MCSR1 and the SUMO-modified protein PML [132]. Of importance, CK2 also phosphorylates PML at three residues, Ser560-Ser562, localized to the SIM ( $V^{556}VVISSS^{562}$ ) domain, which is necessary for recruitment of sumoylated transcription factors to the PML-bodies [133]. CK2-dependent phosphorylation of PML at Ser517 residue has also been described and regulates the levels of PML by promoting its ubiquitin-dependent degradation [134]. Overall, experimental data points to CK2 as an important regulator of DAXX co-repressor activity not only by directly changing its affinity towards SUMOs but also by modulating PML levels. An overlap between CK2 phosphorylation and SIM motifs has also been described for the transcriptional regulator PIAS1 [135].

IKZF1 is a lymphoid lineage transcriptional regulator that acts as a repressor in the context of SIN3-HDAC, NuRD, and CtBP (C-terminal binding protein) complexes.

IKZF1 is also an activator when bound to the SWI/SNF complex (P/R complex) [136–138]. The presence of dominant-negative IKZF1 isoforms (while retaining a wild-type copy) has been linked to the development of hematological malignancies (e.g., ALL, AML, and myelodysplastic syndrome) with evidence supporting a tumor suppressor role of the full-length protein [139]. CK2 phosphorylates IKZF1 at Ser13 and Ser294 affecting its binding to the DNA and its localization to pericentromeric heterochromatin; the latter results in a diffuse pattern of the protein across the nucleus [140,141]. Phosphorylation by CK2 inhibits IKZF1-HDAC1 complex-mediated repression of the histone H3K4 demethylase KDM5B (JARID1B), resulting in KDM5B overexpression and in turn the expression of leukemia-associated genes [20]. CK2 is overexpressed in leukemia and CK2 inhibition promotes the transcriptional repressor activity of the wild-type IKZF1 allele by inhibiting the expression of genes involved in cell cycle and the PI3K pathway [139,142]. In addition, CK2 inhibition promotes IKZF1-dependent expression of the lysine demethylase PHF2 a tumor suppressor whose expression is low in ALL [143]. The data presented provide a rationale for CK2 inhibition in leukemia dependent on changes in the chromatin modification status.

CK2 also phosphorylates the transcriptional regulator MAX at Ser2 and Ser11 residues [144]. MAX is a component of the MXD1-MAX (MAD-MAX) heterodimer that associates with a complex containing SIN3A and HDAC1/2 and mediates repression of MYC-MAX target genes during differentiation and enhancing cell growth inhibition by MXD1 [145,146]. Phosphorylation by CK2 results in enhanced DNA binding of both MAX-MAX homodimer and MYC-MAX heterodimers [144]; thus, antagonizing MXD1-MAX heterodimer function and inducing expression of MYC-target genes. Interestingly, the Ser11 site is adjacent to a CASP5 cleavage site at E10, and phosphorylation of this residue has been described to block CASP5-mediated cleavage [147]. Furthermore, CK2 phosphorylates MYC at several phosphosites (Ser249, Ser250, Ser252, Ser347, Ser348) which are thought to promote MYC stability and regulate the DNA binding properties of MYC-MAX dimers [148–150].

### 3.5 Chromatin dynamics intersect with the CK2 protein-protein interaction network

Three major groups of proteins are involved in chromatin organization, those that methylate DNA, those that modify histones, and those that regulate DNA-histone interactions in an ATP-dependent manner [7]. Several of these proteins have been mentioned here as bonafide CK2 substrates but many others have been identified as CK2 interactors with no phosphorylation link provided to the kinase. To further explore the CK2 regulatory role, I retrieved the interactors of the CK2 subunits using the Bioconductor package resource PSICQUIC v1.24.0 [151] without filtering by detection method. Next, I filtered the interactors to find the chromatin organization proteins represented. As a result, 57, 23, and 22 unique hits for CSNK2A1, CSNK2A2, and CSNK2B, were found respectively (Suppl. Table 3.5). The merged interactors for all subunits represented 69 proteins of the chromatin organization proteins included in the data sets. This analysis is a snapshot and can change with new protein-protein interaction data added to the PSICQUIC participating databases. As a summary, Table 3.2 shows the GO cellular compartments functional classification of all the CSNK2A1 interactors (Suppl. Table 3.6) retrieved from PSICQUIC filtered by complexes relevant to chromatin organization.

**Table 3.2.** CSNK2A1 subunit interactors members of chromatin organization complexes (GO cellular compartment, level 5).

Description	Count	Genes
transcription factor complex (GO:0005667)	38	SIN3A, ATF1, FOS, JUN, DDIT3, TAF1, TAF1D, SOX2, WWOX, TP53, KDM1A, MED23, HDAC1, PRKDC, BRF1, RELA, SPI1, GTF2F1, HCLS1, SFPO, HMGA1, SUB1, HSP90AB1, AIP, XRCC6, NFYA, ASCC1, LHX3, HIF1A, HES6, GTF2A1, TCF7L2, MAX, MYF5, TAF7, PROP1, NKX2-5, REL
PcG protein complex (GO:0031519)	19	HDAC2, SIRT1, RYBP, CBX2, CBX4, CBX7, CBX8, RNF2, PCGF5, PCGF3, PCGF6, PHC1, RBBP4, RING1, PHF19, BMI1
methyltransferase complex (GO:0034708)	18	HDAC2, TAF1, SIRT1, PAXIP1, RNF2, RBBP4, CLNS1A, ZC3H13, KMT2A, WDR5, PRMT1, PHF19, MCRS1, RBBP5, DPY30, KMT2B, MAX, TAF7
histone methyltransferase complex (GO:0035097)	15	HDAC2, TAF1, SIRT1, PAXIP1, RNF2, RBBP4, KMT2A, WDR5, PHF19, MCRS1, RBBP5, DPY30, KMT2B, MAX, TAF7

nuclear DNA-directed RNA polymerase complex (GO:0055029)	13	TAF1, TP53, POLR2E, GTF2F1, CTR9, RTF1, PAF1, LEO1, RPRD2, MCM3, POLR2A, GTF2A1, TAF7
SWI/SNF superfamily-type complex (GO:0070603)	10	HDAC2, CHD3, HDAC1, DPF3, RBBP4, CHD4, SMARCC2, RSF1, MCRS1
Sin3-type complex (GO:0070822)	7	HDAC2, SIN3A, NCOR1, HDAC1, RBBP4, SUDS3
protein acetyltransferase complex (GO:0031248)	6	CREBBP, OGT, WDR5, MCRS1, TAF7, JADE2
NuRD complex (GO:0016581)	6	HDAC2, CHD3, HDAC1, RBBP4, CHD4
nBAF complex (GO:0071565)	2	DPF3, SMARCC2
DNA topoisomerase complex (ATP-hydrolyzing) (GO:0009330)	2	TOP2A, TOP1
protein N-acetylglucosaminyltransferase complex (GO:0017122)	1	OGT
DNA helicase complex (GO:0033202)	1	MCRS1
npBAF complex (GO:0071564)	1	SMARCC2
Ino80 complex (GO:0031011)	1	MCRS1
chromatin silencing complex (GO:0005677)	1	SIRT1

Furthermore, I identified several chromatin-modifying enzymes as CK2 interactors of the CSNK2A1 subunit, i.e., the lysine-specific demethylase KDM5C, the lysine methyltransferases KMT2A and KMT2B, the arginine N-methyltransferase PRMT1, and the histone deacetylase HDAC5. Other CK2 interactors found linked to chromatin organization are CHD3, a helicase incorporated to the NuRD complex [152]; SMARCAD1, an SWI/SNF-like protein that mediates ATP-dependent nucleosome-remodeling [153,154]; JADE2, an E3 ubiquitin-protein ligase component of the histone acetyltransferase complex HBO1 involved in KDM1A degradation during neurogenesis [155,156].

Regarding repressor complexes, CSNK2A1 directly interacts with SUDS3 and CHD4 proteins, which are members of the SIN3 and NuRD complexes, respectively. SUDS3 is a regulatory protein known to enhance HDAC1 activity in the context of the SIN3 complex [157]. CHD4 provides ATP-dependent DNA helicase activity to the NuRD complex and functions by combining ATP-dependent nucleosome remodeling activity with “reader” domains that bind histone tails [158]. Interestingly, in the lymphoid system CHD4-NuRD complex interact with IKZF1 (IKAROS), a CK2 substrate, which provides

sequence recognition specificity resulting in nucleosome mobilization, histone deacetylation, and promoter decommissioning [158]. This highlights once again the relevance of CK2-dependent signaling in transcriptional regulation of the lymphoid system.

Coactivator complexes were also represented with two members of the SWI/SNF superfamily-type complex: SMARCC2 and RSF1 identified as CSNK2A1 interactors [93]. The SWI/SNF chromatin remodeling complexes change DNA-histone interaction to allow gene expression. SMARCC2 (BAF170) functions as a scaffolding protein for the SWI/SN family with domains important in chromatin targeting and DNA specificity [159]. RSF1 (HBXAP) is a histone chaperone for the ISWI complex family that along with SMARCA5 forms the RSF complex (Remodeling and spacing factor). RSF1 is frequently overexpressed in cancer, which impacts genome stability by inducing DNA strand breaks [125]. Another CSNK2A1 interactor involved in chromatin activation is the lysine acetyltransferase CREBBP (CBP).

### 3.6 CK2-dependent phosphoproteome and chromatin dynamics

Although several functional associations have been exposed here, the phosphorylation of HDAC1-3 by CK2 remains the best characterized in the context of chromatin organization. To further explore the involvement of CK2-dependent phosphorylation we performed a motif scan of the proteins in the merged datasets followed by search in two orthogonal phosphoproteomic datasets describing a global *in vitro* screening of CK2 substrates [128] and the treatment of mitotic HeLa cells with the CK2 inhibitor CX-4945 [85], respectively. As a result, 21 proteins involved in chromatin organization that contains at least one CK2 consensus sequence hit were found to be phosphorylated *in vitro* by CK2 and/or its phosphorylation was inhibited by CX-4945 treatment (Table 3.3). The functional classification pointed to a tentative regulatory role of CK2 in histone (and other proteins) methylation and demethylation, which reinforces the notion that the CK2 roles in chromatin organization extend beyond modulation of HDAC activity. It also expands the CK2 role in transcriptional repression by exposing candidate CK2 targets including GATAD2A, GATAD2B, and CHD4, core components of the NuRD complex;

SUDS3 and ARID4B, components of SIN3A complex (Table 3.3). It also ratifies potential CK2 roles in chromatin remodeling by revealing tentative targets like SMARCA4, a core component of SWI/SNF chromatin remodeling complexes; and CHD1, a component of the histone acetylation complex SAGA (Table 3.3). In fact, a recent study suggests a regulatory role of CK2-dependent phosphorylation in the subcellular compartmentalization of SMARCA4 as the hyperphosphorylated protein was found to localize to the soluble chromatin [160]. It also pointed to the possibility of interactors of CK2 like RSF1 and CDH4 as targets of CK2-dependent phosphorylation (Table 3.3).

**Table 3.3.** Proteins involved in chromatin organization that contains at least one hit for CK2 consensus sequence with changed phosphorylation when CK2 is manipulated.

Protein	Function	CK2 motif (-7/+7)	CK2 manipulation
<b>Histone-modifying enzymes</b>			
EZH2	Histone methyltransferase	Ser87 (TRECSVTSDLDFPTQ) Thr378 (NVLESKDTDSREAG)	Thr487 (APAEDVDTPPRKKK)***
NSD1	Histone methyltransferase	More than 10 hits	Ser483 and Ser486 (CLKSLAFDSEHSADEK)**
SETD2	Histone methyltransferase	More than 10 hits	Ser323 (GIFLGSESEDESVRT)**
KDM2A	Histone demethylase	Ser28 (RYEDDGISDDEIEGK) Ser692 (QKRKMEESDEEAVQA) Ser787 (PSGKKELSEVEKAKI)	Ser28**
<b>Transcriptional modulator</b>			
AEBP2	Repressor	Ser211 (RGSLEMSSDGEPLSR)	Ser515 (KVFNLYLSKQ____)**
ATF7IP	Activator/Repressor	Ser445 (ILEKLAPSEDELTCF) Ser496 (KEAFLVLSDEEDISG)	Thr677 (PPVSPGKTVNDVNSN)**
GATAD2A	Repressor	Thr20 (RALERDPTEDDVESK) Thr42 (LLASDLNTDGD MRVT) Ser107 (SPDVIVLSDNEQPSS)	Ser107**
GATAD2B	Repressor	Ser112 (VDMSARRSEPERGRL) Ser129 (SPDIIVLSDNEASSP)	Ser129**
SAP30BP	Corepressor	Ser18 (LAVYAEDSEPESDG) Ser22 (AEDSEPESDGEAGIE) Ser72 (EDENSRQSEDDDDSET) Ser77 (RQSEDDDDSETEKPEA) Thr91 (ADDPKDNTEAEKRDP)	Ser43 (EEKGGLVSDAYGEDD)**
CBX3	Repressor	Thr55 (FLKWKGF TDADNTWE) Ser95 (GTKRKSLSDESDDDS) Ser97 (KRKSLSDESDDSKS)	Ser95 and Ser97*
SMARCA4 (BRG1)	Activator/Repressor	Ser1627 (SRAKPVVSDDDSEEE) Ser1631 (PVSDDDDSEEEQEED)	Ser1627 and Ser1631*
ARID4B	Repressor	More than 10 hits	Thr793 (IEVLS EDTDYEEDEV)*
SUDS3	Repressor	Ser45 (RSCRGRESD EDEDTEA) Ser53 (DEDEDASETDLAKH)	Ser45*
<b>Chromatin remodeling</b>			
BAZ1A	Accessory unit. Nucleosome spacing	Ser164 (DGETIIISDSDDDSET) Thr498 (EVAKEQLTDADTKDL) Thr731 (ELDQDMVTEDEDDPG) Ser1214 (RQRPSLESD EDDVEDS)	Thr731**

		Ser1238 (GDEEEGQSEEEYEV)	
BPTF	Nucleosome sliding	More than 10 hits	Thr300 (VVLREEDTSNTTFGP)* Thr301 (VLREEDTSNTTFGPA)* Ser303 (REEDTSNTTFGPADL)* Thr1763 (TTVTKLSTPSTGGSV)**
ATRX	Helicase	More than 10 hits	Ser677 (RMSHSSSDTDINEI)* Ser1156 (IQSGSSSDAEESSE)* Ser850 (NTKDFDSEDEKHSK)* Ser1061 (SKKKDELSDYAEKST)**
CHD1	Helicase	More than 10 hits	Ser216 (KKRQIDSSEEDDDEE)*
CHD4	Helicase	More than 5 hits	Ser103 (EEEVALRSDSEGSY)*
RSF1	Histone chaperone	More than 10 hits	Thr1305 (KRLHRIETDEEESCD)* Thr1332 (SQPRVLPSEQESTKK)* Ser1345 (KKPYRIESDEEEDFE)*
SUPT6H	Transcription elongation factor	More than 10 hits	Ser78 (EGSDSGDSEDDVGHK)*
<b>DNA methylation</b>			
DNMT3B	DNA methyltransferase	Thr803 (KEDVLWCTELERIFG)	Thr136 (GRNHVDESPVEFPAT)**

\*phosphorylated by CK2 *in vitro* (Bian *et al.* 2013 [128])

\*\*responsive to CX-4945 treatment (Rusin *et al.* 2017 [85])

After gathering information on CK2 modulation and CK2 interactors we looked beyond repressive complexes and HDACs focusing on ATP-dependent chromatin-remodeling complexes. These complexes contain an “ATP motor” that breaks the DNA-protein contacts to allow nucleosome assembly (ISWI and CHD subfamily), access (SWI/SNF subfamily), or editing (INO80 subfamily, SRCAP) through different mechanisms of action [161]. The diverse composition of these complexes determines their interaction with transcription activators, repressors, and histone modifications, and thus chromatin targeting [161]. Interestingly, regulatory links between CK2 and a number of core and auxiliary components of the SNF/SWI [162], ISWI [125], CHD [163,164], and INO80 [165,166] remodeling complexes (Supp. Fig. 2) were identified. Notice that some of these complexes can contain enzymatic activities that function in gene repression [167,168]. For instance, the CHD complex NuRD contains the histone deacetylases HDAC1 and HDAC2, both CK2 substrates as discussed before. The NuRD “ATP motor” and deacetylase activities promote heterochromatin formation and rapid termination of gene expression [168]. Given that, CK2 phosphorylation enhances HDAC1 and HDAC2 (and HDAC3) deacetylase activity and regulates their binding to repressor complexes, CK2 stands at the junction of this chromatin remodeling-deacetylase link.



### 3.6.1 Asp-N phosphoproteome

Trypsin is widely used for phosphoproteome generation. To uncover other possible sites modulated by CK2 manipulation we performed a SILAC phosphoproteomic study in U2OS treated for 6h with 30  $\mu$ M CX-4945 using Asp-N as the protein digestion enzyme [169] (three technical replicates) (Supp. Table 3.7). Interestingly we identified changes in several chromatin organization proteins. Importantly, we observed a potential regulatory role of CK2 in the regulation of the PRC2 complex through the modulation of phosphorylation of EED and EZH2 core proteins (Supp. Table 3.7). This is in addition to the regulatory role described for PRC1. Thus, pointing to a global role of CK2 in regulating the function of the Polycomb group (PcG) proteins, chromatin-associated proteins that assemble into diverse complexes that maintain transcriptional repression [170,171]. The modulation of KAT7 phosphorylation was also observed (Supp. Table 3.7).

PRC2 complex mediates H3K27me1, H3K27me2, and H3K27me3, through the action of the methyltransferases EZH2 and EZH1; H3K27me3, is a repressive histone mark [170,171]. The core subunits of PRC2 are the adaptor protein SUZ12, EED, and the lysine methyltransferase EZH2. Both SUZ12 and EED are required for binding to chromatin and for catalytic activity [170]. The residues Ser378 and Ser380 of EZH2, the catalytic subunit of PRC2, were found inhibited  $\sim$  -3 fold. These sites localize to the region 329–522 where interaction with the chromatin reader CDYL has been mapped [172]. Importantly, phosphorylation at Thr350 by CDK1, a residue in this region, modulates the recruitment of PRC2 enhancing its binding to ncRNA, which directs the complex to defined target genes [173,174]. Increased phosphorylation of Ser43 and Ser46 residues of EED at  $\sim$  1.54 fold was also observed. EED is required for the catalytic activity by allosteric activation of the PRC2 complex via binding of H3K27me3 and potentially JARID2 K116me3 [170]. The N-terminal region of EED is required for interaction with H3 [175], thus increased phosphorylation of this region could directly impact EED function in PRC2.

The change of several phospho-proteins directly or indirectly involved in histone modification was recorded (UniProt source), such as HUWE1, histone deubiquitination;

ISWI, part of the SUPT6H-IWS1-CTD complex with recruits histone methyltransferases and function in elongation-coupled H3K36 methylation; KAT7, H4-specific acetyltransferase activity, which also mediates H3K14 acetylation; LEO1, a member of the PAF1 complex which mediates the formation of H2BK120ub1 and H3K4me3; MPHOSPH8, a member of the HUSH complex that recognizes the H3K9me3 and mediates transcriptional silencing by promoting H3K9me3 deposition through recruiting of the histone methyltransferase SETDB1 also promotes DNA methylation through DNMT3A; SART3, recruitment of USP15 for histone deubiquitination.

### 3.7 Methods and Materials

**Data sets.** Chromatin organization data sets were retrieved from Reactome [26] (R-HSA-3247509, R-HSA-212165) and AmiGO [25] (GO:0006338) portals to perform functional association to CK2.

**Protein function.** Protein function was retrieved from the UniProt database [176] to determine the role of proteins of interest in chromatin remodeling and CK-dependent signaling.

**Literature search.** Automated literature retrieval and extraction were performed with Chilobot text mining tool [24] and the reutils Bioconductor R package [23] to discover known associations between proteins of interest and CK2.

**Database search.** Automated database searches were performed using the Bioconductor R packages UniProt.ws v2.26.0 (UniProt) [177], curl v4.3 (UniProt) [178], iptmnetr v0.1.8 (PTM and kinase-substrate relationships) [179], and PSICQUIC v1.24.0 (molecular interaction) [151]. CORUM (protein complexes) database [180] search was performed on data downloaded in August 2019. This allowed for retrieval of the protein sequence, modification, kinase-substrate, and protein interaction data used in the experiments described in the chapter.

**Phosphoproteomic data.** Phosphoproteomic data for CK2 manipulation was obtained from: **1)** Bian *et al.* 2013 ([128], all sites previously reported in databases); **2)** Rusin *et al.* 2017 [85] supplementary materials; **3)** SILAC phosphoproteomic study by Edward

Cruise *et al.* (<https://ir.lib.uwo.ca/etd/5164>). Light: U2OS cells over-expressing CSNK2A1-WT treated for 4 h with DMSO. Medium: U2OS cells over-expressing CSNK2A1-WT treated for 4 h with 30  $\mu$ M CX-4945. Heavy: U2OS cells over-expressing CSNK2A1-TM (V66A/H160D/I174A) treated for 4 h with 30  $\mu$ M CX-4945. Site localization was selected Localization prob  $\geq 0.75$ . Differentially modulated sites were selected from with a log<sub>2</sub> Medium/Light fold change cut-off  $\leq -1$ . **4)** 30  $\mu$ M CX-4945 treatment of U2OS cells for 6 h followed by sample preparation and protein digestion with the endoproteinase Asp-N [169] (see Chapter 2).

**Protein complex representation.** Protein complex data retrieved from CORUM was overlapped with CK2-specific information and plotted using the RCy3 Bioconductor R package [181]. Nodes style scheme: yellow circle (protein), white square (complex), dark red (CK2 substrate), medium red (Bian *et al.* 2013, fold change  $\geq 2$ ), light red (Bian *et al.* 2013), blue (Rusin *et al.* 2017), green (PPI). This analysis was performed to find core complexes proteins functionally linked to CK2 such as interactors and substrates.

### 3.8 Concluding remarks

The genomic DNA, histone proteins, and other associated proteins assemble into the chromatin structure formed by the nucleosome as its basic unit. The composition and conformation of the chromatin structure, i.e., chromatin organization, is decided by events like the modification of histones and DNA and by transcription, which dictate its specification, formation, and/or maintenance. The information relating to the regulatory role of protein kinase CK2 in chromatin organization remains sparse. As shown here, there is little information describing the direct relevance of CK2-dependent phosphorylation events in this context. However, considering the wide distribution in the proteome of the CK2 consensus motif, several target sites are likely yet to be discovered among proteins involved in chromatin organization. In addition, evidence obtained in Chapter 2 indicates a functional association between CK2 and lysine acetylation networks in the cell involving chromatin remodeling proteins among other activities.

Briefly, here we discussed the role of CK2 as a histone H4 and histone H3 kinase impacting DNA repair and the function of the SAGA transcriptional co-activator,

respectively. We also summarized lesser understood roles of CK2 as a modulator of heterochromatin methylation patterns, chromatin-remodeling, and transcriptional repression. Regarding the latter, we gathered novel evidence supporting a regulatory role of CK2 on the function of the repressor complex PRC2 as treatment with CX-4945 inhibited the phosphorylation of its core subunits EZH2 and EED at sites matching the CK2 consensus. Finally, we also observed that inhibition of CK2-dependent signaling could result in the modulation of the phosphorylation of the lysine acetyltransferase KAT7. Overall, these novel observations require further validation using other chemical probes targeting CK2, but they expand the breadth of potential regulatory roles to consider when studying CK2.

All in all, our analysis highlights the importance of conducting/mining both low and high-throughput experiments for finding and validating new connections between associated biological processes and CK2-dependent signaling. Of special interest should be the phosphoproteomic profiling of existing and novel CK2 inhibitors. But also, the necessity of novel studies profiling the changes in the phosphoproteome resulting from CK2 protein level manipulation either by over-expressing or knocking-down CK2 subunits. Regardless, here we were able to summarize and identify essential aspects of the regulatory role of CK2 in chromatin organization.

### 3.9 References

1. Giancotti, F.G. Deregulation of cell signaling in cancer. *FEBS Lett.* **2014**, *588*, 2558–2570, doi:10.1016/j.febslet.2014.02.005.
2. Liu, F.; Wang, L.; Perna, F.; Nimer, S.D. Beyond transcription factors: how oncogenic signalling reshapes the epigenetic landscape. *Nat. Rev. Cancer* **2016**, *16*, 359–372, doi:10.1038/nrc.2016.41.
3. Mast, F.D.; Ratushny, A. V; Aitchison, J.D. Systems cell biology. *J. Cell Biol.* **2014**, *206*, 695–706, doi:10.1083/jcb.201405027.
4. Roy, D.M.; Walsh, L.A.; Chan, T.A. Driver mutations of cancer epigenomes. *Protein Cell* **2014**, *5*, 265–296, doi:10.1007/s13238-014-0031-6.
5. Sawan, C.; Herceg, Z. Histone modifications and cancer. *Adv. Genet.* **2010**, *70*, 57–85, doi:10.1016/B978-0-12-380866-0.60003-4.
6. Kouzarides, T. Chromatin modifications and their function. *Cell* **2007**, *128*, 693–705, doi:10.1016/j.cell.2007.02.005.
7. Finley, A.; Copeland, R.A. Small molecule control of chromatin remodeling. *Chem. Biol.* **2014**, *21*, 1196–1210, doi:10.1016/j.chembiol.2014.07.024.
8. Bannister, A.J.; Kouzarides, T. Regulation of chromatin by histone modifications.

- Cell Res.* 2011, *21*, 381–395.
9. Musselman, C.A.; Lalonde, M.-E.; Côté, J.; Kutateladze, T.G. Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1218–27, doi:10.1038/nsmb.2436.
  10. Badeaux, A.I.; Shi, Y. Emerging roles for chromatin as a signal integration and storage platform. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 211–224, doi:10.1038/nrm3545.
  11. Kohli, R.M.; Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **2013**, *502*, 472–479, doi:10.1038/nature12750.
  12. Breiling, A.; Lyko, F. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin* **2015**, *8*, 24, doi:10.1186/s13072-015-0016-6.
  13. Guo, C.; Davis, A.T.; Ahmed, K. Dynamics of protein kinase CK2 association with nucleosomes in relation to transcriptional activity. *J. Biol. Chem.* **1998**, *273*, 13675–13680, doi:10.1074/jbc.273.22.13675.
  14. Barz, T.; Ackermann, K.; Dubois, G.; Eils, R.; Pyerin, W. Genome-wide expression screens indicate a global role for protein kinase CK2 in chromatin remodeling. *J. Cell Sci.* **2003**, *116*, 1563–1577, doi:10.1242/jcs.00352.
  15. Salvi, M.; Sarno, S.; Cesaro, L.; Nakamura, H.; Pinna, L.A. Extraordinary pleiotropy of protein kinase CK2 revealed by weblogo phosphoproteome analysis. *Biochim. Biophys. Acta* **2009**, *1793*, 847–59, doi:10.1016/j.bbamcr.2009.01.013.
  16. Vilk, G.; Weber, J.E.; Turowec, J.P.; Duncan, J.S.; Wu, C.; Derksen, D.R.; Zien, P.; Sarno, S.; Donella-Deana, A.; Lajoie, G.; et al. Protein kinase CK2 catalyzes tyrosine phosphorylation in mammalian cells. *Cell. Signal.* **2008**, *20*, 1942–1951, doi:10.1016/j.cellsig.2008.07.002.
  17. Niefind, K.; Guerra, B.; Ermakowa, I.; Issinger, O.G. Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme. *EMBO J.* **2001**, *20*, 5320–31, doi:10.1093/emboj/20.19.5320.
  18. Trembley, J.H.; Wang, G.; Unger, G.; Slaton, J.; Ahmed, K. Protein Kinase CK2 in Health and Disease. *Cell. Mol. Life Sci.* **2009**, *66*, 1858–1867, doi:10.1007/s00018-009-9154-y.
  19. Pluemsampant, S.; Safronova, O.S.; Nakahama, K.; Morita, I. Protein kinase CK2 is a key activator of histone deacetylase in hypoxia-associated tumors. *Int. J. cancer* **2008**, *122*, 333–341, doi:10.1002/ijc.23094.
  20. Wang, H.; Song, C.; Ding, Y.; Pan, X.; Ge, Z.; Tan, B.H.; Gowda, C.; Sachdev, M.; Muthusami, S.; Ouyang, H.; et al. Transcriptional regulation of JARID1B/KDM5B histone demethylase by ikaros, histone deacetylase 1 (HDAC1), and casein kinase 2 (CK2) in B-cell acute lymphoblastic leukemia. *J. Biol. Chem.* **2016**, *291*, 4004–4018, doi:10.1074/jbc.M115.679332.
  21. Kayhanian, H.; Smyth, E.C.; Braconi, C. Emerging molecular targets and therapy for cholangiocarcinoma. *World J. Gastrointest. Oncol.* **2017**, *9*, 268–280, doi:10.4251/wjgo.v9.i7.268.
  22. Cole, P.A. Chemical probes for histone-modifying enzymes. *Nat. Chem. Biol.* **2008**, *4*, 590–597, doi:10.1038/nchembio.111.
  23. Schöfl, G. reutils: Talk to the NCBI EUtils. 2016.
  24. Chen, H.; Sharp, B.M. Content-rich biological network constructed by mining

- PubMed abstracts. *BMC Bioinformatics* **2004**, *5*, 147, doi:10.1186/1471-2105-5-147.
25. Munoz-Torres, M.; Carbon, S. Get GO! retrieving GO data using AmiGO, QuickGO, API, files, and tools. In *Methods in Molecular Biology*; Humana Press Inc., 2017; Vol. 1446, pp. 149–160.
  26. Fabregat, A.; Jupe, S.; Matthews, L.; Sidiropoulos, K.; Gillespie, M.; Garapati, P.; Haw, R.; Jassal, B.; Korninger, F.; May, B.; et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **2018**, *46*, D649–D655, doi:10.1093/nar/gkx1132.
  27. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **2012**, *16*, 284–287, doi:10.1089/omi.2011.0118.
  28. Talbert, P.B.; Henikoff, S. Histone variants on the move: substrates for chromatin dynamics. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 115–126, doi:10.1038/nrm.2016.148.
  29. Lalonde, M.-E.; Cheng, X.; Côté, J. Histone target selection within chromatin: an exemplary case of teamwork. *Genes Dev.* **2014**, *28*, 1029–1041, doi:10.1101/gad.236331.113.
  30. Goll, M.G.; Bestor, T.H. Histone modification and replacement in chromatin activation. *Genes Dev.* **2002**, *16*, 1739–1742, doi:10.1101/gad.1013902.
  31. Rossetto, D.; Avvakumov, N.; Côté, J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics* **2012**, *7*, 1098–1108, doi:10.4161/epi.21975.
  32. Cheung, W.L.; Turner, F.B.; Krishnamoorthy, T.; Wolner, B.; Ahn, S.-H.; Foley, M.; Dorsey, J.A.; Peterson, C.L.; Berger, S.L.; Allis, C.D. Phosphorylation of histone H4 serine 1 during DNA damage requires casein kinase II in *S. cerevisiae*. *Curr. Biol.* **2005**, *15*, 656–660, doi:10.1016/j.cub.2005.02.049.
  33. Utley, R.T.; Lacoste, N.; Jobin-Robitaille, O.; Allard, S.; Côté, J. Regulation of NuA4 histone acetyltransferase activity in transcription and DNA repair by phosphorylation of histone H4. *Mol. Cell. Biol.* **2005**, *25*, 8179–8190, doi:10.1128/MCB.25.18.8179-8190.2005.
  34. Sarmiento, O.F.; Digilio, L.C.; Wang, Y.; Perlin, J.; Herr, J.C.; Allis, C.D.; Coonrod, S.A. Dynamic alterations of specific histone modifications during early murine development. *J. Cell Sci.* **2004**, *117*, 4449–4459, doi:10.1242/jcs.01328.
  35. Basnet, H.; Su, X.B.; Tan, Y.; Meisenhelder, J.; Merkurjev, D.; Ohgi, K.A.; Hunter, T.; Pillus, L.; Rosenfeld, M.G. Tyrosine phosphorylation of histone H2A by CK2 regulates transcriptional elongation. *Nature* **2014**, *516*, 267–271, doi:10.1038/nature13736.
  36. Koutelou, E.; Hirsch, C.L.; Dent, S.Y.R. Multiple faces of the SAGA complex. *Curr. Opin. Cell Biol.* **2010**, *22*, 374–382, doi:10.1016/j.ceb.2010.03.005.
  37. Oh, S.; Sukanuma, T.; Gogol, M.M.; Workman, J.L. Histone H3 threonine 11 phosphorylation by Sch9 and CK2 regulates chronological lifespan by controlling the nutritional stress response. *Elife* **2018**, *7*, doi:10.7554/eLife.36157.
  38. Yang, Q.; Zhu, Q.; Lu, X.; Du, Y.; Cao, L.; Shen, C.; Hou, T.; Li, M.; Li, Z.; Liu, C.; et al. G9a coordinates with the RPA complex to promote DNA damage repair and cell survival. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E6054–E6063,

- doi:10.1073/pnas.1700694114.
39. Becker, J.S.; Nicetto, D.; Zaret, K.S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet.* **2016**, *32*, 29–41, doi:10.1016/j.tig.2015.11.001.
  40. Park, J.-W.; Kim, J.J.; Bae, Y.-S. CK2 downregulation induces senescence-associated heterochromatic foci formation through activating SUV39h1 and inactivating G9a. *Biochem. Biophys. Res. Commun.* **2018**, *505*, 67–73, doi:10.1016/j.bbrc.2018.09.099.
  41. Nishibuchi, G.; Machida, S.; Osakabe, A.; Murakoshi, H.; Hiragami-Hamada, K.; Nakagawa, R.; Fischle, W.; Nishimura, Y.; Kurumizaka, H.; Tagami, H.; et al. N-terminal phosphorylation of HP1 $\alpha$  increases its nucleosome-binding specificity. *Nucleic Acids Res.* **2014**, *42*, 12498–12511, doi:10.1093/nar/gku995.
  42. Hiragami-Hamada, K.; Shinmyozu, K.; Hamada, D.; Tatsu, Y.; Uegaki, K.; Fujiwara, S.; Nakayama, J.-I. N-terminal phosphorylation of HP1{alpha} promotes its chromatin binding. *Mol. Cell. Biol.* **2011**, *31*, 1186–1200, doi:10.1128/MCB.01012-10.
  43. Gouot, E.; Bhat, W.; Rufiange, A.; Fournier, E.; Paquet, E.; Nourani, A. Casein kinase 2 mediated phosphorylation of Spt6 modulates histone dynamics and regulates spurious transcription. *Nucleic Acids Res.* **2018**, *46*, 7612–7630, doi:10.1093/nar/gky515.
  44. Seo, S.; Macfarlan, T.; McNamara, P.; Hong, R.; Mukai, Y.; Heo, S.; Chakravarti, D. Regulation of histone acetylation and transcription by nuclear protein pp32, a subunit of the INHAT complex. *J. Biol. Chem.* **2002**, *277*, 14005–14010, doi:10.1074/jbc.M112455200.
  45. Schneider, R.; Bannister, A.J.; Weise, C.; Kouzarides, T. Direct binding of INHAT to H3 tails disrupted by modifications. *J. Biol. Chem.* **2004**, *279*, 23859–23862, doi:10.1074/jbc.C400151200.
  46. Seo, S.B.; McNamara, P.; Heo, S.; Turner, A.; Lane, W.S.; Chakravarti, D. Regulation of histone acetylation and transcription by INHAT, a human cellular complex containing the set oncoprotein. *Cell* **2001**, *104*, 119–130, doi:10.1016/s0092-8674(01)00196-9.
  47. Reilly, P.T.; Yu, Y.; Hamiche, A.; Wang, L. Cracking the ANP32 whips: important functions, unequal requirement, and hints at disease implications. *Bioessays* **2014**, *36*, 1062–1071, doi:10.1002/bies.201400058.
  48. Estève, P.-O.; Chang, Y.; Samaranyake, M.; Upadhyay, A.K.; Horton, J.R.; Feehery, G.R.; Cheng, X.; Pradhan, S. A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. *Nat. Struct. Mol. Biol.* **2011**, *18*, 42–48, doi:10.1038/nsmb.1939.
  49. VanDemark, A.P.; Kasten, M.M.; Ferris, E.; Heroux, A.; Hill, C.P.; Cairns, B.R. Autoregulation of the rsc4 tandem bromodomain by gcn5 acetylation. *Mol. Cell* **2007**, *27*, 817–828, doi:10.1016/j.molcel.2007.08.018.
  50. Shahbazian, M.D.; Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* **2007**, *76*, 75–100, doi:10.1146/annurev.biochem.76.052705.162114.
  51. Eom, G.H.; Cho, Y.K.; Ko, J.H.; Shin, S.; Choe, N.; Kim, Y.; Joung, H.; Kim, H.S.; Nam, K. II; Kee, H.J.; et al. Casein kinase-2 $\alpha$ 1 induces hypertrophic

- response by phosphorylation of histone deacetylase 2 S394 and its activation in the heart. *Circulation* **2011**, *123*, 2392–2403, doi:10.1161/CIRCULATIONAHA.110.003665.
52. Khan, D.H.; He, S.; Yu, J.; Winter, S.; Cao, W.; Seiser, C.; Davie, J.R. Protein kinase CK2 regulates the dimerization of histone deacetylase 1 (HDAC1) and HDAC2 during mitosis. *J. Biol. Chem.* **2013**, *288*, 16518–28, doi:10.1074/jbc.M112.440446.
  53. Kim, H.S.; Chang, Y.G.; Bae, H.J.; Eun, J.W.; Shen, Q.; Park, S.J.; Shin, W.C.; Lee, E.K.; Park, S.; Ahn, Y.M.; et al. Oncogenic potential of CK2 $\alpha$  and its regulatory role in EGF-induced HDAC2 expression in human liver cancer. *FEBS J.* **2014**, *281*, 851–61, doi:10.1111/febs.12652.
  54. Jia, Z.-M.; Ai, X.; Teng, J.-F.; Wang, Y.-P.; Wang, B.-J.; Zhang, X. p21 and CK2 interaction-mediated HDAC2 phosphorylation modulates KLF4 acetylation to regulate bladder cancer cell proliferation. *Tumour Biol.* **2016**, *37*, 8293–304, doi:10.1007/s13277-015-4618-1.
  55. Reynolds, N.; O’Shaughnessy, A.; Hendrich, B. Transcriptional repressors: multifaceted regulators of gene expression. *Development* **2013**, *140*, 505–512, doi:10.1242/dev.083105.
  56. Zschoernig, B.; Mahlknecht, U. Carboxy-terminal phosphorylation of SIRT1 by protein kinase CK2. *Biochem. Biophys. Res. Commun.* **2009**, *381*, 372–377, doi:10.1016/j.bbrc.2009.02.085.
  57. Kang, H.; Jung, J.-W.; Kim, M.K.; Chung, J.H. CK2 is the regulator of SIRT1 substrate-binding affinity, deacetylase activity and cellular response to DNA-damage. *PLoS One* **2009**, *4*, e6611, doi:10.1371/journal.pone.0006611.
  58. Coin, F.; Auriol, J.; Tapias, A.; Clivio, P.; Vermeulen, W.; Egly, J.-M. Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity. *EMBO J.* **2004**, *23*, 4835–4846, doi:10.1038/sj.emboj.7600480.
  59. Shah, P.; Zhao, B.; Qiang, L.; He, Y.-Y. Phosphorylation of xeroderma pigmentosum group C regulates ultraviolet-induced DNA damage repair. *Nucleic Acids Res.* **2018**, *46*, 5050–5060, doi:10.1093/nar/gky239.
  60. Bandyopadhyay, K.; Li, P.; Gjerset, R.A. CK2-mediated hyperphosphorylation of topoisomerase I targets serine 506, enhances topoisomerase I-DNA binding, and increases cellular camptothecin sensitivity. *PLoS One* **2012**, *7*, e50427, doi:10.1371/journal.pone.0050427.
  61. Seto, E.; Yoshida, M. Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a018713, doi:10.1101/cshperspect.a018713.
  62. Luo, R.X.; Dean, D.C. Chromatin remodeling and transcriptional regulation. *J. Natl. Cancer Inst.* **1999**, *91*, 1288–1294, doi:10.1093/jnci/91.15.1288.
  63. Ropero, S.; Esteller, M. The role of histone deacetylases (HDACs) in human cancer. *Mol. Oncol.* **2007**, *1*, 19–25, doi:10.1016/j.molonc.2007.01.001.
  64. Cai, R.; Kwon, P.; Yan-Neale, Y.; Sambuccetti, L.; Fischer, D.; Cohen, D. Mammalian histone deacetylase 1 protein is posttranslationally modified by phosphorylation. *Biochem. Biophys. Res. Commun.* **2001**, *283*, 445–453, doi:10.1006/bbrc.2001.4786.
  65. Pflum, M.K.; Tong, J.K.; Lane, W.S.; Schreiber, S.L. Histone deacetylase 1



- phosphorylation promotes enzymatic activity and complex formation. *J. Biol. Chem.* **2001**, *276*, 47733–47741, doi:10.1074/jbc.M105590200.
66. Tsai, S.-C.; Seto, E. Regulation of histone deacetylase 2 by protein kinase CK2. *J. Biol. Chem.* **2002**, *277*, 31826–33, doi:10.1074/jbc.M204149200.
  67. Sun, J.-M.; Chen, H.Y.; Moniwa, M.; Litchfield, D.W.; Seto, E.; Davie, J.R. The transcriptional repressor Sp3 is associated with CK2-phosphorylated histone deacetylase 2. *J. Biol. Chem.* **2002**, *277*, 35783–35786, doi:10.1074/jbc.C200378200.
  68. Zhang, X.; Ozawa, Y.; Lee, H.; Wen, Y.-D.; Tan, T.-H.; Wadzinski, B.E.; Seto, E. Histone deacetylase 3 (HDAC3) activity is regulated by interaction with protein serine/threonine phosphatase 4. *Genes Dev.* **2005**, *19*, 827–839, doi:10.1101/gad.1286005.
  69. Watabe, M.; Nakaki, T. Protein kinase CK2 regulates the formation and clearance of aggresomes in response to stress. *J. Cell Sci.* **2011**, *124*, 1519–1532, doi:10.1242/jcs.081778.
  70. Längst, G.; Manelyte, L. Chromatin Remodelers: From Function to Dysfunction. *Genes (Basel)*. **2015**, *6*, 299–324, doi:10.3390/genes6020299.
  71. Buontempo, F.; Orsini, E.; Lonetti, A.; Cappellini, A.; Chiarini, F.; Evangelisti, C.; Evangelisti, C.; Melchionda, F.; Pession, A.; Bertaina, A.; et al. Synergistic cytotoxic effects of bortezomib and CK2 inhibitor CX-4945 in acute lymphoblastic leukemia: turning off the prosurvival ER chaperone BIP/Grp78 and turning on the pro-apoptotic NF- $\kappa$ B. *Oncotarget* **2016**, *7*, 1323–1340, doi:10.18632/oncotarget.6361.
  72. Patil, H.; Wilks, C.; Gonzalez, R.W.; Dhanireddy, S.; Conrad-Webb, H.; Bergel, M. Mitotic Activation of a Novel Histone Deacetylase 3-Linker Histone H1.3 Protein Complex by Protein Kinase CK2. *J. Biol. Chem.* **2016**, *291*, 3158–3172, doi:10.1074/jbc.M115.643874.
  73. Choi, H.-K.; Choi, Y.; Park, E.S.; Park, S.-Y.; Lee, S.-H.; Seo, J.; Jeong, M.-H.; Jeong, J.-W.; Cheong, J.-H.; Lee, P.C.W.; et al. Corrigendum: Programmed cell death 5 mediates HDAC3 decay to promote genotoxic stress response. *Nat. Commun.* **2015**, *6*, 8225.
  74. Lai, A.Y.; Wade, P.A. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat. Rev. Cancer* **2011**, *11*, 588–596, doi:10.1038/nrc3091.
  75. Nair, S.S.; Kumar, R. Chromatin remodeling in cancer: a gateway to regulate gene transcription. *Mol. Oncol.* **2012**, *6*, 611–619, doi:10.1016/j.molonc.2012.09.005.
  76. Ceccacci, E.; Minucci, S. Inhibition of histone deacetylases in cancer therapy: lessons from leukaemia. *Br. J. Cancer* **2016**, *114*, 605–611, doi:10.1038/bjc.2016.36.
  77. Purwin, M.; Hernández-Toribio, J.; Coderch, C.; Panchuk, R.; Skorokhyd, N.; Filipiak, K.; De Pascual-Teresa, B.; Ramos, A. Design and synthesis of novel dual-target agents for HDAC1 and CK2 inhibition. *RSC Adv.* **2016**, *6*, 66595–66608, doi:10.1039/c6ra09717k.
  78. Laugesen, A.; Helin, K. Chromatin repressive complexes in stem cells, development, and cancer. *Cell Stem Cell* **2014**, *14*, 735–751, doi:10.1016/j.stem.2014.05.006.
  79. Kadamb, R.; Mittal, S.; Bansal, N.; Batra, H.; Saluja, D. Sin3: insight into its

- transcription regulatory functions. *Eur. J. Cell Biol.* **2013**, *92*, 237–246, doi:10.1016/j.ejcb.2013.09.001.
80. Torchy, M.P.; Hamiche, A.; Klaholz, B.P. Structure and function insights into the NuRD chromatin remodeling complex. *Cell. Mol. Life Sci.* **2015**, *72*, 2491–2507, doi:10.1007/s00018-015-1880-8.
  81. Liu, Y.; Amin, E.B.; Mayo, M.W.; Chudgar, N.P.; Bucciarelli, P.R.; Kadota, K.; Adusumilli, P.S.; Jones, D.R. CK2 $\alpha$ ' Drives Lung Cancer Metastasis by Targeting BRMS1 Nuclear Export and Degradation. *Cancer Res.* **2016**, *76*, 2675–2686, doi:10.1158/0008-5472.CAN-15-2888.
  82. Liu, Y.; Smith, P.W.; Jones, D.R. Breast cancer metastasis suppressor 1 functions as a corepressor by enhancing histone deacetylase 1-mediated deacetylation of RelA/p65 and promoting apoptosis. *Mol. Cell. Biol.* **2006**, *26*, 8683–8696, doi:10.1128/MCB.00940-06.
  83. Hurst, D.R. Metastasis suppression by BRMS1 associated with SIN3 chromatin remodeling complexes. *Cancer Metastasis Rev.* **2012**, *31*, 641–651, doi:10.1007/s10555-012-9363-y.
  84. Dominguez, I.; Sonenshein, G.E.; Seldin, D.C. Protein kinase CK2 in health and disease: CK2 and its role in Wnt and NF-kappaB signaling: linking development and cancer. *Cell. Mol. Life Sci.* **2009**, *66*, 1850–7, doi:10.1007/s00018-009-9153-z.
  85. Rusin, S.F.; Adamo, M.E.; Kettenbach, A.N. Identification of candidate casein kinase 2 substrates in mitosis by quantitative phosphoproteomics. *Front. Cell Dev. Biol.* **2017**, *5*, doi:10.3389/fcell.2017.00097.
  86. D'Oto, A.; Tian, Q.-W.; Davidoff, A.M.; Yang, J. Histone demethylases and their roles in cancer epigenetics. *J. Med. Oncol. Ther.* **2016**, *1*, 34–40.
  87. Hayami, S.; Kelly, J.D.; Cho, H.-S.; Yoshimatsu, M.; Unoki, M.; Tsunoda, T.; Field, H.I.; Neal, D.E.; Yamaue, H.; Ponder, B.A.J.; et al. Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int. J. cancer* **2011**, *128*, 574–586, doi:10.1002/ijc.25349.
  88. Sareddy, G.R.; Viswanadhapalli, S.; Surapaneni, P.; Suzuki, T.; Brenner, A.; Vadlamudi, R.K. Novel KDM1A inhibitors induce differentiation and apoptosis of glioma stem cells via unfolded protein response pathway. *Oncogene* **2017**, *36*, 2423–2434, doi:10.1038/onc.2016.395.
  89. Nicholson, T.B.; Chen, T. LSD1 demethylates histone and non-histone proteins. *Epigenetics* **2009**, *4*, 129–132, doi:10.4161/epi.4.3.8443.
  90. Peng, B.; Wang, J.; Hu, Y.; Zhao, H.; Hou, W.; Zhao, H.; Wang, H.; Liao, J.; Xu, X. Modulation of LSD1 phosphorylation by CK2/WIP1 regulates RNF168-dependent 53BP1 recruitment in response to DNA damage. *Nucleic Acids Res.* **2015**, *43*, 5936–5947, doi:10.1093/nar/gkv528.
  91. Costa, R.; Arrigoni, G.; Cozza, G.; Lolli, G.; Battistutta, R.; Izipisua Belmonte, J.C.; Pinna, L.A.; Sarno, S. The lysine-specific demethylase 1 is a novel substrate of protein kinase CK2. *Biochim. Biophys. Acta* **2014**, *1844*, 722–729, doi:10.1016/j.bbapap.2014.01.014.
  92. Zhang, J.; Kalkum, M.; Chait, B.T.; Roeder, R.G. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol. Cell* **2002**, *9*, 611–623, doi:10.1016/s1097-2765(02)00468-9.
  93. Glass, C.K.; Rosenfeld, M.G. The coregulator exchange in transcriptional

- functions of nuclear receptors. *Genes Dev.* **2000**, *14*, 121–141.
94. Underhill, C.; Qutob, M.S.; Yee, S.P.; Torchia, J. A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor KAP-1. *J. Biol. Chem.* **2000**, *275*, 40463–40470, doi:10.1074/jbc.M007864200.
  95. Yoo, J.-Y.; Choi, H.-K.; Choi, K.-C.; Park, S.-Y.; Ota, I.; Yook, J.I.; Lee, Y.-H.; Kim, K.; Yoon, H.-G. Nuclear hormone receptor corepressor promotes esophageal cancer cell invasion by transcriptional repression of interferon- $\gamma$ -inducible protein 10 in a casein kinase 2-dependent manner. *Mol. Biol. Cell* **2012**, *23*, 2943–2954, doi:10.1091/mbc.E11-11-0947.
  96. Oswald, F.; Rodriguez, P.; Giaimo, B.D.; Antonello, Z.A.; Mira, L.; Mittler, G.; Thiel, V.N.; Collins, K.J.; Tabaja, N.; Cizelsky, W.; et al. A phospho-dependent mechanism involving NCoR and KMT2D controls a permissive chromatin state at Notch target genes. *Nucleic Acids Res.* **2016**, *44*, 4703–4720, doi:10.1093/nar/gkw105.
  97. Zhou, Y.; Gross, W.; Hong, S.H.; Privalsky, M.L. The SMRT corepressor is a target of phosphorylation by protein kinase CK2 (casein kinase II). *Mol. Cell. Biochem.* **2001**, *220*, 1–13, doi:10.1023/a:1011087910699.
  98. Mikami, S.; Kanaba, T.; Takizawa, N.; Kobayashi, A.; Maesaki, R.; Fujiwara, T.; Ito, Y.; Mishima, M. Structural insights into the recruitment of SMRT by the corepressor SHARP under phosphorylative regulation. *Structure* **2014**, *22*, 35–46, doi:10.1016/j.str.2013.10.007.
  99. Aranda, S.; Mas, G.; Di Croce, L. Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* **2015**, *1*, e1500737, doi:10.1126/sciadv.1500737.
  100. Chittock, E.C.; Latwiel, S.; Miller, T.C.R.; Müller, C.W. Molecular architecture of polycomb repressive complexes. *Biochem. Soc. Trans.* **2017**, *45*, 193–205, doi:10.1042/BST20160173.
  101. Gao, Z.; Lee, P.; Stafford, J.M.; von Schimmelmann, M.; Schaefer, A.; Reinberg, D. An AUTS2-Polycomb complex activates gene expression in the CNS. *Nature* **2014**, *516*, 349–354, doi:10.1038/nature13921.
  102. Banerjee Mustafi, S.; Chakraborty, P.K.; Dwivedi, S.K.D.; Ding, K.; Moxley, K.M.; Mukherjee, P.; Bhattacharya, R. BMI1, a new target of CK2 $\alpha$ . *Mol. Cancer* **2017**, *16*, 56, doi:10.1186/s12943-017-0617-8.
  103. Kawaguchi, T.; Machida, S.; Kurumizaka, H.; Tagami, H.; Nakayama, J.-I. Phosphorylation of CBX2 controls its nucleosome-binding specificity. *J. Biochem.* **2017**, *162*, 343–355, doi:10.1093/jb/mvx040.
  104. Jin, B.; Li, Y.; Robertson, K.D. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer* **2011**, *2*, 607–617, doi:10.1177/1947601910393957.
  105. Deplus, R.; Blanchon, L.; Rajavelu, A.; Boukaba, A.; Defrance, M.; Luciani, J.; Rothé, F.; Dedeurwaerder, S.; Denis, H.; Brinkman, A.B.; et al. Regulation of DNA methylation patterns by CK2-mediated phosphorylation of Dnmt3a. *Cell Rep.* **2014**, *8*, 743–753, doi:10.1016/j.celrep.2014.06.048.
  106. Rank, G.; Cerruti, L.; Simpson, R.J.; Moritz, R.L.; Jane, S.M.; Zhao, Q. Identification of a PRMT5-dependent repressor complex linked to silencing of human fetal globin gene expression. *Blood* **2010**, *116*, 1585–1592,

- doi:10.1182/blood-2009-10-251116.
107. Bronner, C.; Trotzler, M.-A.; Filhol, O.; Cochet, C.; Rochette-Egly, C.; Schöller-Guinard, M.; Klein, J.-P.; Mousli, M. The antiapoptotic protein ICBP90 is a target for protein kinase 2. *Ann. N. Y. Acad. Sci.* **2004**, *1030*, 355–360, doi:10.1196/annals.1329.044.
  108. Bostick, M.; Kim, J.K.; Estève, P.-O.; Clark, A.; Pradhan, S.; Jacobsen, S.E. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **2007**, *317*, 1760–1764, doi:10.1126/science.1147939.
  109. Mousli, M.; Hopfner, R.; Abbady, A.-Q.; Monté, D.; Jeanblanc, M.; Oudet, P.; Louis, B.; Bronner, C. ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. *Br. J. Cancer* **2003**, *89*, 120–127, doi:10.1038/sj.bjc.6601068.
  110. McStay, B.; Grummt, I. The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.* **2008**, *24*, 131–157, doi:10.1146/annurev.cellbio.24.110707.175259.
  111. Goodfellow, S.J.; Zomerdijk, J.C.B.M. Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA genes. *Subcell. Biochem.* **2013**, *61*, 211–236, doi:10.1007/978-94-007-4525-4\_10.
  112. Grummt, I.; Längst, G. Epigenetic control of RNA polymerase I transcription in mammalian cells. *Biochim. Biophys. Acta* **2013**, *1829*, 393–404, doi:10.1016/j.bbagr.2012.10.004.
  113. Yang, L.; Song, T.; Chen, L.; Kabra, N.; Zheng, H.; Koomen, J.; Seto, E.; Chen, J. Regulation of SirT1-nucleomethylin binding by rRNA coordinates ribosome biogenesis with nutrient availability. *Mol. Cell. Biol.* **2013**, *33*, 3835–3848, doi:10.1128/MCB.00476-13.
  114. Vaquero, A.; Scher, M.; Erdjument-Bromage, H.; Tempst, P.; Serrano, L.; Reinberg, D. SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature* **2007**, *450*, 440–444, doi:10.1038/nature06268.
  115. Voit, R.; Seiler, J.; Grummt, I. Cooperative Action of Cdk1/cyclin B and SIRT1 Is Required for Mitotic Repression of rRNA Synthesis. *PLoS Genet.* **2015**, *11*, e1005246, doi:10.1371/journal.pgen.1005246.
  116. Guetg, C.; Lienemann, P.; Sirri, V.; Grummt, I.; Hernandez-Verdun, D.; Hottiger, M.O.; Fussenegger, M.; Santoro, R. The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. *EMBO J.* **2010**, *29*, 2135–2146, doi:10.1038/emboj.2010.17.
  117. Panova, T.B.; Panov, K.I.; Russell, J.; Zomerdijk, J.C.B.M. Casein kinase 2 associates with initiation-competent RNA polymerase I and has multiple roles in ribosomal DNA transcription. *Mol. Cell. Biol.* **2006**, *26*, 5957–5968, doi:10.1128/MCB.00673-06.
  118. Voit, R.; Schnapp, A.; Kuhn, A.; Rosenbauer, H.; Hirschmann, P.; Stunnenberg, H.G.; Grummt, I. The nucleolar transcription factor mUBF is phosphorylated by casein kinase II in the C-terminal hyperacidic tail which is essential for transactivation. *EMBO J.* **1992**, *11*, 2211–2218.
  119. Kappes, F.; Damoc, C.; Knippers, R.; Przybylski, M.; Pinna, L.A.; Gruss, C. Phosphorylation by protein kinase CK2 changes the DNA binding properties of the

- human chromatin protein DEK. *Mol. Cell. Biol.* **2004**, *24*, 6011–20, doi:10.1128/MCB.24.13.6011-6020.2004.
120. Vintermist, A.; Böhm, S.; Sadeghifar, F.; Louvet, E.; Mansén, A.; Percipalle, P.; Ostlund Farrants, A.-K. The chromatin remodelling complex B-WICH changes the chromatin structure and recruits histone acetyl-transferases to active rRNA genes. *PLoS One* **2011**, *6*, e19184, doi:10.1371/journal.pone.0019184.
  121. Cavellán, E.; Asp, P.; Percipalle, P.; Farrants, A.-K.O. The WSTF-SNF2h chromatin remodeling complex interacts with several nuclear proteins in transcription. *J. Biol. Chem.* **2006**, *281*, 16264–16271, doi:10.1074/jbc.M600233200.
  122. Wise-Draper, T.M.; Mintz-Cole, R.A.; Morris, T.A.; Simpson, D.S.; Wikenheiser-Brokamp, K.A.; Currier, M.A.; Cripe, T.P.; Grosveld, G.C.; Wells, S.I. Overexpression of the cellular DEK protein promotes epithelial transformation in vitro and in vivo. *Cancer Res.* **2009**, *69*, 1792–1799, doi:10.1158/0008-5472.CAN-08-2304.
  123. von Lindern, M.; Fornerod, M.; van Baal, S.; Jaegle, M.; de Wit, T.; Buijs, A.; Grosveld, G. The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, results in the fusion of two genes, *dek* and *can*, and the expression of a chimeric, leukemia-specific *dek-can* mRNA. *Mol. Cell. Biol.* **1992**, *12*, 1687–1697, doi:10.1128/mcb.12.4.1687.
  124. Kappes, F.; Waldmann, T.; Mathew, V.; Yu, J.; Zhang, L.; Khodadoust, M.S.; Chinnaiyan, A.M.; Luger, K.; Erhardt, S.; Schneider, R.; et al. The DEK oncoprotein is a Su(var) that is essential to heterochromatin integrity. *Genes Dev.* **2011**, *25*, 673–678, doi:10.1101/gad.2036411.
  125. Aydin, Ö.Z.; Vermeulen, W.; Lans, H. ISWI chromatin remodeling complexes in the DNA damage response. *Cell Cycle* **2014**, *13*, 3016–3025, doi:10.4161/15384101.2014.956551.
  126. Sadeghifar, F.; Böhm, S.; Vintermist, A.; Östlund Farrants, A.-K. The B-WICH chromatin-remodelling complex regulates RNA polymerase III transcription by promoting Max-dependent c-Myc binding. *Nucleic Acids Res.* **2015**, *43*, 4477–4490, doi:10.1093/nar/gkv312.
  127. Poot, R.A.; Bozhenok, L.; van den Berg, D.L.C.; Steffensen, S.; Ferreira, F.; Grimaldi, M.; Gilbert, N.; Ferreira, J.; Varga-Weisz, P.D. The Williams syndrome transcription factor interacts with PCNA to target chromatin remodelling by ISWI to replication foci. *Nat. Cell Biol.* **2004**, *6*, 1236–1244, doi:10.1038/ncb1196.
  128. Bian, Y.; Ye, M.; Wang, C.; Cheng, K.; Song, C.; Dong, M.; Pan, Y.; Qin, H.; Zou, H. Global screening of CK2 kinase substrates by an integrated phosphoproteomics workflow. *Sci. Rep.* **2013**, *3*, 3460, doi:10.1038/srep03460.
  129. Mukhopadhyay, D.; Matunis, M.J. SUMO1-mediated Daxx-mediated repression. *Mol. Cell* **2011**, *42*, 4–5, doi:10.1016/j.molcel.2011.03.008.
  130. Lin, D.-Y.; Huang, Y.-S.; Jeng, J.-C.; Kuo, H.-Y.; Chang, C.-C.; Chao, T.-T.; Ho, C.-C.; Chen, Y.-C.; Lin, T.-P.; Fang, H.-I.; et al. Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Mol. Cell* **2006**, *24*, 341–354, doi:10.1016/j.molcel.2006.10.019.
  131. Chang, C.-C.; Naik, M.T.; Huang, Y.-S.; Jeng, J.-C.; Liao, P.-H.; Kuo, H.-Y.; Ho,

- C.-C.; Hsieh, Y.-L.; Lin, C.-H.; Huang, N.-J.; et al. Structural and functional roles of Daxx SIM phosphorylation in SUMO paralogue-selective binding and apoptosis modulation. *Mol. Cell* **2011**, *42*, 62–74, doi:10.1016/j.molcel.2011.02.022.
132. Shih, H.-M.; Chang, C.-C.; Kuo, H.-Y.; Lin, D.-Y. Daxx mediates SUMO-dependent transcriptional control and subnuclear compartmentalization. *Biochem. Soc. Trans.* **2007**, *35*, 1397–1400, doi:10.1042/BST0351397.
133. Lallemand-Breitenbach, V.; de Thé, H. CK2 and PML: regulating the regulator. *Cell* **2006**, *126*, 244–245, doi:10.1016/j.cell.2006.07.004.
134. Scaglioni, P.P.; Yung, T.M.; Cai, L.F.; Erdjument-Bromage, H.; Kaufman, A.J.; Singh, B.; Teruya-Feldstein, J.; Tempst, P.; Pandolfi, P.P. A CK2-dependent mechanism for degradation of the PML tumor suppressor. *Cell* **2006**, *126*, 269–283, doi:10.1016/j.cell.2006.05.041.
135. Stehmeier, P.; Muller, S. Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. *Mol. Cell* **2009**, *33*, 400–409, doi:10.1016/j.molcel.2009.01.013.
136. Koipally, J.; Renold, A.; Kim, J.; Georgopoulos, K. Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. *EMBO J.* **1999**, *18*, 3090–3100, doi:10.1093/emboj/18.11.3090.
137. Koipally, J.; Georgopoulos, K. Ikaros interactions with CtBP reveal a repression mechanism that is independent of histone deacetylase activity. *J. Biol. Chem.* **2000**, *275*, 19594–19602, doi:10.1074/jbc.M000254200.
138. O’Neill, D.W.; Schoetz, S.S.; Lopez, R.A.; Castle, M.; Rabinowitz, L.; Shor, E.; Krawchuk, D.; Goll, M.G.; Renz, M.; Seelig, H.P.; et al. An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Mol. Cell. Biol.* **2000**, *20*, 7572–7582, doi:10.1128/mcb.20.20.7572-7582.2000.
139. Song, C.; Gowda, C.; Pan, X.; Ding, Y.; Tong, Y.; Tan, B.-H.; Wang, H.; Muthusami, S.; Ge, Z.; Sachdev, M.; et al. Targeting casein kinase II restores Ikaros tumor suppressor activity and demonstrates therapeutic efficacy in high-risk leukemia. *Blood* **2015**, *126*, 1813–1822, doi:10.1182/blood-2015-06-651505.
140. Dovat, S.; Song, C.; Payne, K.J.; Li, Z. Ikaros, CK2 kinase, and the road to leukemia. *Mol. Cell. Biochem.* **2011**, *356*, 201–207, doi:10.1007/s11010-011-0964-5.
141. Song, C.; Li, Z.; Erbe, A.K.; Savic, A.; Dovat, S. Regulation of Ikaros function by casein kinase 2 and protein phosphatase 1. *World J. Biol. Chem.* **2011**, *2*, 126–131, doi:10.4331/wjbc.v2.i6.126.
142. Song, C.; Pan, X.; Ge, Z.; Gowda, C.; Ding, Y.; Li, H.; Li, Z.; Yochum, G.; Muschen, M.; Li, Q.; et al. Epigenetic regulation of gene expression by Ikaros, HDAC1 and Casein Kinase II in leukemia. *Leukemia* **2016**, *30*, 1436–1440.
143. Ge, Z.; Gu, Y.; Han, Q.; Sloane, J.; Ge, Q.; Gao, G.; Ma, J.; Song, H.; Hu, J.; Chen, B.; et al. Plant homeodomain finger protein 2 as a novel IKAROS target in acute lymphoblastic leukemia. *Epigenomics* **2018**, *10*, 59–69, doi:10.2217/epi-2017-0092.
144. Bousset, K.; Henriksson, M.; Lüscher-Firzlaff, J.M.; Litchfield, D.W.; Lüscher, B. Identification of casein kinase II phosphorylation sites in Max: effects on DNA-binding kinetics of Max homo- and Myc/Max heterodimers. *Oncogene* **1993**, *8*, 3211–3220.

145. Laherty, C.D.; Yang, W.M.; Sun, J.M.; Davie, J.R.; Seto, E.; Eisenman, R.N. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* **1997**, *89*, 349–356, doi:10.1016/s0092-8674(00)80215-9.
146. Sommer, A.; Hilfenhaus, S.; Menkel, A.; Kremmer, E.; Seiser, C.; Loidl, P.; Lüscher, B. Cell growth inhibition by the Mad/Max complex through recruitment of histone deacetylase activity. *Curr. Biol.* **1997**, *7*, 357–365, doi:10.1016/s0960-9822(06)00183-7.
147. Krippner-Heidenreich, A.; Talanian, R. V.; Sekul, R.; Kraft, R.; Thole, H.; Ottleben, H.; Lüscher, B. Targeting of the transcription factor Max during apoptosis: phosphorylation-regulated cleavage by caspase-5 at an unusual glutamic acid residue in position P1. *Biochem. J.* **2001**, *358*, 705–715, doi:10.1042/0264-6021:3580705.
148. Pelengaris, S.; Khan, M.; Evan, G. c-MYC: more than just a matter of life and death. *Nat. Rev. Cancer* **2002**, *2*, 764–776, doi:10.1038/nrc904.
149. Hagiwara, T.; Nakaya, K.; Nakamura, Y.; Nakajima, H.; Nishimura, S.; Taya, Y. Specific phosphorylation of the acidic central region of the N-myc protein by casein kinase II. *Eur. J. Biochem.* **1992**, *209*, 945–950, doi:10.1111/j.1432-1033.1992.tb17367.x.
150. Vervoorts, J.; Lüscher-Firzlaff, J.; Lüscher, B. The ins and outs of MYC regulation by posttranslational mechanisms. *J. Biol. Chem.* **2006**, *281*, 34725–34729, doi:10.1074/jbc.R600017200.
151. Shannon, P. PSICQUIC: Proteomics Standard Initiative Common QUery InterfaCe. 2019.
152. Hoffmeister, H.; Fuchs, A.; Erdel, F.; Pinz, S.; Gröbner-Ferreira, R.; Bruckmann, A.; Deutzmann, R.; Schwartz, U.; Maldonado, R.; Huber, C.; et al. CHD3 and CHD4 form distinct NuRD complexes with different yet overlapping functionality. *Nucleic Acids Res.* **2017**, *45*, 10534–10554, doi:10.1093/nar/gkx711.
153. Rowbotham, S.P.; Barki, L.; Neves-Costa, A.; Santos, F.; Dean, W.; Hawkes, N.; Choudhary, P.; Will, W.R.; Webster, J.; Oxley, D.; et al. Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1. *Mol. Cell* **2011**, *42*, 285–296, doi:10.1016/j.molcel.2011.02.036.
154. Doiguchi, M.; Nakagawa, T.; Imamura, Y.; Yoneda, M.; Higashi, M.; Kubota, K.; Yamashita, S.; Asahara, H.; Iida, M.; Fujii, S.; et al. SMARCAD1 is an ATP-dependent stimulator of nucleosomal H2A acetylation via CBP, resulting in transcriptional regulation. *Sci. Rep.* **2016**, *6*, 20179, doi:10.1038/srep20179.
155. Doyon, Y.; Cayrou, C.; Ullah, M.; Landry, A.-J.; Côté, V.; Selleck, W.; Lane, W.S.; Tan, S.; Yang, X.-J.; Côté, J. ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol. Cell* **2006**, *21*, 51–64, doi:10.1016/j.molcel.2005.12.007.
156. Han, X.; Gui, B.; Xiong, C.; Zhao, L.; Liang, J.; Sun, L.; Yang, X.; Yu, W.; Si, W.; Yan, R.; et al. Destabilizing LSD1 by Jade-2 promotes neurogenesis: an antibraking system in neural development. *Mol. Cell* **2014**, *55*, 482–494, doi:10.1016/j.molcel.2014.06.006.
157. Ramakrishna, S.; Suresh, B.; Lee, E.-J.; Lee, H.-J.; Ahn, W.-S.; Baek, K.-H. Lys-63-specific deubiquitination of SDS3 by USP17 regulates HDAC activity. *J. Biol.*

- Chem.* **2011**, 286, 10505–10514, doi:10.1074/jbc.M110.162321.
158. Arends, T.; Dege, C.; Bortnick, A.; Danhorn, T.; Knapp, J.R.; Jia, H.; Harmacek, L.; Fleenor, C.J.; Straign, D.; Walton, K.; et al. CHD4 is essential for transcriptional repression and lineage progression in B lymphopoiesis. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 10927–10936, doi:10.1073/pnas.1821301116.
  159. Tang, L.; Nogales, E.; Ciferri, C. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog. Biophys. Mol. Biol.* **2010**, *102*, 122–128, doi:10.1016/j.pbiomolbio.2010.05.001.
  160. Padilla-Benavides, T.; Haokip, D.T.; Yoon, Y.; Reyes-Gutierrez, P.; Rivera-Pérez, J.A.; Imbalzano, A.N. CK2-Dependent Phosphorylation of the Brg1 Chromatin Remodeling Enzyme Occurs during Mitosis. *Int. J. Mol. Sci.* **2020**, *21*, doi:10.3390/ijms21030923.
  161. Clapier, C.R.; Iwasa, J.; Cairns, B.R.; Peterson, C.L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 407–422, doi:10.1038/nrm.2017.26.
  162. Mohrmann, L.; Verrijzer, C.P. Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta* **2005**, *1681*, 59–73, doi:10.1016/j.bbexp.2004.10.005.
  163. Marfella, C.G.A.; Imbalzano, A.N. The Chd family of chromatin remodelers. *Mutat. Res.* **2007**, *618*, 30–40, doi:10.1016/j.mrfmmm.2006.07.012.
  164. Murawska, M.; Brehm, A. CHD chromatin remodelers and the transcription cycle. *Transcription* **2011**, *2*, 244–253, doi:10.4161/trns.2.6.17840.
  165. Bao, Y.; Shen, X. INO80 subfamily of chromatin remodeling complexes. *Mutat. Res.* **2007**, *618*, 18–29, doi:10.1016/j.mrfmmm.2006.10.006.
  166. Morrison, A.J.; Shen, X. Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 373–384, doi:10.1038/nrm2693.
  167. Sif, S.; Saurin, A.J.; Imbalzano, A.N.; Kingston, R.E. Purification and characterization of mSin3A-containing Brg1 and hBrm chromatin remodeling complexes. *Genes Dev.* **2001**, *15*, 603–618, doi:10.1101/gad.872801.
  168. Allen, H.F.; Wade, P.A.; Kutateladze, T.G. The NuRD architecture. *Cell. Mol. Life Sci.* **2013**, *70*, 3513–3524, doi:10.1007/s00018-012-1256-2.
  169. Giansanti, P.; Tsiatsiani, L.; Low, T.Y.; Heck, A.J.R. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protoc.* **2016**, *11*, 993–1006, doi:10.1038/nprot.2016.057.
  170. Laugesen, A.; Højfeldt, J.W.; Helin, K. Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Mol. Cell* **2019**, *74*, 8–18, doi:10.1016/j.molcel.2019.03.011.
  171. Tamburri, S.; Lavarone, E.; Fernández-Pérez, D.; Conway, E.; Zanotti, M.; Manganaro, D.; Pasini, D. Histone H2AK119 Mono-Ubiquitination Is Essential for Polycomb-Mediated Transcriptional Repression. *Mol. Cell* **2020**, *77*, 840–856.e5, doi:10.1016/j.molcel.2019.11.021.
  172. Zhang, Y.; Yang, X.; Gui, B.; Xie, G.; Zhang, D.; Shang, Y.; Liang, J. Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27. *J. Biol. Chem.* **2011**, 286, 42414–42425, doi:10.1074/jbc.M111.271064.



173. Margueron, R.; Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **2011**, *469*, 343–349, doi:10.1038/nature09784.
174. Kaneko, S.; Li, G.; Son, J.; Xu, C.-F.; Margueron, R.; Neubert, T.A.; Reinberg, D. Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes Dev.* **2010**, *24*, 2615–2620, doi:10.1101/gad.1983810.
175. Tie, F.; Stratton, C.A.; Kurzhals, R.L.; Harte, P.J. The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. *Mol. Cell. Biol.* **2007**, *27*, 2014–2026, doi:10.1128/MCB.01822-06.
176. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515, doi:10.1093/nar/gky1049.
177. Carlson, M.; Bioconductor Package Maintainer UniProt.ws: R Interface to UniProt Web Services. 2019.
178. Ooms, J. curl: A Modern and Flexible Web Client for R. 2019.
179. Gavali, S.; Cowart, J.; Chen, C.; Ross, K.E.; Arighi, C.; Wu, C.H. RESTful API for iPTMnet: a resource for protein post-translational modification network discovery. *Database (Oxford)*. **2020**, *2020*, doi:10.1093/database/baz157.
180. Ruepp, A.; Waegele, B.; Lechner, M.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Mewes, H.-W. CORUM: the comprehensive resource of mammalian protein complexes--2009. *Nucleic Acids Res.* **2010**, *38*, D497–D501, doi:10.1093/nar/gkp914.
181. Gustavsen, J.A.; Pai, S.; Isserlin, R.; Demchak, B.; Pico, A.R. RCy3: Network biology using Cytoscape from within R. *F1000Research* **2019**, *8*, 1774, doi:10.12688/f1000research.20887.2.

### 3.10 Supplemental Materials

#### Tables

**Supp. Table 3.1.** Literature text mining for co-mentions of CK2 and chromatin organization and histone proteins.

Sheet	Name	Content
1	Table 1	Reutils searches. Chilibot search: CK2 and histone. Sentences containing either CK2 or HISTONE.
Appendix A		Supplemental Material – Chapter 3

**Supp. Table 3.2. Chromatin organization data sets.**

Sheet	Name	Content
1	Table 2	AmiGO: chromatin organization (GO:0006338) data set. Reactome: chromatin-modifying enzymes (R-HSA-3247509) data set. Reactome: epigenetic regulation of gene expression (R-HSA-212165) data set.
Appendix A		Supplemental Material – Chapter 3

**Supp. Table 3.3.** Functional ranking of chromatin organization proteins to CK2.

Sheet	Name	Content
1	Table 3	AmiGO: chromatin organization (GO:0006338) data set similarity ranking to CK2.

		Reactome: chromatin-modifying enzymes (R-HSA-3247509) data set similarity ranking to CK2. Reactome: epigenetic regulation of gene expression (R-HSA-212165) data set similarity ranking to CK2.
Appendix A		Supplemental Material – Chapter 3

**Supp. Table 3.4.** Subunits of complexes linked to CK2 and chromatin organization.

Sheet	Name	Content
1	Table 4	CORUM complexes. Complex Portal complexes.
Appendix A		Supplemental Material – Chapter 3

**Supp. Table 3.5.** CK2 interactors by subunits involved in chromatin organization.

Sheet	Name	Content
1	Table 5	CK2 interactors in chromatin remodeling. CSNK2A1 interactors. CSNK2A2 interactors. CSNK2B interactors.
Appendix A		Supplemental Material – Chapter 3

**Supp. Table 3.6.** GO subcellular compartment classification of CK2 interactors involved in chromatin organization.

Sheet	Name	Content
1	Table 6	CK2 interactors retrieved from PSICQUIC service, June 2020. GO CC functional classification, level 5.
Appendix A		Supplemental Material – Chapter 3

**Supp. Table 3.7.** Phosphosites modulated when CK2 is manipulated using Asp-N as protein digestion enzyme involved in chromatin organization and transcription.

Sheet	Name	Content
1	Table 7	Phosphosites modulated when CK2 is manipulated using Asp-N as protein digestion enzyme involved in chromatin organization and transcription.
Appendix A		Supplemental Material – Chapter 3
Appendix B		MS identification and quantification

**Figures**

**PSP Regulatory sites**

Source	Dataset	No.substrates	No.regulatory.sites	No.domain
Reactome	Epigenetics	29	110	13
AmiGO	Chromatin remodeling	49	238	32
Reactome	Chromatin remodeling	63	230	27
Total	Merged	115	472	55

Source	Dataset	p	m2	m1	ac	ub	me	sm	m3	gl
Reactome	Epigenetics	82	8	7	15	3	8	NA	5	NA
AmiGO	Chromatin remodeling	204	7	3	21	3	3	1	NA	NA
Reactome	Chromatin remodeling	186	9	5	22	5	7	4	7	3
Total	Merged	401	15	8	36	8	9	5	7	3

**Merged kinase-substrate**

Source	Dataset	No.substrates	No.kinases	No.kinase.substrate	No.phosphosites
Total	Merged	168	179	1142	658

**iPTMnet enzyme-substrate**

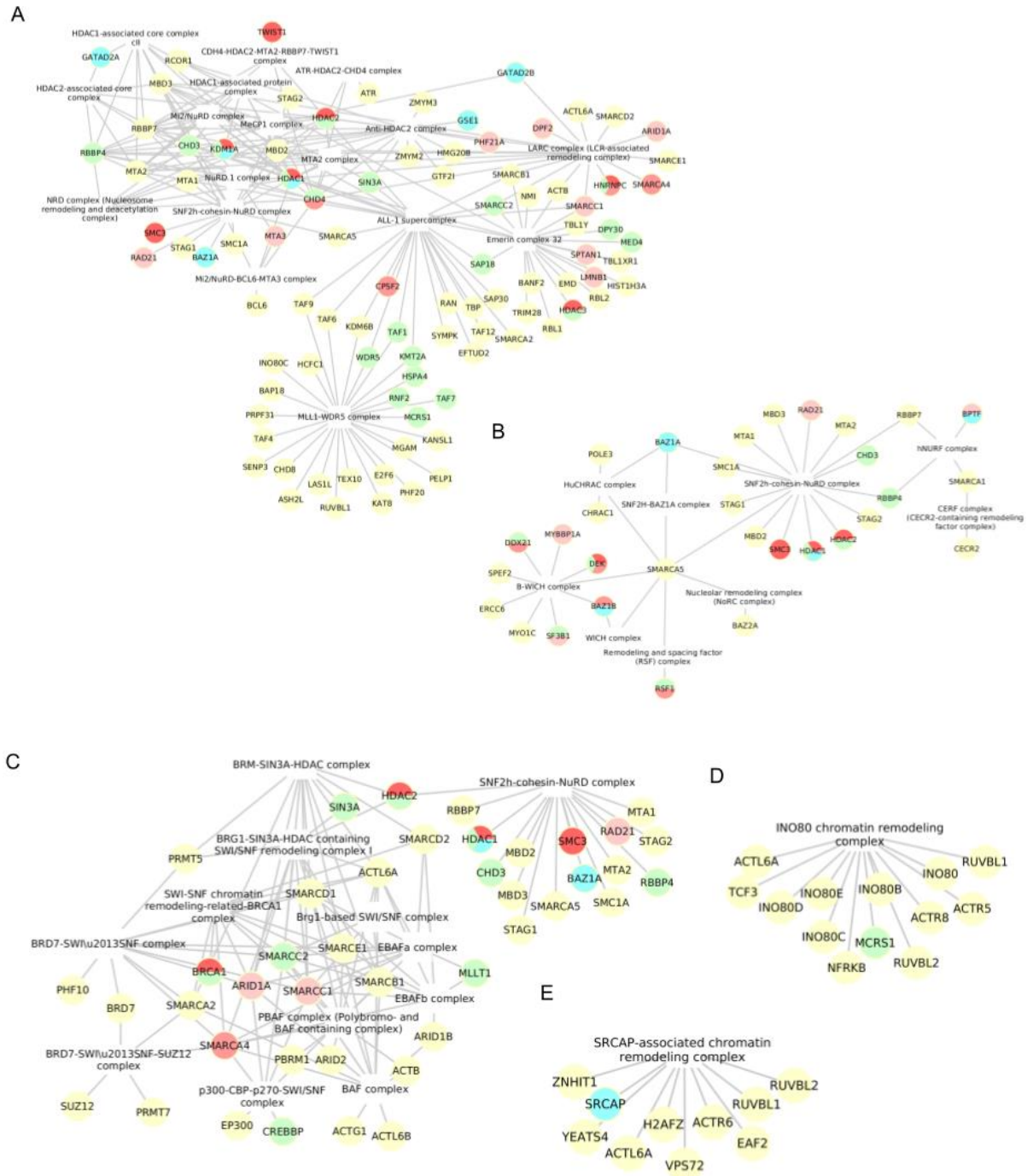
Source	Dataset	No.substrates	No.enzymes	No.sites	No.substrate.enzyme
Reactome	Epigenetics	40	83	135	236
Reactome	Chromatin remodeling	77	99	250	364
AmiGO	Chromatin remodeling	78	136	318	579
Total	Merged	158	200	593	1004

Source	Dataset	Interaction	Acetylation	Phosphorylation	Methylation
Reactome	Epigenetics	4	5	126	NA
Reactome	Chromatin remodeling	11	14	222	3
AmiGO	Chromatin remodeling	13	8	297	NA
Total	Merged	25	22	543	3

**PSP kinase-substrate**

Source	Dataset	No.substrates	No.kinases	No.sites	No.kinase.substrate	No.domain
Reactome	Chromatin remodeling	73	85	219	318	30
Reactome	Epigenetics	39	77	128	206	15
AmiGO	Chromatin remodeling	54	98	215	392	30
Total	Merged	128	140	463	765	59

**Supp. Figure 3.1.** Summary of enzyme-substrate relationships and site information retrieved from iPTMnet and PhosphoSitePlus (PSP) databases. Annotation ID: Reactome (R-HSA-3247509, R-HSA-212165) and AmiGO (GO:0006338). NA: not available.



**Supp. Figure 3.2.** Chromatin-remodeling complexes and CK2 highlighting core subunits that are CK2 substrates and/or are modulated upon CK2 inhibition. A) CHD complexes. B) ISWI complexes. C) SNF/SWI complexes. D) INO80 complex. E) SRCAP complex. Nodes style scheme: yellow circle (protein), white square (complex), dark red (CK2 substrate), medium red (Bian *et al.* 2013, fold change  $\geq 2$ ), light red (Bian *et al.* 2013), blue (Rusin *et al.* 2017), green (PPI).

## Chapter 4

### 4 Strategies for the functional analysis of lists of modification sites

Several bioinformatics tools are available to functionally contextualize long lists of proteins by focusing, for instance, on the representation of annotations of interest such as GO ontologies and pathways [1,2]. However, fewer tools are available for the contextualization of lists of post-translational modifications (PTMs) sites, for example, those identified or found differentially modulated in phosphoproteome [3] and acetylome experiments [4]. In such cases, the tools for analyzing protein lists are of use but the specificity decreases since information on the modified sites are disregarded [2]. Only downstream analysis pipelines that incorporate site-specific functional and quantitative data can assist in uncovering the patterns and dynamics of protein modifications. Thus, providing insights into signal integration and the regulation of biological processes and pathways in the cell; potentiating the development of novel molecular-targeted therapies.

Here we describe strategies that implement and interconnect existing bioinformatic tools for the functional analysis of lists of modification sites. We cover the topics: pattern matching and discovery [5,6]; retrieval and/or extraction of information from databases and the literature concerning known PTM information [7], kinase-substrate relationships and molecular interactions [8–10]; prediction of kinase-substrate relationships [11]; prediction of PTM functional relevance [12]; functional association to PTM writing enzymes [13]; motif and domain representation [14]. We also remark on the advantages of designing and executing the analysis strategies in R for which Bioconductor [15] is available. Bioconductor is an open software initiative that provides a suite of packages for the analysis of high-throughput genomic data [15]. A notable example of analysis implementation in R is the INKA (Integrative Inferred Kinase Activity) pipeline [16], which integrates kinase-centric and substrate-centric information for inferring kinase activity data from lists of phosphosites and other phosphoproteomic data. In order to do so, INKA brings together experimental information on phosphopeptides; data from UniProt (protein identifiers), PhosphoSitePlus (phosphorylation data), KinBase (protein kinases) and HGNC (gene symbols) databases; and the tools Phomics (kinase activation

loops) and NetworKIN (prediction of kinase-substrate relationships). Another example is the web service SELPHI [17], which performs GO term enrichment on clusters of phosphopeptides, motif enrichment, and phosphosite-phosphatase correlation, etc. SELPHI integrates information from more than 10 different tools and databases. Here, capitalizing on our group's practical experience studying signaling by protein kinase CK2, we provide examples that illustrate the use of analysis pipelines complementary to INKA and SELPHI for uncovering biologically relevant information for this and kinase substrates. More importantly, we describe the implementation in R of a report and functional analysis tool, visualRepo, that features strategies for exploring PTM interplay and for the functional association of PTM lists to writer enzymes.

## 4.1 Functional analysis strategies

### 4.1.1 Pattern scan and discovery

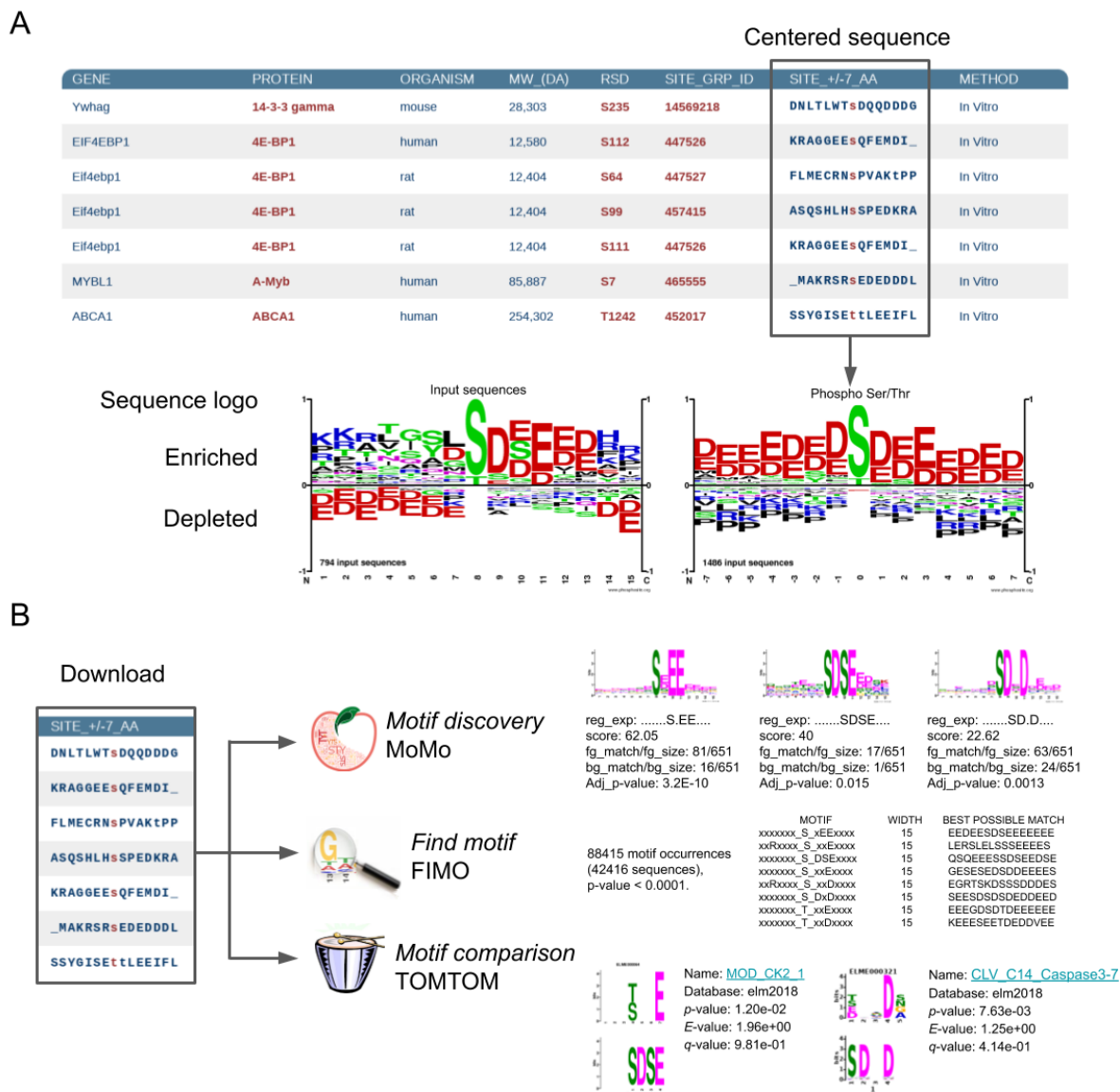
High-throughput studies, such as phosphoproteome studies, identify hundreds of modified peptides that change in the cell in response to perturbations. These lists of peptides are frequently represented as sequences centered at the modified residue with a range of flanking residues anywhere from 6 to 30. The centered sequences are frequently visualized as sequence logos and analyzed for pattern discovery. This allows making inferences regarding the writer enzyme(s) modulated in the system of interest. Examples of available sequence logo generators are PhosphoSitePlus (PSP) sequence Logo Analysis [9,18], WebLogo 3 [19], and Two Sample Logo [20]. Tools that can be used for pattern (motif) discovery in PTM data are MoMo [6] from the Meme suite [21] (implements motif-x algorithm) and PSP Motif Analysis Generator [9,18]. Additionally, if a previous hypothesis exists, such as inhibition of a given enzyme with a chemical probe, scanning the sequences for specific consensus motifs is usually performed. Tools that can be used to scan a protein database or specific sequences with user-defined motifs are ScanProsite [22], EMBOSS' patmatdb [23], and FIMO [5] from the MeMe suite. A distinctive feature of ScanProsite search is that pattern matches can not only be performed against the UniProt Knowledgebase (protein sequence database) [24] but also against the PDB (protein structure database) [25]; this way researchers can also retrieve structural information on the observed hits. Another useful functionality is provided by

PSP Site Search which retrieves a list of modified sites matching a given pattern and the output can be filtered by molecular weight, cell type, cell line, tissue, treatment, GO ontologies, etc. [9,18]. Finally, using R the linear patterns of interest can be defined as regular expressions and the search performed using the `grep` function, and its variants `regexpr` and `gregexpr`, on a list of protein/peptide sequences. The sequences can be uploaded for processing from a local copy of UniProt database in fasta format using Bioconductor's package `Biostrings` [26] or directly from UniProt database using Bioconductor's package `UniProt.ws` [27]. The package `Biostrings` also provides the functions `(v)matchPattern` and `(v)countPattern` for pattern scan. Sequence information from PDB about secondary structure [25] can also be uploaded into R. The sequences can be searched and based on the position of the hits inside the sequence, the corresponding secondary structure information of the hits can be obtained.

#### 4.1.1.1 Pattern scan and motif discovery strategies applied to CK2 datasets

The list of CK2 target sites was downloaded from the PSP database and the corresponding +/- 7 centered sequences (Fig. 4.1A, Supp. Table 4.1) were analyzed using pattern scan and motif discovery strategies. The sequences were represented in a sequence logo (Fig. 4.1A). As expected of an acidophilic kinase, the input sequences were found enriched on negatively charged residues when using all the annotated Ser/Thr phosphosites in PSP as background (Fig. 4.1A). However, when using the input sequences as background we further determined that the positions downstream of the phosphosite (C-terminal) were significantly more enriched in negatively charged residues (Fig. 4.1A). Thus, by including both background settings in our analysis we were able to recapitulate the working model for CK2 substrate specificity where positions +1 to +3 are deemed the most relevant for phosphorylation [28,29]. Next, the CK2 target sequences were searched for motif discovery using `MoMo v.5.0.5`. Interestingly, several motifs were discovered (adjusted P value  $\leq 0.015$ ) making it possible to cluster the CK2 target sequences (Fig. 4.1B, Supp. Table 4.2). The observed motifs were then used to scan the human proteome for occurrences using `FIMO v.5.0.5` and we found that the discovered motifs were widely represented, > 80,000 hits in > 42,000 sequences (P value < 0.0001)

(Fig. 4.1B, Supp. Table 4.3). The global scope of CK2-dependent phosphorylation has been shown in phosphoproteomic studies and in a wealth of low-throughput experiments describing novel CK2 sites [28,30].



**Figure 4.1.** Motif discovery and pattern matching tools and strategies. A) Dataset centered around the modified site, known CK2 phosphorylation target sites (PhosphoSitePlus (PSP)) centered at +/-7 amino acids, and the corresponding sequence logo. The sequence logo was generated using the PSP Production algorithm (sum of absolute values equal to one, background selection: input sequences or Phosho Ser/Thr at <http://www.phosphosite.org/sequenceLogoAction.do>). B) Motif discovery (MoMo v.5.0.5), search (FIMO v.5.0.5), and comparison (Tomtom v.5.0.5) performed on the



centered list of CK2 target sites. Selected examples of the observed output are shown for illustration purposes.

The discovery of the motif SDSE (adjusted P value 0.015) among CK2 substrates suggested that other phosphorylatable residues can be found in the vicinity of CK2 target sites. Upon close examination of the hits, we found that several of these sites also match the CK2 consensus or are reported as known CK2 sites. Thus, by performing motif discovery, we found that CK2 target sites can appear in clusters for example Ser421/Ser423 of HDAC1, Ser773/Ser775 of FURIN, and Ser250/Ser252 of MYC (Supp. Table 4.1-2). Other motifs discovered also indicated the occurrence of Ser nearby the phosphorylation site (Supp. Table 4.2). An interesting observation made by our group points to hierarchical phosphorylation as a regulatory event on CK2-dependent phosphorylation where a given kinase can phosphorylate and prime a sequence for CK2-mediated phosphorylation [31]. Thus, by applying motif discovery to a set of sequences centered at the modulated sites PTM interplay could also be uncovered; this refers to the coordination of proteins' PTMs in cis or trans to determine a function [32]. To further explore PTM interplay, we performed a motif comparison using Tomtom [33] which compares discovered motifs against known motifs annotated in the Eukaryotic Linear Motif [34] (ELM 2018) resource. We found that not only our discovery motif matched the CK2 motif (P value 0.012), as expected, but it also matched the Caspase-3 and -7 cleavage motif (P value 0.00763) (Fig. 4.1B). Interestingly, our group described a regulatory role of CK2 on preventing cleavage by caspases via phosphorylation of caspase cleavage motifs [35].

#### 4.1.2 Post-translational modification data retrieval

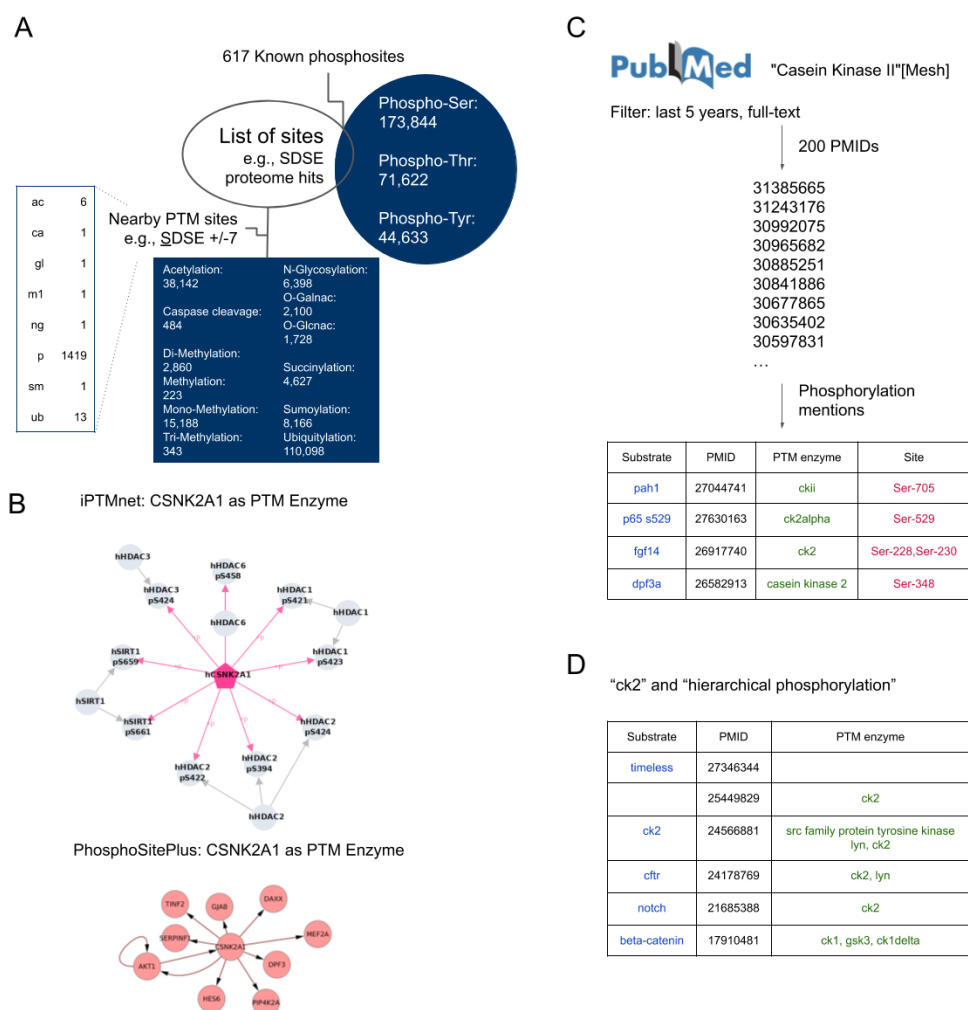
The quality of identifications as well as the functional relevance of the sites found in each experiment can be assessed by retrieving information from PTM databases and extracting sites mentioned in the literature. In addition, the patterns/motifs discovered in your list of interest can be explored to retrieve the PTMs and enzymes targeting them. For this purpose, several PTM databases are available; some include annotations on low- and/or high-throughput experiments, expert-curated data from the literature, enzyme-substrate relationships, and/or information regarding PTM interplay. Examples of PTM databases

are PSP [9], iPTMnet [10], and PTMcode v2 [36]. PSP data can be accessed using the Site Search functionality and the retrieved information downloaded and processed in R or it can be automatically accessed from Cytoscape via the PhosphoSitePlugin [9]. Programmatic access to iPTMnet can be achieved using the CRAN package `iptmnetr` [37], an API client for importing iPTMnet data into R. The iPTMnet data can also be accessed using Batch Retrieval but is limited to 500 items (sites) per query. The iPTMnet database collects information from > 10 PTM resources including PSP. The PTM interplay data in PTMcode v2 data can be downloaded and processed in R. Importantly, PTM databases might not be up-to-date, thus mining the literature to extract kinase-substrate relationships becomes important. For this purpose, the information extraction system RLIMS-P 2.0 [7] can be used for mining kinase, substrate, and phosphorylation information from abstracts in MEDLINE and open-access manuscripts in PubMed Central. With RLIMS-P a maximum of 200 PMIDs of interest can be analyzed in a single run or a keyword search can be defined. The information extracted with RLIMS-P can be downloaded in table format and processed and analyzed in R. Finally, if some of the observed sites or the sites of interest could not be found in the databases/literature a global prediction of PTM sites could be performed using the tool PTMscape [38] for gathering *in silico* evidence.

#### 4.1.2.1 Post-translational modification data retrieval strategies applied to CK2 datasets

The SDSE motif, represented in a subset of CK2 target sites, was overlapped to the phosphorylation data in PSP and 617 known phosphosites were retrieved as matches (Fig. 4.2A, Supp. Table 4.4). Given the relevance of PTMs co-occurring nearby CK2 sites we also looked for other modification sites at a window of  $\pm 7$  amino acids from the phosphorylatable residue in the SDSE motif. This analysis pointed to the occurrence of 8 different PTM types including a caspase cleavage site, which exemplifies what we observed in the motif comparison experiment using Tomtom (Fig. 4.2A, Supp. Table 4.5). A complementary analysis could also be performed by searching the iPTMnet database for the sites matching the SDSE motif. Both iPTMnet and PSP provide the means to visualize kinase-substrate relationship data using network format (Fig. 4.2B).

The corresponding kinase-substrate relationships of the 617 matches were retrieved from PSP and iPTMnet databases and 11 CK2 target sites, one CDC7 target site, one CSNK1A1 target site, and one MAPK14 target site were found (Supp. Table 4.6). Furthermore, we searched for SDSE matches in a phosphoproteomic study of *in vitro* phosphorylation by CK2 [30]. As a result, we found a total of 35 phosphosites matching the SDSE motif found phosphorylated *in vitro* by CK2 in Bian *et al.* 2013 (Supp. Table 4.6). Thus, by searching the motif(s) discovered against PTM databases and phosphoproteomic studies from the literature we were able to uncover information regarding potential CK2 target sites in our list of interest.



**Figure 4.2.** PTM profile of CK2 substrates. A) SDSE proteome hits annotated as phosphorylated in PhosphoSitePlus (PSP) database. PTM types annotated in the vicinity (+/-7 amino acids) of phosphorylated SDSE hits. B) iPTMnet data network representation

of lysine deacetylases phosphorylated by CK2; PSP data using the Cytoscape plugin C) RLIMS-P phosphorylation mention extraction from full-text articles on CK2 published in the last 5 years. D) Keyword RLIMS-P information extraction using the query “ck2” and “hierarchical phosphorylation”.

Finally, we also extracted phosphorylation mentions from the literature using RLIMS-P to complement the existing database information. To use recent publications, we performed a PubMed search using the “Casein Kinase II” MesH term and selected 200 PMIDs belonging to the last five years and representing full-text publications (Fig. 4.2C). This RLIMS-P search yielded at least four recent papers describing CK2-dependent phosphorylation (Fig. 4.2C). Since hierarchical phosphorylation provides a regulatory mechanism at the substrate level for specifying CK2-dependent phosphorylation we also extracted relevant mentions from the literature using RLIMS-P by defining the query: “ck2” and “hierarchical phosphorylation” (Fig. 4.2D). As a result, we identified seven instances where hierarchical phosphorylation was described in the context of CK2 (Fig. 4.2D). For example, we found a regulatory role for hierarchical phosphorylation at the CK2 target site Ser511 of CFTR [39,40].

### 4.1.3 Kinase enrichment and activity inference analyses\*

Several tools apply the known kinase-substrate relationships retrieved from the databases/literature and the direction of the change of the phosphosites/phosphopeptides to find enriched or depleted kinase activities. Some of the available kinase enrichment tools are KSEA (KSEAapp R package [41]), KinSwing (KinSwingR R package [42]), PHOXRACK [43], KEA2 [44], KinasePA (directPA R package [45]), and CLUster Evaluation or Clue (ClueR R package [46]). Another useful tool is the algorithm KARP [47] (no software implementation available), which uses the normalized sum of the intensities of the phosphopeptides containing known kinase target sites that were identified in a phosphoproteomic experiment to rank kinase activity. KSEA and KinSwing are useful for analyzing single conditions of phosphoproteomic studies, whereas Clue was developed for the analysis of time-course experiments.

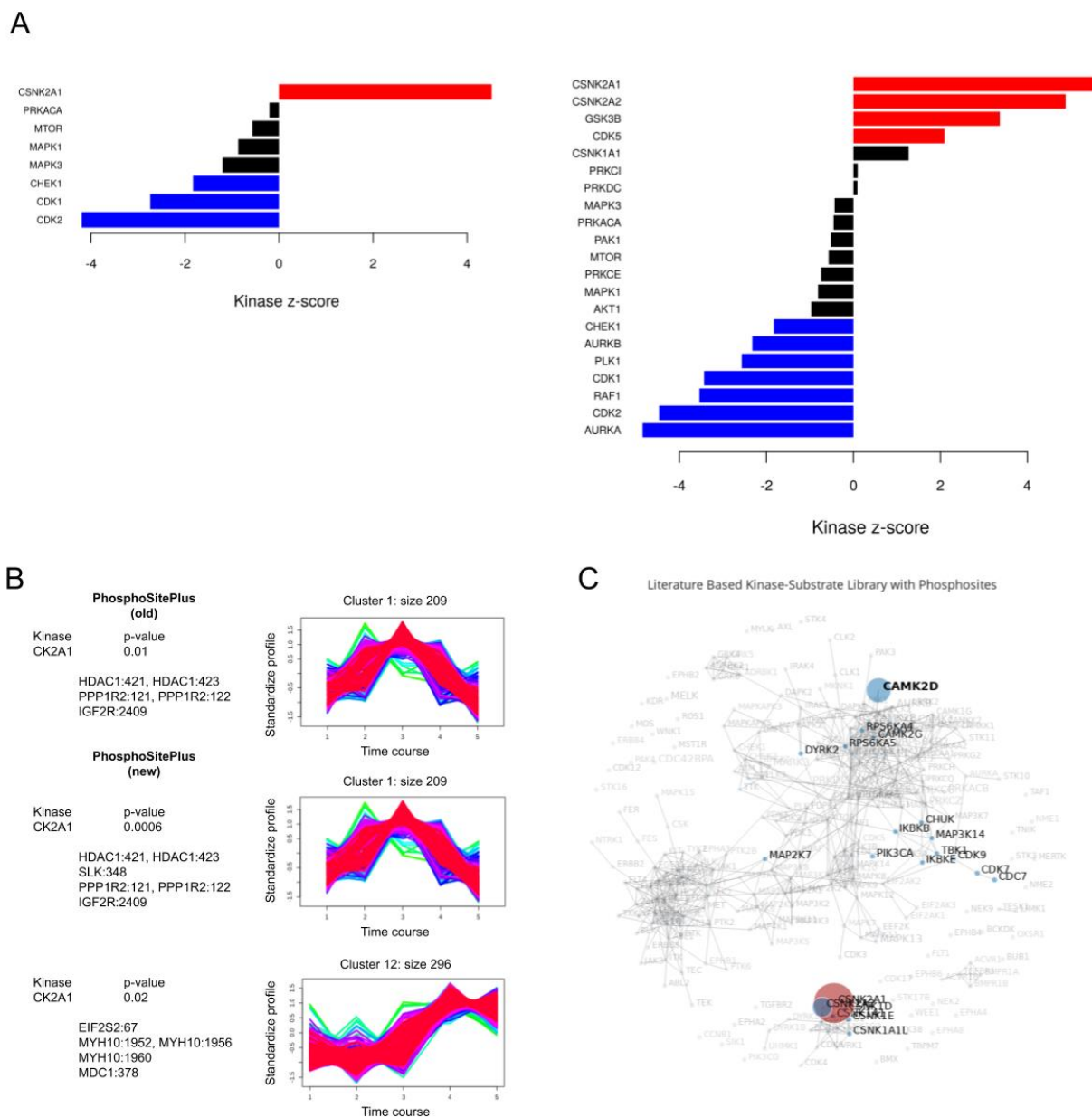
Some important considerations when analyzing phosphoproteome data are: the enrichment and kinase activity prediction results are limited to available PTM database

annotations, not all of the sites identified on multiply modified peptides may be responding to the treatment, and not all of the substrates of a kinase might be impacted by the treatment. Tools like KSEA deal with incomplete database annotation by including kinase-substrate prediction [41]. On the other hand, kinase enrichment analysis can also be performed with kinase-substrate data sets provided by the researcher using the `enricher` function of the Bioconductor package `clusterProfiler` [48]. The uncertainty introduced by considering multiply phosphorylated peptides can be reduced by comparing the kinase enrichment results obtained for all peptides and for the singly modified peptides. Finally, researchers should look at evidence from other functional analysis strategies before eliminating any modified sites from downstream functional analysis and before obviating the contribution of kinases that did not appear modulated.

#### 4.1.3.1 Kinase enrichment and activity inference strategies applied to CK2 datasets\*

The sample dataset provided with KSEA [41], already formatted with the correct input, was analyzed with KSEA's web implementation using either the PSP kinase-substrate dataset or the PSP + NetworKIN dataset (Supp. Table 4.7). In both cases, we observed significantly upregulated (red) and downregulated (blue) kinases (Fig. 4.3A). As expected, selecting the PSP + NetworKIN option (cutoff: 5) impacted the results with more relationships identified compared to PSP only (Fig. 4.3A). Among the kinases found significantly enriched we observed CK2. Since the log<sub>2</sub> fold change represents treatment vs control we could observe that collectively the sites annotated as CK2 were upregulated by the treatment (Fig. 4.3A). We also applied KinSwingR to infer kinase activity modulation in a phosphoproteome study of CK2 inhibition using the ATP-competitive inhibitor CX-4945 (study by Edward Cruise *et al.*, <https://ir.lib.uwo.ca/etd/5164>). As expected, the prediction results indicated CK2 activity inhibition (P value = 0.001) by CX-4945 (Supp. Table 4.7). Next, the sample data provided with ClueR (R package for Clue), a profile of temporal changes in the phosphoproteome of human embryonic stem cells during differentiation, was analyzed [49] (Supp. Table 4.7). Several clusters were observed with cluster 1 containing 5 known CK2 substrates (P value = 0.01) (Fig. 4.3B). Since the PSP release used by ClueR is at

least 5 years old we modified the data for CK2 substrates to reflect the current release of PSP. In correspondence, we identified the same cluster (cluster 1) with an additional CK2 substrate (P value = 0.0006) (Fig. 4.3B). We also observed a new cluster, cluster 12, with 5 new CK2 substrates (Fig. 4.3B). The latter highlights the importance of including up-to-date information and should be extended to other kinases and tools like KinSwingR. The changes of the CK2 sites across time can be followed by looking at the overall trend of each of the clusters containing them (Fig. 4.3B). Of note the P value calculated for the kinase in each cluster and can be significant or not depending if the kinase substrates are spread across different clusters like CK2, in this case.



**Figure 4.3.** Kinase set enrichment analysis. A) KSEA analysis using PhosphoSitePlus (PSP) kinase-substrate dataset (right; default parameters) or PSP kinase-substrate dataset + NetworkKIN prediction (left; NetworkKIN score cutoff: 5 and set substrate count cutoff: 15); in red, black, and blue kinases whose activity was found increased, unchanged, and decreased, respectively. B) CLUSTER Evaluation (Clue) analysis of a time-course experiment profiling the changes to the phosphoproteome during embryonic stem cell development; clusters identified containing CK2 substrates using the old PSP release provided with the ClueR package or a modified version to include the CK2 substrates from the latest release; P value for a given kinase calculated using Fisher's exact test. C) KEA2 analysis using the literature-based kinase-substrate package as background and the CK2 target sites reported in PSP as input.

Finally, the list of CK2 substrates was analyzed using the literature-based kinase-substrate library provided by KEA2. Besides CK2, other kinases were found to target the CK2 sites reported in PSP (Fig. 4.3C). This provided information on kinase integration and consensus sequence overlap.

#### 4.1.4 Kinase-substrate relationship prediction\*

Although PTM databases provide information on kinase-substrate relationships the information available is limited. This imposes a knowledge gap when trying to assess the functional relevance and the biological context of the changes observed in an experiment. As an alternative, tools that perform kinase-substrate relationship predictions are available in addition to those for pattern matching and motif discovery. A widely used kinase-substrate prediction tool is NetworKIN/NetPhorest [8]. Other available tools are GPS [50], iGPS [51], Musite [52], NetPhos [53], HPRD's PhosphoMotif Finder [54], KinasePhos 2.0 [55], CoPhosK [56], KSP-PUEL [57], PhosphoPICK [58], and PKIS [59]. CoPhosK combines the NetworKIN approach of motif matching and proximity in STRING [60] with co-phosphorylation networks. In the latter, a phosphosite is likely to be associated with a given kinase depending on the proportion of its neighbors that are known to be associated with the kinase [56]. Of notice, the sequence databases used by some of these tools might not be up-to-date.

##### 4.1.4.1 Kinase-substrate relationship prediction strategies applied to CK2 datasets\*

The list of CK2 substrates retrieved from PSP was used for kinase-substrate relationship prediction with NetworKIN. A total of 447 out of 502 known sites (89.0%) were predicted by NetworKIN as matches to the CK2\_group (Table 4.1, Supp. Table 4.8). After applying the  $\geq 2$  score cutoff, tool default, a total of 251 sites remained (56.2%) (Table 4.1, Supp. Fig. 4.1A). By including the matches for the CSNK2A2 catalytic subunit (CK2a2 id) only two sequences were found in addition to those matched to CSNK2A1 (CK2alpha id) (Supp. Fig. 4.1B). The netphorest score was the same for sites matching either subunit since the specificity of both CK2 catalytic subunits is considered equal as observed in the sequence logo generated from the hit sequences (Supp. Fig.



4.1C). The difference in the number of hits between subunits was attributed to differences in the STRING functional profiles of both CK2 subunits. As reflected in the STRING score portion of the NetworKIN score (Supp. Fig. 4.1D). This behavior was expected, by incorporating STRING we hoped to include data that allows us to differentiate subunit-specific substrate pools. On the other hand, the sequence logo analysis also showed that most of the sequences below the threshold matched the CK2 motif, as expected due to the input data used (Supp. Fig. 4.1C). Thus, the number of sites with scoring  $< 2$  could stem from a lack of information on the STRING database regarding CK2 and the input proteins to be classified as substrates.

We also performed a CoPhosK prediction using the list of CK2 substrates as input and were able to link 50 sites back to CK2 (Supp. Table 4.8). The limited number of hits compared to the NetworKIN prediction might be because the query sites were not identified in the context-specific phosphoproteomic experiments used by CoPhosK to build the co-phosphorylation network [56]. Next, we applied the iGPS (*in vivo* GPS) predictor, which uses STRING and experimental molecular interaction information to reduce false positives [51]. The inclusion of experimental protein-protein interaction information allowed the assignment of 126 and 55 sites to CSNK2A1 and CSNK2A2, respectively (6% false-positive rate) (Supp. Table 4.8).

Finally, we used KSP-PUEL to predict CK2 substrates using as input the sites identified in an insulin stimulation of 3T3-L1 adipocyte cells (example dataset). This tool does not rely on STRING or other biological annotations for prediction instead, it uses a positive-unlabeled ensemble learning model [57]. As a result, we found 1,546 hits for CK2 out of 12,062 identified sites in the input list. The chosen cut-off was 0.70 since it includes the majority of the known substrates found in the PSP dataset of bonafide CK2 targets, 17 substrates out of 18 (Table 4.1, Supp. Table 4.8). The lowest score for all the known CK2 substrates was 0.48 like what was observed for mTOR substrates [57]. The motif found enriched in  $> 1,500$  hit sequences found by KSP-PUEL was [ST]D[SEDT]E which fits the known CK2 substrates provided as input. KSP-PUEL can be used in conjunction with CLUE or other kinase enrichment tools for finding the kinases represented and then choosing a kinase of interest for substrate prediction with KSP-PUEL [57].

**Table 4.1.** Prediction of kinase-substrate relationships for known human CK2 substrates (PhosphoSitePlus).

Tool	Input and Database	Prediction results	Cutoff
NetworKIN 3.0 (Web)	List of Protein, Site, and AA Human - UniProt 2013/01 (MaxQuant)	Proteins found: 244/251 (97.2%) CK2_group: CK2alpha (447 hits), CK2a2 (447 hits) $\geq 2$ (cutoff): 251 (56.2%) $< 2$ (cutoff): 196 (43.8%) CK2alpha: Median: 3.02907 Mean: 9.87922 $\geq 2$ (cutoff): 249 (55.7%) $< 2$ (cutoff): 198 (44.3%) CK2a2: Median: 0.6063 Mean: 1.9800 $\geq 2$ (cutoff): 87 (19.5%) $< 2$ (cutoff): 360 (80.3%)	Minimum score: 2.00 (0.99) Max. Difference: 4.00 Max. No. of predictions: 2
CoPhosK (Web)	List of Protein, Site	Sites found: 50/502 (9.96%) CK2 sites: 44/50	-
iGPS 1.0 (Standalone, Java)	List of peptides with phosphosite as pS/pT/pY	Proteins found: 86/251 (34.26%) Predictor Other/CK2: CK2a1 (176 hits) CK2a2 (144 hits) CK2a1 interaction: Exp.: 126 String: 50 CK2a2 interaction: Exp.: 1 Exp & String: 54 String: 89	Score cut-off: 3.57 (Medium: FPRs of 6%)
KSP-PUEL (Standalone, Java)	Phosphoproteomics data with a column of centered peptides. List of substrates for a given kinase as Substrate;Site;	Known CK2 sites: 17 Predicted CK2 sites: 1,546	Score cut-off: 0.70 (17 known substrates $\geq$ cut-off)

Overall, since different tools use different biological annotations (e.g., functional association with STRING, co-regulation in phosphoproteomics, protein-protein interaction) the inclusion of more than one tool in the analysis strategy will provide useful to uncover kinase-substrate relationships by considering both the overlap and the union of their outputs. Tools like CoPhosK are of great interest since it can be adapted to include context-specific phosphoproteomic studies that will allow for the generation of novel co-phosphorylation networks (dynamic) and thus for the identification of previously unknown relationships.

#### 4.1.5 Functional association to writer enzymes

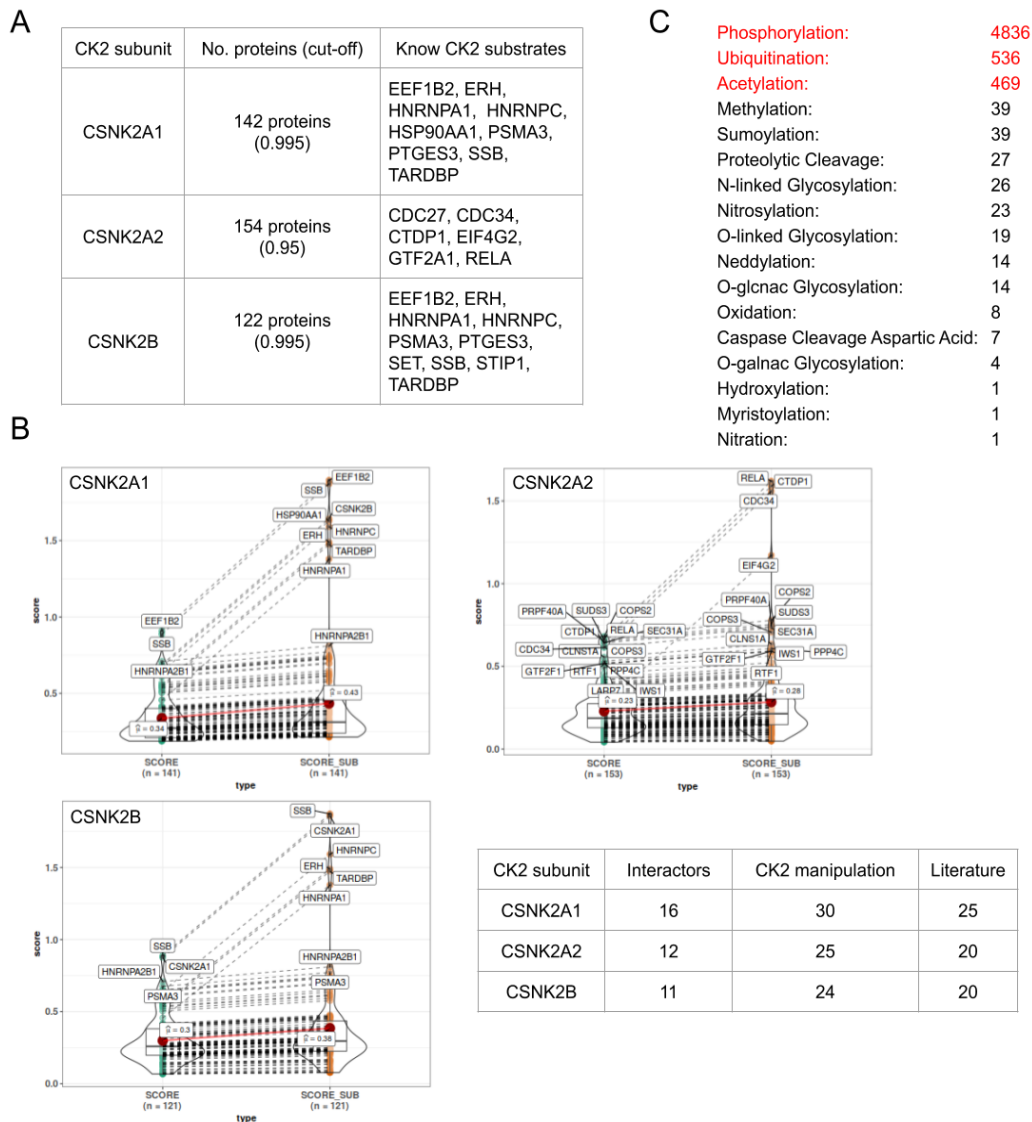
A novel and essential resource for exploring kinase-substrate relationships comes from the ProteomeHD project [13], a co-regulation map of the human proteome. This

proteomic knowledge base provides quantitative information regarding changes in abundance of > 10,000 human proteins in response to 294 biological perturbations. By searching for a given writer enzyme, such as a kinase, experimental evidence is retrieved functionally linking the enzyme to a subset of co-regulated proteins. Given that PTMs can impact the stability of substrates, it is safe to assume that the list of co-regulated proteins may include known but also unknown substrates of the enzyme of interest as previously observed for substrates of APC/C-dependent degradation [61]. This strategy does not provide information at the site level but in combination with any of the other strategies discussed here provides a functional hypothesis for further validating candidate substrates. Other methods to functionally link a writer enzyme to a list of proteins is to apply GO similarity using the Bioconductor's package GoSemSim [62] or to find the proteins in the list that are co-mentioned in Pubmed abstracts and full-text PMC articles using R's package reutils [63]. Another way to estimate the functional association between sites of interest and a given writer enzyme is to find potential interplaying residues with the known target sites of the enzyme using PTMcode v2 data [36] or PTMscape crosstalk prediction for neighboring sites [38].

#### 4.1.5.1 Functional association to CK2

The list of known CK2 substrates from PSP and iPTMnet was overlapped to the proteins co-regulated with the CK2 subunits in ProteomeHD (Fig. 4.4A, Supp. Table 4.9). The ProteomeHD cutoff was selected in a way to retrieve approximately the top 150 co-regulated proteins for each of the CK2 subunits (Fig. 4.4A). The top co-regulated proteins were searched for known CK2 substrates (PSP and iPTMnet databases) and we found 8 overlapping hits for CSNK2A1 and CSNK2B and 1 and 2 unique hits for CSNK2A1 and CSNK2B, respectively (Fig. 4.4A). Interestingly, none of the top co-regulated CK2 substrates were shared between CSNK2A1/CSNK2B and CSNK2A2 (Fig. 4.4A); 6 hits were found for CSNK2A2 (Fig. 4.4A). Next, the list of co-regulated proteins of each CK2 subunit was functionally linked to CK2 considering interactors, literature mention, GO similarity, CK2 motif and subcellular localization, CK2 manipulation, etc. (see Chapter 2). As expected, several functionally linked proteins to CK2 was observed (outliers) (Fig. 4.4B) beyond the known CK2 substrates (Fig. 4.4B, SCORE\_SUB > 1).

As before, the difference observed in the outliers between CSNK2A2 and CSNK2A1/CSNK2B exemplify the functional differences between CK2 subunits (Fig. 4.4A-B).



**Figure 4.4.** Functional association of a list of modulated phosphoproteins to CK2. A) CK2 subunits co-regulated proteins in ProteomeHD as per the selected cutoff; known CK2 substrates found co-regulated. B) Functional link to CK2 of co-regulated proteins by CK2 subunits (outliers: higher ranking proteins, CK2 substrates: SCORE\_SUB > 1); summary table. C) PTMcode v2 interplay information available for known CK2 target sites; the number of interplaying sites for each represented PTM type. Venn diagrams generated with jvenn [64].

Finally, we also looked at the potential interplay between CK2 target sites and other PTM sites using information retrieved from PTMcode v2 and found 6,064 potential interplaying sites for 267 known human CK2 target sites (Fig. 4.4D). Three PTM types: phosphorylation, ubiquitination, and acetylation provided the most information followed by 14 other PTM types also found represented (Fig. 4.4D). The sources for the identified functional interplay are co-evolution evidence (6,042 hits), same residue competition (4 hits), manual evidence (2 hits), and structural distance (36 hits). This analysis points to the need to study CK2-dependent signaling in the context of other writers and erases such as lysine acetyltransferases and lysine deacetylases, respectively (see Chapter 2).

#### 4.1.6 Molecular interaction

Protein-protein interactions [65,66] are necessary for cellular signaling and the maintenance of all cellular functions. Several public databases provide molecular interaction information for proteins like BioGrid [67]; a comprehensive list of molecular interaction databases can be found on Pathguide [68]. The IMEx Consortium is a collaborative effort that integrates information from major public databases providing a non-redundant set of protein interactions [69]. The IMEx application allows retrieval of information using the PSICQUIC service [11] (PSI (Proteome Standard Initiative) Common Query Interface), which accesses all IMEx participating databases. The obtained hits can be clustered to get rid of redundant binary interactions (undirected) annotated using identifiers from a wealth of databases. A package for accessing PSICQUIC services is available in Bioconductor. Cytoscape interface also provides built-in access to PSICQUIC services using “import network from database” [11].

##### 4.1.6.1 Molecular interaction retrieval strategies applied to CK2\*

Using PSICQUIC services from Cytoscape or R [70] we interrogated the IMEx database for retrieving known CK2 interactors. We observed that although common interactors are reported, clear differences exist in the interactions established by each of the CK2 subunits [71] (see Chapter 3). The molecular interaction was overlapped with co-regulation information (Fig. 4.4B), highlighting functional differences between subunits again. In this regard, CSNK2A1 and CSNK2B shared 10 co-regulated proteins as

interactors but none was shared between these and CSNK2A2 or between the three subunits (Supp. Table 4.10). Finally, since protein-protein interaction can be readily retrieved it can be easily integrated to any functional analysis strategy and uncover potential writer-substrate associations. For instance, here we identified 8, 5, and 6 CK2 co-regulated interactor proteins that have been mentioned in the literature along with CK2 for CSNK2A1, CSNK2A2, and CSNK2B, respectively (Supp. Table 4.10). This means that likely there is available information for exploring their association to CK2.

#### 4.1.7 Functional relevance of modification

Proteins contain functional motifs (UniProt: a short conserved sequence, < 20 amino acid, with associated experimental evidence) and domains (UniProt: a specific combination of secondary structure with a characteristic 3D structure of fold). The occurrence of modulated phosphosites inside these protein features could then indicate an effect on protein function. Positional and non-positional information regarding protein motifs and domains can be retrieved from UniProt, which uses InterPro [14] for predicting the occurrence of domains [72]. InterPro provides a framework to functionally classify proteins by predicting domains and important sites using member databases like PROSITE, SMART, and Pfam [14]. This information can be mined to retrieve the domain information for a given list of sites. Alternatively, the PSP database provides domain and PTM regulatory effect information for annotated phosphosites [9].

If modification sites on kinases were expected to be modulated or to test this hypothesis, an Activation loop analysis using Phomics [12] can be performed. This will output those peptides that contain phosphosites localized to the activation loop of kinases which could imply a direct modulation of the kinase's activity and will likely impact the regulation of dependent signaling pathways and biological processes [12]. Another useful resource is the database PTMfunc [73], a repository for predicting the functional relevance of confidently localized PTMs identified by mass-spectrometry for a given protein. PTMfunc considers if a site is in a globular or unstructured region of the protein and makes use of structural information or the frequency of nearby modifications to assess functional relevance, respectively [73]. It annotates residues that are likely found in the interface and if their modification impacts protein interaction. Furthermore, it also

provides information on the presence of PTM hotspots, regions of a protein that are enriched in PTMs [73].

#### 4.1.7.1 Functional relevance strategies applied to CK2 datasets

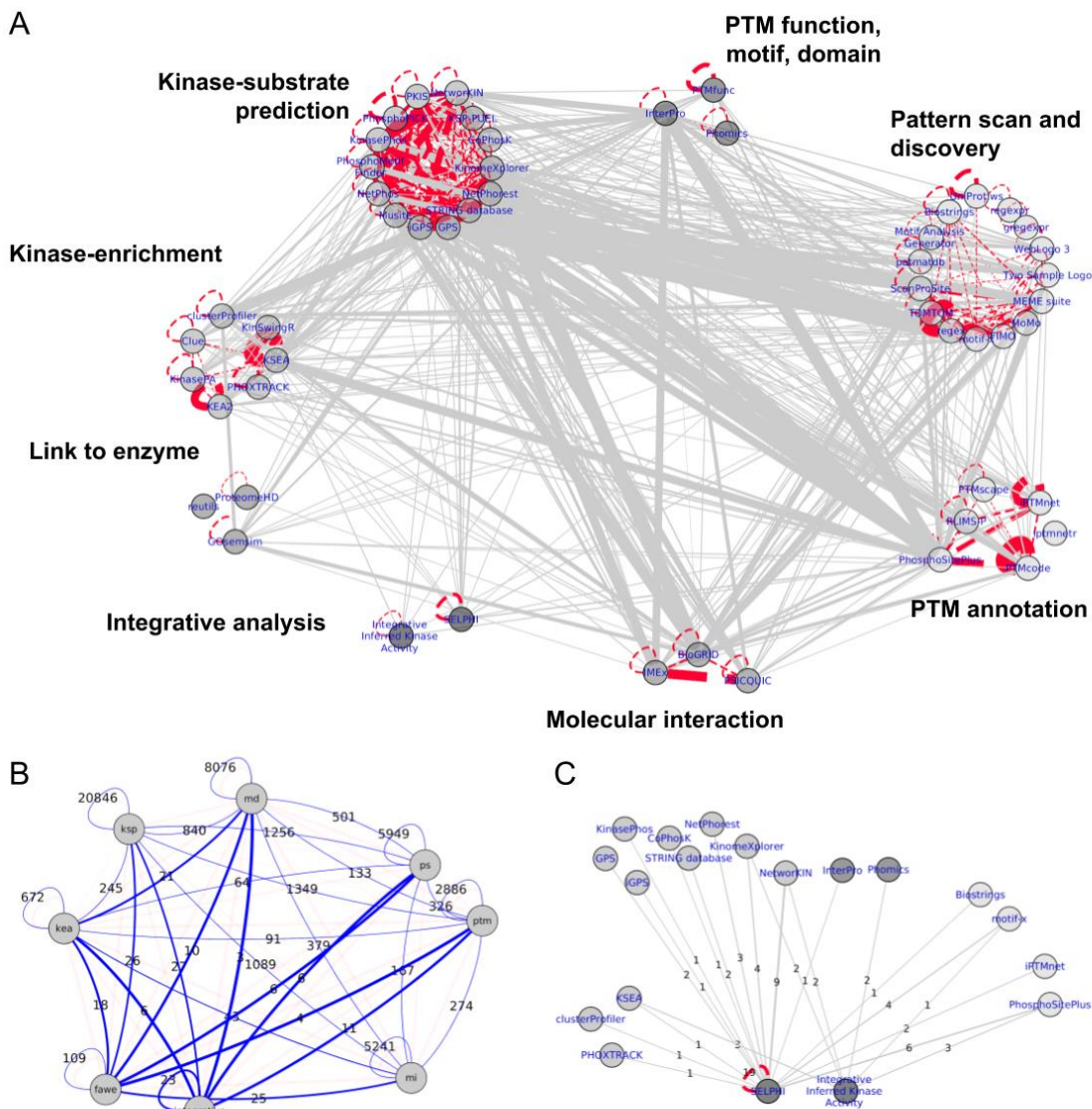
An example of how these tools can be integrated to retrieve information for the CK2 target sites Ser421 and Ser423 of the histone deacetylase HDAC1 can be found in Supp. Fig. 4.2. Briefly, information indicates that these sites are conserved and although they localize to a disordered region outside the deacetylase domain their phosphorylation induces HDAC1 activity and regulated HDAC interactions (Supp. Fig. 4.2). Regulation of phosphorylation occurs in cell cycle and differentiation and is involved in cell adhesion, differentiation, and cytoskeletal reorganization (Supp. Fig. 4.2). Additionally, we also looked at CK2 target sites annotated in motifs or as regulatory using the information provided by PSP and found the representation of > 60 different domains and 364 regulatory associations (Supp. Table 4.11). After mining the regulatory information provided by PSP, it is evident that more experimental work is needed to understand the impact of CK2-dependent phosphorylation in the function of these domains.

#### 4.1.8 Frequently used tools

To explore the functional analysis strategies commonly applied by researchers we automated Google Scholar searches and retrieved from the scholarly literature the occurrences of the tools/databases discussed here (December 2019) (Supp. Table 4.12), in studies mentioning both kinase and proteomic. The tools/databases with the most (> 1,200) hits found were: InterPro, GPS, STRING, BioGRID, NetPhos, PSP, motif-x, MeMe suite, and ScanProSite (Fig. 4.5A). These data suggest that up to December 2019 these were among the most frequently used tools/databases by researchers in the field. Furthermore, since applying different strategies was useful for extracting biologically relevant information from CK2 datasets, we also looked for studies that offer information on the integration of the analysis strategies proposed here. For this, we retrieved the co-mentions (pairs) between tools/databases belonging to the same or different strategies discussed here (Fig. 4.5A). The mentions retrieved were visualized in a network and we found that six strategies out of eight were highly interconnected (Supp. Table 4.13, Fig.

4.5B). We could observe that ksp-ptm (kinase-substrate prediction and PTM resources: 1,349), ksp-ps (kinase-substrate prediction and pattern scan: 1256), and ksp-mi (kinase-substrate prediction and molecular interaction: 1,089), ksp-md (kinase-substrate prediction and motif, domain, and function: 840), md-ps (motif, domain, and function and pattern scan: 501), and kea-ksp (kinase enrichment analysis and kinase-substrate prediction: 245) had the highest number of co-occurrences (co-mention) between groups. Whereas the less represented co-occurrences were fawe-md (functional association with writer enzyme and motif, domain and function: 10), fawe-ps (functional association with writer enzyme and pattern scan: 6), fawe-ptm (functional association with writer enzyme and PTM resource: 4). Suggesting that fewer studies apply tools that functionally link proteins to modifying enzymes by mining the literature or by considering GO similarities in combination with other strategies. We also found that the tools SELPHI and INKA (Integrative Inferred Kinase Activity) since implemented, i.e., 2015 and 2019, respectively, have been mentioned along with other databases/tools (Fig. 4.5C). These results point to the complementarity of the available tools but also highlights the need to perform analysis strategies that include as many tools/databases from different strategies as possible. Here, by doing so we were able to recapitulate known experimental information on protein kinase CK2.





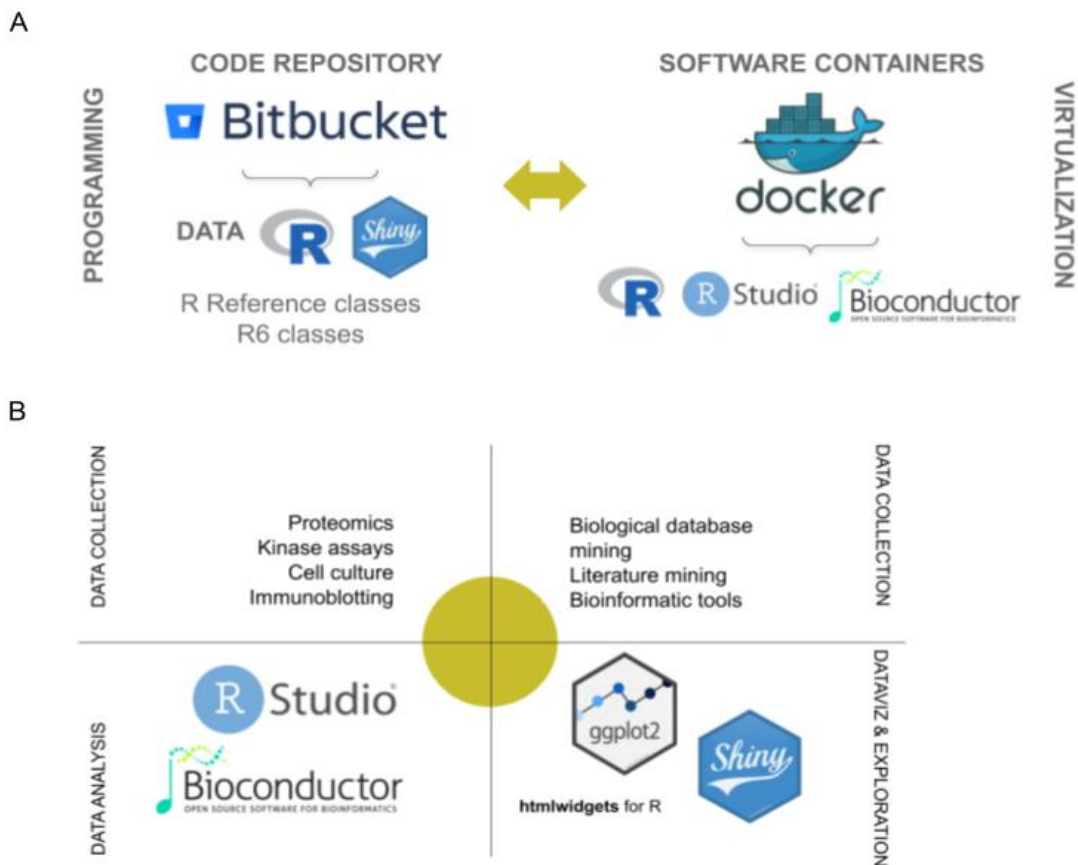
**Figure 4.5.** Scholarly literature that co-mention (pairs) the tools and databases described here. A) Cluster of strategies where nodes represent the tools/databases and the edges the number of co-mentions retrieved from the literature; the edge width represents the number of associations retrieved from low (thin) to high (thick). B) The number of observations shared intra- and inter-group; the edge width represents the number of associations retrieved from high (thin) to low (thick). C) Subnetwork representing the number of observations retrieved for the integrative strategies SELPHI and INKA. ksp: kinase-substrate prediction, ps: pattern scan, ptm: PTM resources, kea: kinase enrichment analysis, fawe: functional association with enzyme writer, mi: molecular interaction, md: motif, domain, and function.

## 4.2 visualRepo implementation

A data-driven R web-based application, that I named visualRepo (currently at version 2), was developed by me to summarize the data processing and data analysis steps of the functional analyses performed during my studies on CK2-dependent phosphorylation and PTM interplay. By implementing visualRepo we made sure that the data sets obtained and the bioinformatic methods implemented could be of use to others studying CK2, phosphorylation, acetylation, or other kinases and PTM types. Thus, visualRepo not only facilitates the reanalysis and the reuse of data but also that of the R code. The application visualRepo is available at [www.visualrepo.cloverpath.ca/visualRepo](http://www.visualrepo.cloverpath.ca/visualRepo). Alternatively, a local copy can be started using the RStudio image provided in the repository rstudio-vrepo:0.2 available at docker hub along with the code in the Appendix C (see the readme file).

### 4.2.1 Development environment

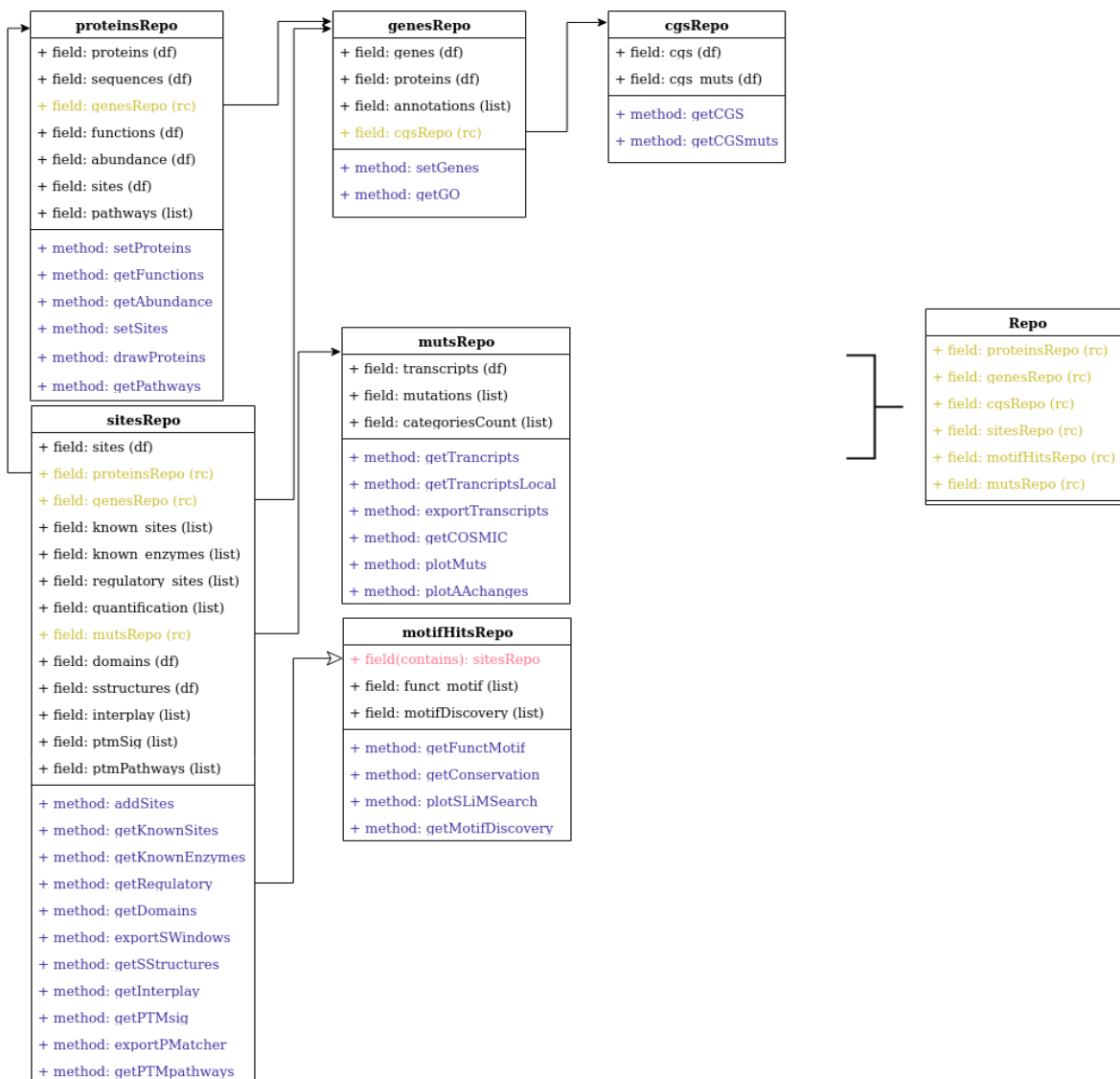
The R programming language R v.3.6.3 (2020-02-29, "Holding the Windssock") [74] was selected for coding visualRepo because it is powerful in statistics, the packages extending R, are well documented, and can be adapted to the analysis of genomics data by installing Bioconductor v.3.10 [15] packages. The integrated development environment or IDE used was RStudio v.1.2.5033 [75] because it is free for academic use and is actively maintained. The R package shiny v.1.4.0.2 [76] was used for building the application interface and server-side. Data visualization was achieved using the R package ggplot2 [77] and the htmlwidget R packages visNetwork [78] and plotly [79]. The changes in the application were kept track of by using the version controller program git v.2.17.1 [80] and safely stored using Bitbucket [81], a free version control repository hosting service in the cloud. To facilitate the use of the application, an Ubuntu Docker container image was prepared with RStudio and Shiny installed featuring all the necessary Bioconductor [15] and R package dependencies with the correct version. The Docker container image can be found in dockerhub [82] as rstudio-vrepo. A summary of the technologies used is shown in Fig. 4.6.



**Figure 4.6.** Development environment for visualRepo application implementation. A) code repository and software container technologies. B) Experimental and *in silico* tools for data collection and tools for data analysis, exploration, and visualization.

#### 4.2.2 Report objects

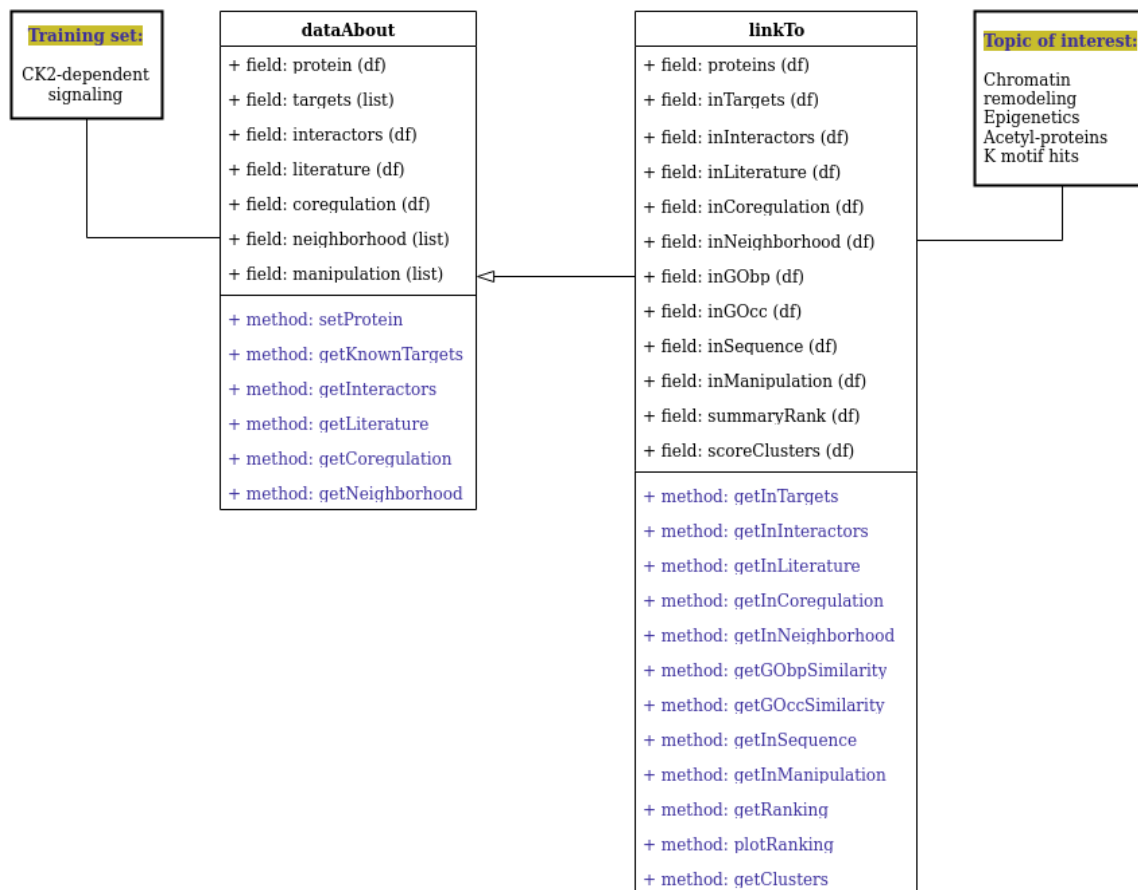
There are several types of biological annotations of interest when performing the functional analysis of PTM data as mentioned in the previous section. Here we organized and classified the information into six groups i.e., proteins, sites, genes, mutations, motif hits, and the COSMIC Cancer Gene Census. Then, we modeled each group into six R classes for storing the information as fields, following an object-oriented programming approach (Fig. 4.7). For each class, we also implemented methods to process and analyze the data to adapt it into a useful tabular or graph format (Fig. 4.7). The fields for each of the groups are summarized in Figure 4.7. The methods and fields of a class can be inherited from other classes using inheritance e.g., the class `motifHitsRepo` contains all the fields and methods of class `sitesRepo` (Fig. 4.7).



**Figure 4.7.** Diagram of classes of reports in the visualRepo application. The fields show the information collected for each category and the methods the steps required to store the information.

We also created two other classes: dataAbout and linkTo, which store information for a given writer of interest and use this information to rank a list of proteins of a topic of interest according to its functional association to the writer, respectively (Fig. 4.8). As an example, data about CK2 such as its substrates, interactors, etc. can be stored in a dataAbout class as specified by its fields. Then this data is used to specify the fields in linkTo using the methods of linkTo. After specifying the fields in linkTo the information can be used to rank the proteins of interest, for example, those involved in chromatin

remodeling, to CK2, in this case (See Chapter 3). The information that can be stored in these classes extended to any writer of interest and to any protein list of interest.



**Figure 4.8.** Diagram of `dataAbout` and `linkTo` classes example of data applied to each class. Training set refers to the information collected to CK2 whereas the topic of interest is the list of proteins or genes of interest to functionally associate to CK2.

### 4.2.3 User Interface

The User Interface (UI) of `visualRepo` is intuitive and simple and consists of a landing page and several tabs in the navigation bar (Fig. 4.9A). The landing page contains general information on the purpose of building the application, images retrieved from work by our group, and some of the main findings described here. It also contains a description of the main biological annotations used in `visualRepo` to perform functional analysis.

The first tab is “Repos” and contains information about the reports and analysis available for exploring and visualizing data in the application under “Overview” (Fig. 4.9B).

Noticeably, the number of reports in visualRepo is not fixed as new data can be added thus making it flexible to incorporating the results of similar analyzes. For visualRepo v.2 the available reports match the analyzes described in the chapters above. The reports available for Chapter 2 are data about CK2, the link of CK2 to acetylated proteins identified by Scholz *et al.* 2015 [4], the link of CK2 to proteome-wide K motif hits, the proteome-wide K motif hits for each position separately, and the sites found in the phospho-acetyl dual enrichment experiments. For Chapter 3 these are the link of CK2 to AmiGO [83] chromatin remodeling, the link of CK2 to Reactome chromatin-modifying enzymes, and the link of CK2 to Reactome epigenetics. These reports can be accessed in tabular format in the same tab under “Data Access” (Fig. 4.9B).

**A**

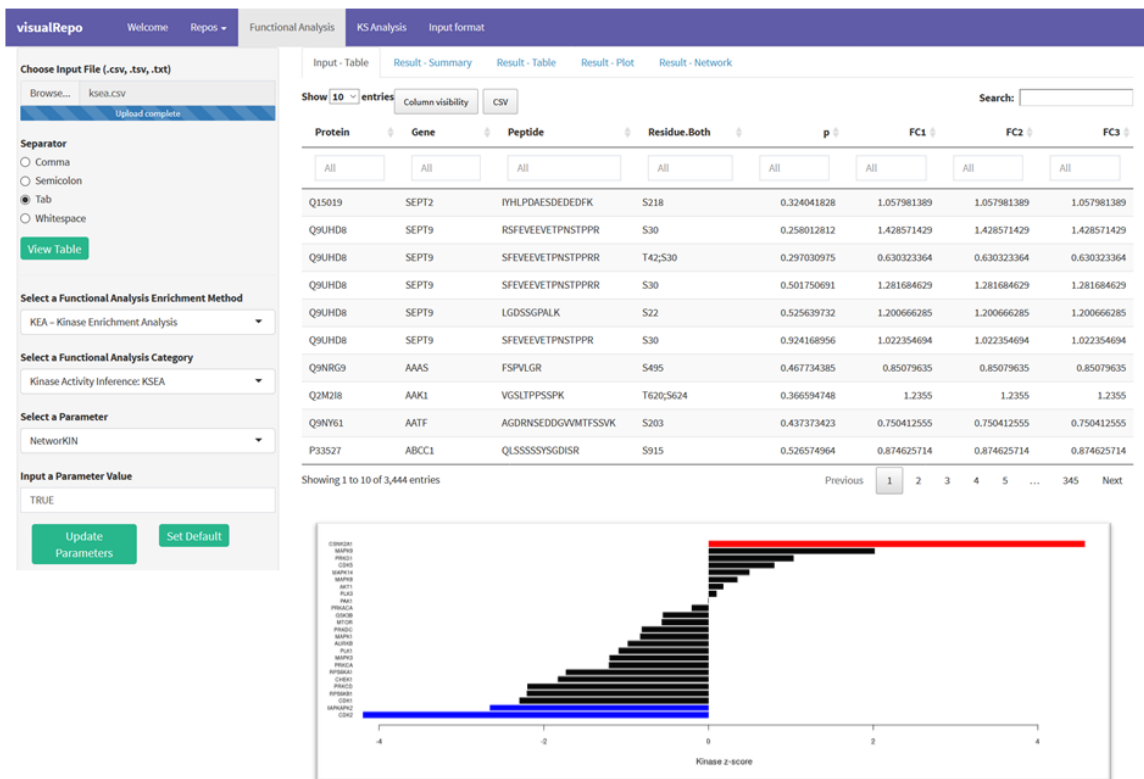
**B**

Data set	Type
CK2- AmiGO chromatin remodeling	Link To
CK2- Reactome chromatin modifying enzymes	Link To
CK2- Reactome epigenetics	Link To
CK2- Acetyl/Phospho Dual Enrichment	Link To
CK2- K motif Hits	Link To
CK2- Acetyl Proteins Scholz et. Al 2015	Link To

Ranking	acc_uniprot	gene_name
1	Q9UGU5	HMGXB4
2	Q8NG31	KNL1
3	A8MT69	CENPX

**Figure 4.9.** visualRepo user interface. A) Landing page and sections of the user interface labeled. B) Repo tab with an Overview tab containing information on the repositories and analysis available and a Data Access tab for accessing the report information in tabular and/or graphic format.

The other two tabs “Functional Analysis” and “KS Analysis” can be used to perform several relevant functional analyses that take as input a list of proteins, phosphorylation sites, and/or relevant quantification data. These were included to perform the most common functional analysis strategies described here. For visualRepo v.2 five types of analysis are available: overrepresentation analysis (ORA), kinase enrichment analysis (KEA, see section 4.1.3), gene set enrichment analysis (GSEA), functional classification, and data set comparison using ORA and functional classification [2,48]. Like the reports the number of analyzes performed by visualRepo is flexible and new analyzes can be added without introducing important changes to the application. An example of tabular and graph output obtained by performing KEA analysis with visualRepo v.2 can be seen in Figure 4.10.

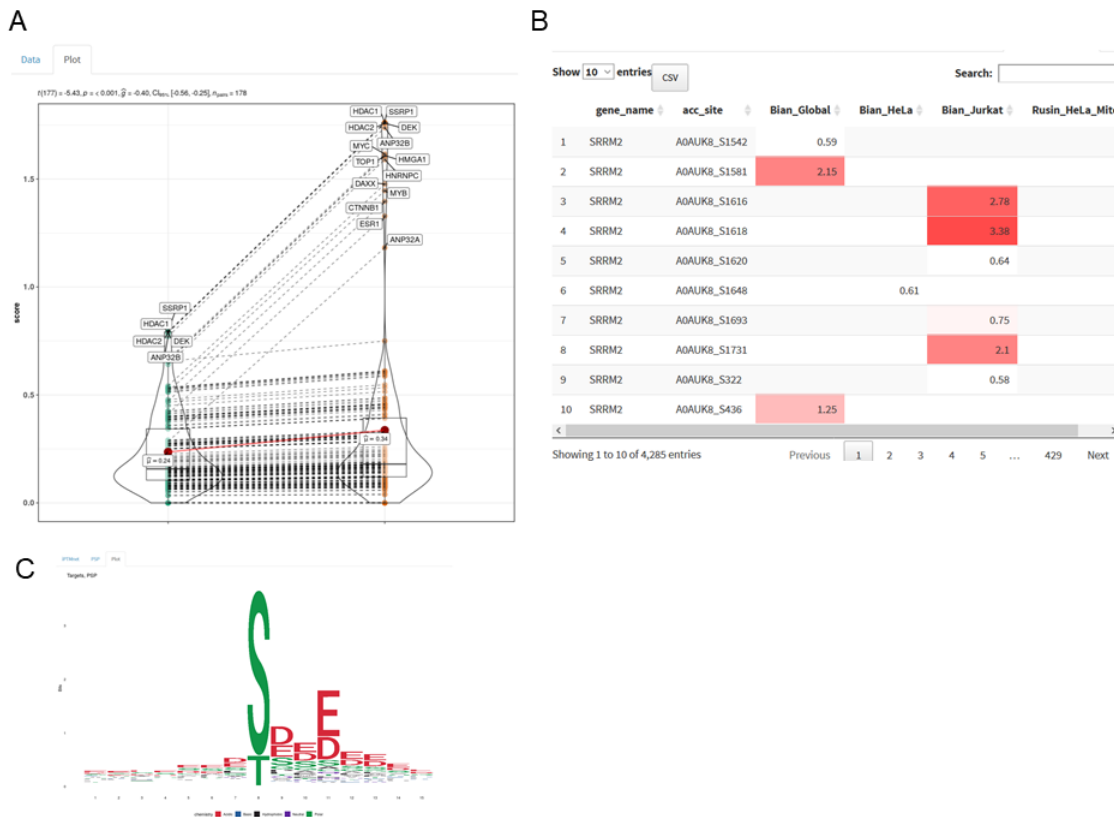


**Figure 4.10.** visualRepo functional analysis example.

#### 4.2.4 Reports

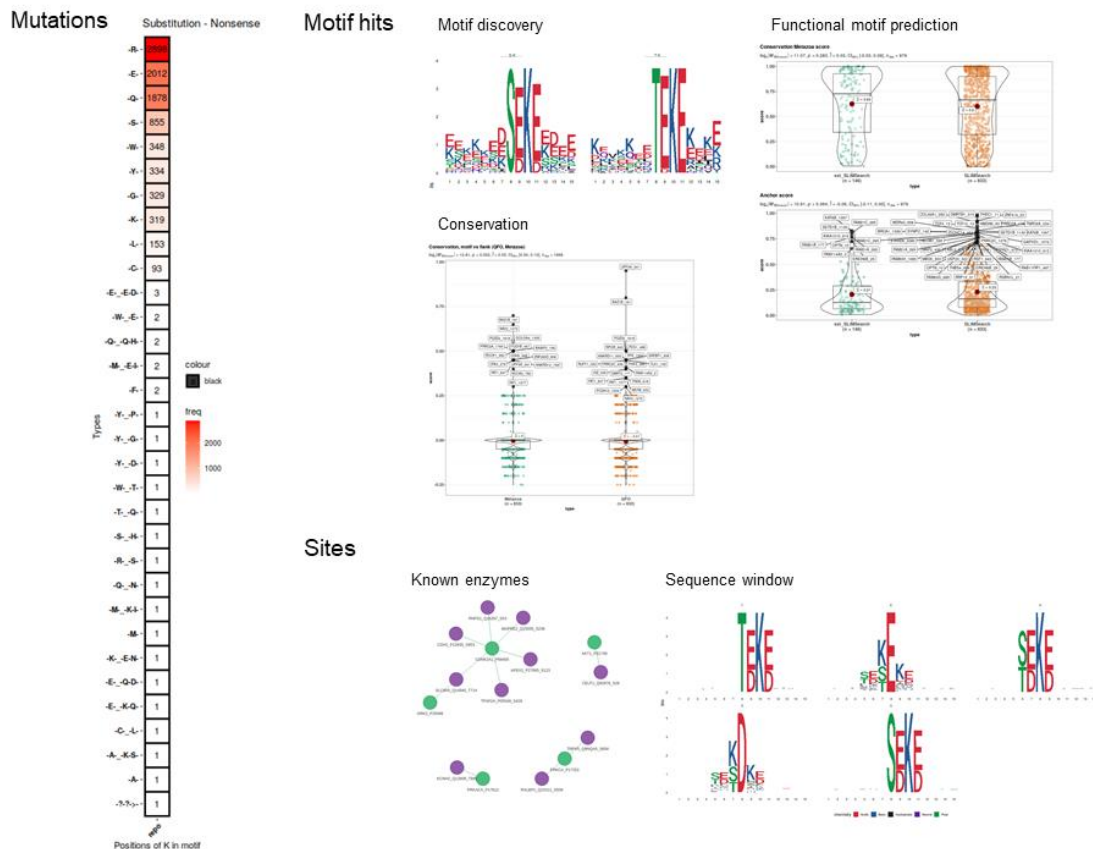
The reports section allows the user to interact with a myriad of tables generated in the analyses performed in Chapters 2 and 3 and to explore the data obtained in tabular and graphic formats. Regarding the latter, the user can also change several predefined values to modify the default parameters of the graphic analyzes. For instance, when browsing the ranking/link of chromatin remodeling proteins to CK2 the user can cluster the results and identify the proteins in the vicinity of a given score (Fig. 4.9B) and/or perform outlier detection with adjustable statistical parameters and the outlier coefficient (Fig. 4.11A). The user can also access quantitative data from CK2 manipulation in a highlight table format where the direction and magnitude of the fold change are represented by a continuous red (up)-to-blue (down) palette (Fig. 4.11B). The sequence logo of CK2 substrates, or in general of the kinase of interest, and of the K motif hits can also be visualized using chemical properties as the color style (Fig. 4.11C).





**Figure 4.11.** Visualization of report data with visualRepo. A) Outliers detection applied to reports on the link of CK2 to protein lists of interest. B) Highlight table of CK2 manipulation data. C) Sequence logo of CK2 substrates in the dataAbout CK2 report.

The advantage of defining R classes storing information of interest, for example, sites, mutations, and motif hits is that it allows us to format any new data set of interest using these. So, to visualRepo, if site information is added to a report or a new report containing this information is added all together, the same tabular and graphic functionalities can still be applied without change. This is important since reports like the proteome-wide K motif hits are complex and require all six groups of information mentioned above. In this case, several graphics are outputted by visualRepo for exploring the information (Fig. 4.12).



**Figure 4.12.** Visualization examples of mutation, motif hit, and site data with visualRepo for K motif hits position +2.

#### 4.2.5 Functional analysis

The “Functional Analysis” and the “KS Analysis” tabs provide access to enrichment analysis, data set comparison, and functional classification utilities. The implementation of these analysis strategies was performed using functions provided in the R libraries for enrichment analysis: clusterProfiler v3.14.3 [48], org.Hs.eg.db v3.10.0 [84], DOSE v3.12.0 [85], msigdb v7.0.1 [86], meshes v1.12.0 [87]; kinase-substrate and kinase activity prediction: KSEAapp v0.99.0 [41], ClueR v1.4 [46], KinSwingR v1.4.0 [42]; pathway analysis: ReactomePA v1.30.0 [88]; and statistical analysis: lsr v0.5 [89], and psych v1.9.12.31 [90]. To start the analysis upload a .csv file containing the list of proteins, sites, and/or site quantification of interest. The input utility recognizes the most common separators besides comma. Sample input files are provided with visualRepo describing the format required for each type of analysis or tool.

#### 4.2.5.1 Enrichment and classification analyses

A total of three enrichment analysis types i.e., ORA, KEA, and GSEA and GO classification can be performed under the “Functional Analysis” tab. ORA was also implemented for data set comparison. The difference between ORA and GSEA was reviewed in [2]; KEA was discussed in section 4.1.3. Briefly, in ORA a list of accession numbers representing a list of entities of biological interest e.g. proteins or phosphoproteins found differentially modulated are searched for overrepresented or enriched terms by using overlap statistics like the hypergeometric distribution for calculating the P value [85,91]. In GSEA, all entities identified in an experiment are considered for enrichment not only those differentially modified since the entities are ranked from high to low based on the quantitative data obtained. For performing ORA analysis, visualRepo applies functions from the clusterProfiler v3.14.3, DOSE v3.12.0, and ReactomePA v1.30.0 packages which rely on a one-sided Fisher’s exact test to determine enrichment. For GSEA the same libraries were used which implement the method described in Subramanian *et al.* 2005 [91]. The proportion of false positives is controlled by FDR. The ORA and GSEA enrichment results are visualized using an enrichment map, organizes enriched terms into a network with edges connecting overlapping gene sets [92], and a ridgeplot, visualizes the expression distributions of core enriched genes for GSEA enriched categories [92], respectively. The ORA results for “compare analysis” are visualized as dotplots [92], depicting the enrichment scores and gene count or ratio by setting the dot color and size, respectively.

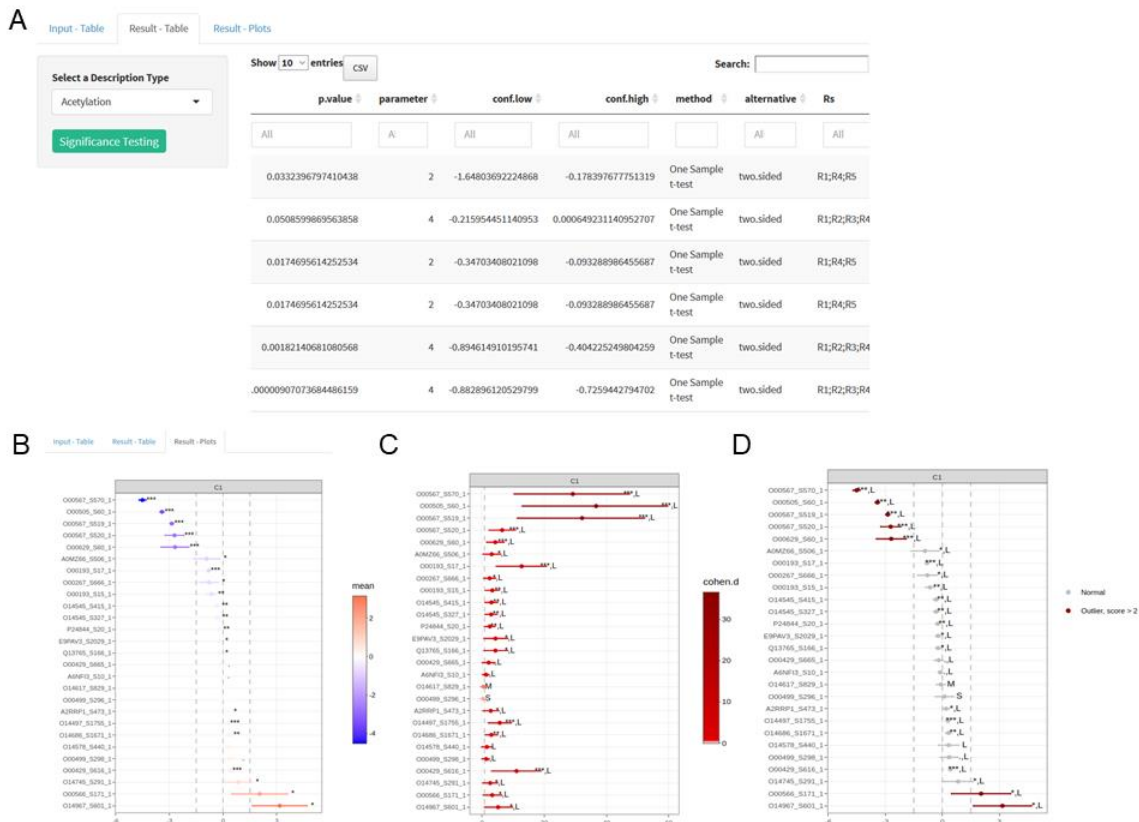
The available biological annotations for performing ORA and GSEA analysis in visualRepo account for the most used including GO, KEGG PATHWAYS, Reactome, Molecular Signature Databases, Wikipathways, Cell Markers, and DisGeNet [48]. Since CK2 data is stored in the dataAbout class the CK2 manipulation data can be extracted and used for ORA analysis as well. This allows the rapid identification of proteins of interest that change in the experiment being analyzed but also when CK2 is manipulated. It also indicates in which past experiments the proteins were observed changed by the name of the annotation, which creates novel hypotheses by providing the overlap with previously generated data by our group.

Finally, the functional classification analysis annotates all the proteins of interest to GO, which is very useful when looking for specific processes, cellular compartments, and molecular functions expected to be targeted. It is also useful because it provides information for single proteins whereas the enrichment analyses are limited to annotations represented, typically, by three or more proteins of interest.

#### 4.2.5.2 Kinase-substrate functional analysis

The “KS Analysis” tab provides several utilities to link a list of proteins and sites to writers. The lists are uploaded to visualRepo as described for the functional analysis. Once uploaded, a list of proteins can be functionally linked to CK2 and/or searched for proteins changing upon CK2 modulation. A list of sites can be also searched for sites that are annotated in the PTM databases PSP and iPTMnet, known kinase-substrates associations, known regulatory roles, and for sites modulated when CK2 is manipulated.

This tab also provides a utility to perform hypothesis testing of sites quantified in a PTM proteomic experiment by selecting a functional category or UniProt keyword of interest (Fig. 4.13). This selection results in a smaller number of comparisons that need to be adjusted for multiple comparisons. Thus when calculating the FDR adjusted P values with the formula  $P \text{ value} \times (n/i)$ , where  $i$  is the rank of the P value being adjusted and  $n$  is the total number of tests [93], the  $n$  value is made smaller which results in smaller adjusted P values and better control on false negatives. After P value calculation and adjustment for the sites belonging to the annotation or keyword of interest e.g., acetylation, phosphorylation, and chromatin remodeling, the results can be visualized (Fig. 4.13B-D). Using the visualization, the sites that increased or decreased above a certain cutoff (Fig. 4.13B) can be identified along with their corresponding size of the effect (Fig. 4.13C), and if they are considered outliers (Fig. 4.13D). For finding the outliers visualRepo implements a robust Z-score where for each site the absolute difference between the log<sub>2</sub>FC value and the median log<sub>2</sub>FC of all sites is divided by the median absolute difference [94]. Any site with a Robust Z-score > 2 is considered an outlier (Fig. 4.13D).



**Figure 4.13.** visualRepo site quantification hypothesis testing. A) Filter by GO annotation or UniProt keywords to perform the tests. B) Log<sub>2</sub>FC order from decreasing to increasing magnitude. C) Effect size. D) Robust Z-score outlier detection.

### 4.3 Concluding remarks

Several complementary strategies can be applied for the functional lists of modified sites observed in PTM proteomic studies. As illustrated here for sites targeted by protein kinase CK2 or differentially modulated by CK2 inhibition. These strategies can be used to expand upon previous working hypotheses (e.g., known consensus sequence or response to inhibition with a chemical probe) or to uncover new trends (e.g., PTM function, PTM interplay, and/or kinase activity in response to perturbations). Equally important, by using R, Bioconductor, and Shiny we integrated several of these tools and methods into the web application visualRepo that allows others to explore, reanalyze and reuse the data generated here. The application can also be adapted to allow for the analysis of other modification types and writer enzymes. As mentioned before, the application visualRepo is available at [www.visualrepo.cloverpath.ca/visualRepo](http://www.visualrepo.cloverpath.ca/visualRepo).

Alternatively, a local copy can be started using the RStudio image provided in the repository `rstudio-vrepo:0.2` available at docker hub along with the code in the Appendix C (see the readme file).

## 4.4 References

1. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13, doi:10.1093/nar/gkn923.
2. Villavicencio-Diaz, T.N.; Rodriguez-Ulloa, A.; Guirola-Cruz, O.; Perez-Riverol, Y. Bioinformatics tools for the functional interpretation of quantitative proteomics results. *Curr. Top. Med. Chem.* **2014**, *14*, 435–49.
3. Francavilla, C.; Lupia, M.; Tsafou, K.; Villa, A.; Kowalczyk, K.; Rakownikow Jersie-Christensen, R.; Bertalot, G.; Confalonieri, S.; Brunak, S.; Jensen, L.J.; et al. Phosphoproteomics of Primary Cells Reveals Druggable Kinase Signatures in Ovarian Cancer. *Cell Rep.* **2017**, *18*, 3242–3256, doi:10.1016/j.celrep.2017.03.015.
4. Schölz, C.; Weinert, B.T.; Wagner, S.A.; Beli, P.; Miyake, Y.; Qi, J.; Jensen, L.J.; Streicher, W.; McCarthy, A.R.; Westwood, N.J.; et al. Acetylation site specificities of lysine deacetylase inhibitors in human cells. *Nat. Biotechnol.* **2015**, *33*, 415–425, doi:10.1038/nbt.3130.
5. Grant, C.E.; Bailey, T.L.; Noble, W.S. FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **2011**, *27*, 1017–1018, doi:10.1093/bioinformatics/btr064.
6. Cheng, A.; Grant, C.E.; Noble, W.S.; Bailey, T.L.; Hancock, J. MoMo: Discovery of statistically significant post-translational modification motifs. *Bioinformatics* **2019**, *35*, 2774–2782, doi:10.1093/bioinformatics/bty1058.
7. Torii, M.; Arighi, C.N.; Li, G.; Wang, Q.; Wu, C.H.; Vijay-Shanker, K. RLIMS-P 2.0: A Generalizable Rule-Based Information Extraction System for Literature Mining of Protein Phosphorylation Information. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **2015**, *12*, 17–29, doi:10.1109/TCBB.2014.2372765.
8. Horn, H.; Schoof, E.M.; Kim, J.; Robin, X.; Miller, M.L.; Diella, F.; Palma, A.; Cesareni, G.; Jensen, L.J.; Linding, R. KinomeXplorer: an integrated platform for kinome biology studies. *Nat. Methods* **2014**, *11*, 603–604.
9. Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J.M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **2015**, *43*, D512–D520, doi:10.1093/nar/gku1267.
10. Ross, K.E.; Huang, H.; Ren, J.; Arighi, C.N.; Li, G.; Tudor, C.O.; Lv, M.; Lee, J.-Y.; Chen, S.-C.; Vijay-Shanker, K.; et al. iPTMnet: Integrative Bioinformatics for Studying PTM Networks. *Methods Mol. Biol.* **2017**, *1558*, 333–353, doi:10.1007/978-1-4939-6783-4\_16.
11. Aranda, B.; Blankenburg, H.; Kerrien, S.; Brinkman, F.S.L.; Ceol, A.; Chautard, E.; Dana, J.M.; De Las Rivas, J.; Dumousseau, M.; Galeota, E.; et al. PSICQUIC and PSISCORE: Accessing and scoring molecular interactions. *Nat. Methods* **2011**, *8*, 528–529.
12. Munk, S.; Refsgaard, J.C.; Olsen, J. V; Jensen, L.J. From Phosphosites to Kinases.

- Methods Mol. Biol.* **2016**, *1355*, 307–321, doi:10.1007/978-1-4939-3049-4\_21.
13. Kustatscher, G.; Grabowski, P.; Schrader, T.A.; Passmore, J.B.; Schrader, M.; Rappsilber, J. Co-regulation map of the human proteome enables identification of protein functions. *Nat. Biotechnol.* **2019**, *37*, 1361–1371, doi:10.1038/s41587-019-0298-5.
  14. Mitchell, A.; Chang, H.-Y.; Daugherty, L.; Fraser, M.; Hunter, S.; Lopez, R.; McAnulla, C.; McMenamin, C.; Nuka, G.; Pesseat, S.; et al. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* **2015**, *43*, D213-21, doi:10.1093/nar/gku1243.
  15. Huber, W.; Carey, V.J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B.S.; Bravo, H.C.; Davis, S.; Gatto, L.; Girke, T.; et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **2015**, *12*, 115–121, doi:10.1038/nmeth.3252.
  16. Beekhof, R.; van Alphen, C.; Henneman, A.A.; Knol, J.C.; Pham, T. V; Rolfs, F.; Labots, M.; Henneberry, E.; Le Large, T.Y.; de Haas, R.R.; et al. INKA, an integrative data analysis pipeline for phosphoproteomic inference of active kinases. *Mol. Syst. Biol.* **2019**, *15*, e8250, doi:10.15252/msb.20188250.
  17. Petsalaki, E.; Helbig, A.O.; Gopal, A.; Pasculescu, A.; Roth, F.P.; Pawson, T. SELPHI: correlation-based identification of kinase-associated networks from global phospho-proteomics data sets. *Nucleic Acids Res.* **2015**, *43*, W276-82, doi:10.1093/nar/gkv459.
  18. Hornbeck, P. V; Kornhauser, J.M.; Tkachev, S.; Zhang, B.; Skrzypek, E.; Murray, B.; Latham, V.; Sullivan, M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **2012**, *40*, D261-70, doi:10.1093/nar/gkr1122.
  19. Crooks, G.E.; Hon, G.; Chandonia, J.-M.; Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res.* **2004**, *14*, 1188–1190, doi:10.1101/gr.849004.
  20. Vacic, V.; Iakoucheva, L.M.; Radivojac, P. Two Sample Logo: a graphical representation of the differences between two sets of sequence alignments. *Bioinformatics* **2006**, *22*, 1536–1537, doi:10.1093/bioinformatics/btl151.
  21. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43*, W39–W49, doi:10.1093/nar/gkv416.
  22. de Castro, E.; Sigrist, C.J.A.; Gattiker, A.; Bulliard, V.; Langendijk-Genevaux, P.S.; Gasteiger, E.; Bairoch, A.; Hulo, N. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* **2006**, *34*, W362-5, doi:10.1093/nar/gkl124.
  23. Rice, P.; Longden, L.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **2000**, *16*, 276–277.
  24. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515, doi:10.1093/nar/gky1049.
  25. Burley, S.K.; Berman, H.M.; Kleywegt, G.J.; Markley, J.L.; Nakamura, H.; Velankar, S. Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. *Methods Mol. Biol.* **2017**, *1607*, 627–641, doi:10.1007/978-1-4939-7000-1\_26.
  26. H. Pagès; Aboyoun, P.; Gentleman, R.; DebRoy, S. Biostrings: Efficient

- manipulation of biological strings. 2019.
27. Carlson, M.; Bioconductor Package Maintainer UniProt.ws: R Interface to UniProt Web Services. 2019.
  28. Meggio, F.; Pinna, L.A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **2003**, *17*, 349–68, doi:10.1096/fj.02-0473rev.
  29. Meggio, F.; Marin, O.; Pinna, L.A. Substrate specificity of protein kinase CK2. *Cell. Mol. Biol. Res.* **1994**, *40*, 401–409.
  30. Bian, Y.; Ye, M.; Wang, C.; Cheng, K.; Song, C.; Dong, M.; Pan, Y.; Qin, H.; Zou, H. Global screening of CK2 kinase substrates by an integrated phosphoproteomics workflow. *Sci. Rep.* **2013**, *3*, 3460, doi:10.1038/srep03460.
  31. St-Denis, N.; Gabriel, M.; Turowec, J.P.; Gloor, G.B.; Li, S.S.C.; Gingras, A.C.; Litchfield, D.W. Systematic investigation of hierarchical phosphorylation by protein kinase CK2. *J. Proteomics* **2015**, *118*, 49–62, doi:10.1016/j.jprot.2014.10.020.
  32. Csizmok, V.; Forman-Kay, J.D. Complex regulatory mechanisms mediated by the interplay of multiple post-translational modifications. *Curr. Opin. Struct. Biol.* **2018**, *48*, 58–67, doi:10.1016/j.sbi.2017.10.013.
  33. Gupta, S.; Stamatoyannopoulos, J.A.; Bailey, T.L.; Noble, W.S. Quantifying similarity between motifs. *Genome Biol.* **2007**, *8*, R24, doi:10.1186/gb-2007-8-2-r24.
  34. Gouw, M.; Sámano-Sánchez, H.; Van Roey, K.; Diella, F.; Gibson, T.J.; Dinkel, H. Exploring Short Linear Motifs Using the ELM Database and Tools. *Curr. Protoc. Bioinforma.* **2017**, *58*, 8.22.1–8.22.35, doi:10.1002/cpbi.26.
  35. Turowec, J.P.; Vilk, G.; Gabriel, M.; Litchfield, D.W. Characterizing the convergence of protein kinase CK2 and caspase-3 reveals isoform-specific phosphorylation of caspase-3 by CK2 $\alpha$ : implications for pathological roles of CK2 in promoting cancer cell survival. *Oncotarget* **2013**, *4*, 560–71, doi:10.18632/oncotarget.948.
  36. Minguez, P.; Letunic, I.; Parca, L.; Bork, P. PTMcode: a database of known and predicted functional associations between post-translational modifications in proteins. *Nucleic Acids Res.* **2013**, *41*, D306–11, doi:10.1093/nar/gks1230.
  37. Gavali, S.; Cowart, J.; Chen, C.; Ross, K.E.; Arighi, C.; Wu, C.H. RESTful API for iPTMnet: a resource for protein post-translational modification network discovery. *Database (Oxford)*. **2020**, *2020*, doi:10.1093/database/baz157.
  38. Li, G.X.H.; Vogel, C.; Choi, H. PTMscape: an open source tool to predict generic post-translational modifications and map modification crosstalk in protein domains and biological processes. *Mol. Omi.* **2018**, *14*, 197–209, doi:10.1039/c8mo00027a.
  39. Cesaro, L.; Marin, O.; Venerando, A.; Donella-Deana, A.; Pinna, L.A. Phosphorylation of cystic fibrosis transmembrane conductance regulator (CFTR) serine-511 by the combined action of tyrosine kinases and CK2: the implication of tyrosine-512 and phenylalanine-508. *Amino Acids* **2013**, *45*, 1423–1429, doi:10.1007/s00726-013-1613-y.
  40. Venerando, A.; Cesaro, L.; Marin, O.; Donella-Deana, A.; Pinna, L.A. A “SYDE” effect of hierarchical phosphorylation: possible relevance to the cystic fibrosis basic defect. *Cell. Mol. Life Sci.* **2014**, *71*, 2193–2196, doi:10.1007/s00018-014-



- 1581-8.
41. Wiredja, D.D.; Koyutürk, M.; Chance, M.R. The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics* **2017**, *33*, 3489–3491, doi:10.1093/bioinformatics/btx415.
  42. Engholm-Keller, K.; Waardenberg, A.J.; Müller, J.A.; Wark, J.R.; Fernando, R.N.; Arthur, J.W.; Robinson, P.J.; Dietrich, D.; Schoch, S.; Graham, M.E. The temporal profile of activity-dependent presynaptic phospho-signalling reveals long-lasting patterns of poststimulus regulation. *PLoS Biol.* **2019**, *17*, e3000170, doi:10.1371/journal.pbio.3000170.
  43. Weidner, C.; Fischer, C.; Sauer, S. PHOXTRACK—a tool for interpreting comprehensive datasets of post-translational modifications of proteins. *Bioinformatics* **2014**, *30*, 3410–3411, doi:10.1093/bioinformatics/btu572.
  44. Lachmann, A.; Ma'ayan, A. KEA: kinase enrichment analysis. *Bioinformatics* **2009**, *25*, 684–686, doi:10.1093/bioinformatics/btp026.
  45. Yang, P.; Patrick, E.; Humphrey, S.J.; Ghazanfar, S.; James, D.E.; Jothi, R.; Yang, J.Y.H. KinasePA: Phosphoproteomics data annotation using hypothesis driven kinase perturbation analysis. *Proteomics* **2016**, *16*, 1868–1871, doi:10.1002/pmic.201600068.
  46. Yang, P.; Zheng, X.; Jayaswal, V.; Hu, G.; Yang, J.Y.H.; Jothi, R. Knowledge-Based Analysis for Detecting Key Signaling Events from Time-Series Phosphoproteomics Data. *PLoS Comput. Biol.* **2015**, *11*, e1004403, doi:10.1371/journal.pcbi.1004403.
  47. Wilkes, E.H.; Casado, P.; Rajeeve, V.; Cutillas, P.R. Kinase activity ranking using phosphoproteomics data (KARP) quantifies the contribution of protein kinases to the regulation of cell viability. *Mol. Cell. Proteomics* **2017**, *16*, 1694–1704, doi:10.1074/mcp.O116.064360.
  48. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **2012**, *16*, 284–287, doi:10.1089/omi.2011.0118.
  49. Rigbolt, K.T.G.; Prokhorova, T.A.; Akimov, V.; Henningsen, J.; Johansen, P.T.; Kratchmarova, I.; Kassem, M.; Mann, M.; Olsen, J. V; Blagoev, B. System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. *Sci. Signal.* **2011**, *4*, rs3, doi:10.1126/scisignal.2001570.
  50. Xue, Y.; Liu, Z.; Cao, J.; Ma, Q.; Gao, X.; Wang, Q.; Jin, C.; Zhou, Y.; Wen, L.; Ren, J. GPS 2.1: enhanced prediction of kinase-specific phosphorylation sites with an algorithm of motif length selection. *Protein Eng. Des. Sel.* **2011**, *24*, 255–260, doi:10.1093/protein/gzq094.
  51. Song, C.; Ye, M.; Liu, Z.; Cheng, H.; Jiang, X.; Han, G.; Songyang, Z.; Tan, Y.; Wang, H.; Ren, J.; et al. Systematic analysis of protein phosphorylation networks from phosphoproteomic data. *Mol. Cell. Proteomics* **2012**, *11*, 1070–1083, doi:10.1074/mcp.M111.012625.
  52. Gao, J.; Thelen, J.J.; Dunker, A.K.; Xu, D. Musite, a tool for global prediction of general and kinase-specific phosphorylation sites. *Mol. Cell. Proteomics* **2010**, *9*, 2586–2600, doi:10.1074/mcp.M110.001388.
  53. Huang, H.-D.; Lee, T.-Y.; Tzeng, S.-W.; Horng, J.-T. KinasePhos: a web tool for

- identifying protein kinase-specific phosphorylation sites. *Nucleic Acids Res.* **2005**, *33*, W226-9, doi:10.1093/nar/gki471.
54. Goel, R.; Harsha, H.C.; Pandey, A.; Prasad, T.S.K. Human Protein Reference Database and Human Proteinpedia as resources for phosphoproteome analysis. *Mol. Biosyst.* **2012**, *8*, 453–463, doi:10.1039/c1mb05340j.
  55. Wong, Y.-H.; Lee, T.-Y.; Liang, H.-K.; Huang, C.-M.; Wang, T.-Y.; Yang, Y.-H.; Chu, C.-H.; Huang, H.-D.; Ko, M.-T.; Hwang, J.-K. 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, doi:10.1093/nar/gkm322.
  56. Ayati, M.; Wiredja, D.; Schlatzer, D.; Maxwell, S.; Li, M.; Koyutürk, M.; Chance, M.R. CoPhosK: A method for comprehensive kinase substrate annotation using co-phosphorylation analysis. *PLoS Comput. Biol.* **2019**, *15*, e1006678, doi:10.1371/journal.pcbi.1006678.
  57. Yang, P.; Humphrey, S.J.; James, D.E.; Yang, Y.H.; Jothi, R. Positive-unlabeled ensemble learning for kinase substrate prediction from dynamic phosphoproteomics data. *Bioinformatics* **2016**, *32*, 252–259, doi:10.1093/bioinformatics/btv550.
  58. Patrick, R.; Horin, C.; Kobe, B.; Cao, K.-A.L.; Bodén, M. Prediction of kinase-specific phosphorylation sites through an integrative model of protein context and sequence. *Biochim. Biophys. Acta* **2016**, *1864*, 1599–1608, doi:10.1016/j.bbapap.2016.08.001.
  59. Zou, L.; Wang, M.; Shen, Y.; Liao, J.; Li, A.; Wang, M. PKIS: computational identification of protein kinases for experimentally discovered protein phosphorylation sites. *BMC Bioinformatics* **2013**, *14*, 247, doi:10.1186/1471-2105-14-247.
  60. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.; Roth, A.; Bork, P.; et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **2016**, gkw937, doi:10.1093/nar/gkw937.
  61. Singh, S.A.; Winter, D.; Kirchner, M.; Chauhan, R.; Ahmed, S.; Ozlu, N.; Tzur, A.; Steen, J.A.; Steen, H. Co-regulation proteomics reveals substrates and mechanisms of APC/C-dependent degradation. *EMBO J.* **2014**, *33*, 385–399, doi:10.1002/embj.201385876.
  62. Yu, G.; Li, F.; Qin, Y.; Bo, X.; Wu, Y.; Wang, S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* **2010**, *26*, 976–8, doi:10.1093/bioinformatics/btq064.
  63. Schöfl, G. reutils: Talk to the NCBI EUtils. 2016.
  64. Bardou, P.; Mariette, J.; Escudié, F.; Djemiel, C.; Klopp, C. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **2014**, *15*, 293, doi:10.1186/1471-2105-15-293.
  65. Bonetta, L. Protein-protein interactions: Interactome under construction. *Nature* **2010**, *468*, 851–854, doi:10.1038/468851a.
  66. Perkins, J.R.; Diboun, I.; Dessailly, B.H.; Lees, J.G.; Orengo, C. Transient protein-protein interactions: structural, functional, and network properties. *Structure* **2010**, *18*, 1233–1243, doi:10.1016/j.str.2010.08.007.

67. Chatr-aryamontri, A.; Oughtred, R.; Boucher, L.; Rust, J.; Chang, C.; Kolas, N.K.; O'Donnell, L.; Oster, S.; Theesfeld, C.; Sellam, A.; et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **2016**, gkw1102, doi:10.1093/nar/gkw1102.
68. Bader, G.D.; Cary, M.P.; Sander, C. Pathguide: a pathway resource list. *Nucleic Acids Res.* **2006**, *34*, doi:10.1093/nar/gkj126.
69. Orchard, S.; Kerrien, S.; Abbani, S.; Aranda, B.; Bhate, J.; Bidwell, S.; Bridge, A.; Briganti, L.; Brinkman, F.S.L.; Cesareni, G.; et al. Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. *Nat. Methods* **2012**, *9*, 345–350, doi:10.1038/nmeth.1931.
70. Shannon, P. PSICQUIC: Proteomics Standard Initiative Common Query InterfaCe. 2019.
71. de Villavicencio-Diaz, T.N.; Rabalski, A.J.; Litchfield, D.W. Protein kinase CK2: Intricate relationships within regulatory cellular networks. *Pharmaceuticals* 2017.
72. UniProt Consortium, T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2018**, *46*, 2699, doi:10.1093/nar/gky092.
73. Beltrao, P.; Albanèse, V.; Kenner, L.R.; Swaney, D.L.; Burlingame, A.; Villén, J.; Lim, W.A.; Fraser, J.S.; Frydman, J.; Krogan, N.J. Systematic functional prioritization of protein posttranslational modifications. *Cell* **2012**, *150*, 413–425, doi:10.1016/j.cell.2012.05.036.
74. R: What is R? Available online: <https://www.r-project.org/about.html> (accessed on Jul 15, 2020).
75. RStudio | Open source & professional software for data science teams - RStudio Available online: <https://rstudio.com/> (accessed on Jul 15, 2020).
76. Chang, W.; Cheng, J.; Allaire, J.; Xie, Y.; McPherson, J. shiny: Web Application Framework for R. 2020.
77. Wickham, H. *ggplot2 -Positioning Elegant Graphics for Data Analysis*; 2016; ISBN 978-0-387-98140-6.
78. B.V., A.; Thieurmel, B.; Robert, T. visNetwork: Network Visualization using “vis.js” Library. 2019.
79. Sievert., C. Interactive Web-Based Data Visualization with R, plotly, and shiny. 2020.
80. Git Available online: <https://git-scm.com/> (accessed on Jul 15, 2020).
81. Bitbucket | The Git solution for professional teams Available online: <https://bitbucket.org/product/> (accessed on Jul 15, 2020).
82. Docker Hub Available online: <https://hub.docker.com/> (accessed on Jul 15, 2020).
83. Munoz-Torres, M.; Carbon, S. Get GO! retrieving GO data using AmiGO, QuickGO, API, files, and tools. In *Methods in Molecular Biology*; Humana Press Inc., 2017; Vol. 1446, pp. 149–160.
84. Carlson, M. org.Hs.eg.db: Genome wide annotation for Human. 2019.
85. Yu, G.; Wang, L.-G.; Yan, G.-R.; He, Q.-Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **2015**, *31*, 608–609, doi:10.1093/bioinformatics/btu684.
86. Dolgalev, I. msgdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format. 2019.
87. Yu, G. Using meshes for MeSH term enrichment and semantic analyses.

- Bioinformatics* **2018**, *34*, 3766–3767, doi:10.1093/bioinformatics/bty410.
88. Yu, G.; He, Q.-Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* **2016**, *12*, 477–9, doi:10.1039/c5mb00663e.
  89. Navarro, D.J. Learning statistics with R: A tutorial for psychology students and other beginners. 2015.
  90. Revelle, W. psych: Procedures for Personality and Psychological Research, Northwestern University. 2019.
  91. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15545–15550, doi:10.1073/pnas.0506580102.
  92. Chapter 12 Visualization of Functional Enrichment Result | clusterProfiler: universal enrichment tool for functional and comparative study Available online: <https://yulab-smu.github.io/clusterProfiler-book/chapter12.html#ridgeline-plot-for-expression-distribution-of-gsea-result> (accessed on Jul 15, 2020).
  93. McDonald, J.H. *HANDBOOK OF BIOLOGICAL STATISTICS SECOND EDITION*;
  94. Lipinski, S.; Grabe, N.; Jacobs, G.; Billmann-Born, S.; Till, A.; Häslér, R.; Aden, K.; Paulsen, M.; Arlt, A.; Kraemer, L.; et al. RNAi screening identifies mediators of NOD2 signaling: implications for spatial specificity of MDP recognition. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 21426–21431, doi:10.1073/pnas.1209673109.

## 4.5 Supplemental Material

### Tables

**Supp. Table 4.1.** CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019.

Sheet	Name	Content
1	Table 1	PhosphoSitePlus CK2 (CK2a1) substrates, September 2019.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.2.** CK2 target sequences motif discovery using MoMo v.5.0.5.

Sheet	Name	Content
1	Table 2	MoMo motif discovery applied to CK2 substrates. MoMo input and output. MoMo motifs.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.3.** FIMO v.5.0.5 proteome scan for motifs discovered with MoMo v.5.0.5.

Sheet	Name	Content
1	Table 3	FIMO proteome search using MoMo output motifs as input.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.4.** SDSE hits that are annotated as phosphorylated in PhosphoSitePlus (PSP) database.

Sheet	Name	Content
1	Table 4	PhosphoSitePlus, sequence or motif: xxxxxxxSDSExxxx. Only phosphorylation sites (December 2019).
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.5.** PTM types annotated in the vicinity ( $\pm 7$  amino acids) of phosphorylated SDSE hits.

Sheet	Name	Content
1	Table5	PhosphoSitePlus, sequence or motif: x{7}SDSEX{4}. Nearby PTM sites to SDSE (December 2019).
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.6.** SDSE hits known kinase-substrate relationships and phosphorylation *in vitro* by CK2.

Sheet	Name	Content
1	Table 6	SDSE hits: known kinase substrate-relationships, iPTMnet and PSP databases (December 2019). SDSE hits: <i>in vitro</i> CK2 phosphorylation (Bian <i>et al.</i> 2013).
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.7.** Kinase set enrichment analysis and activity inference.

Sheet	Name	Content
1	Table7	Kinase set enrichment analysis and activity inference. KSEA (Web, R package), KinSwingR (R package), CLUster Evaluation (Clue) (R package)
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.8.** CoPhosK, iGPS, and NetworKIN prediction using as input CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019.

Sheet	Name	Content
1	Table 8	CoPhosK prediction using as input CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019. iGPS prediction using as input CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019. KSP-PUEL prediction using as input CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019. NetworKIN prediction using as input CK2 substrates retrieved from PhosphoSitePlus (PSP) database, September 2019.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.9.** Co-regulated proteins of CK2 subunits retrieved from ProteomeHD.

Sheet	Name	Content
1	Table 9	proteomehd.net download. It contains proteins co-regulated with your query (above selected score cut-off). You query was: P68400. Your selected score cut-off was: 0.995. You query was: P19784. Your selected score cut-off was: 0.95. You query was: P67870. Your selected score cut-off was: 0.995.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.10.** Co-regulated proteins of CK2 subunits that are known CK2 interactors, June 2020.

Sheet	Name	Content
1	Table 10	Co-regulated proteins that are CK2 interactors by subunit, June 2020.

		CSNK2A1: co-regulated proteins as interactors. CSNK2A2: co-regulated proteins as interactors. CSNK2B: co-regulated proteins as interactors. Overlap.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.11.** Regulatory CK2 target sites and domains represented, September 2019.

Sheet	Name	Content
1	Table 11	Regulatory CK2 target sites retrieved from PSP September 2019. Domain containing regulatory CK2 target sites.
Appendix A		Supplemental Material – Chapter 4

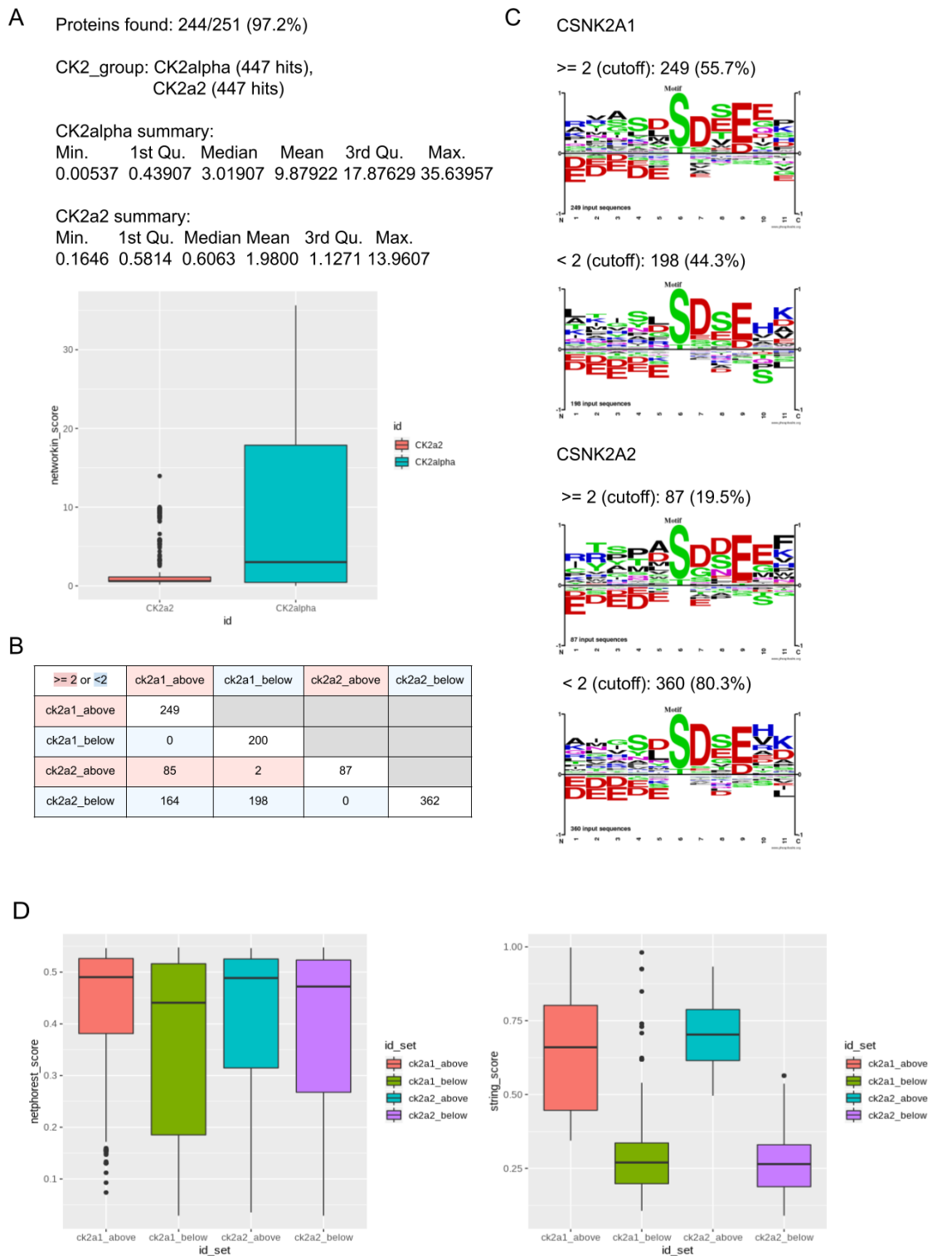
**Supp. Table 4.12.** Summary of strategies and tools discussed in Chapter 4.

Sheet	Name	Content
1	Table 12	Summary of strategies and tools discussed in Chapter 4.
Appendix A		Supplemental Material – Chapter 4

**Supp. Table 4.13.** Network representation of tools and databases included in the strategies that are co-mentioned in the Scholarly literature.

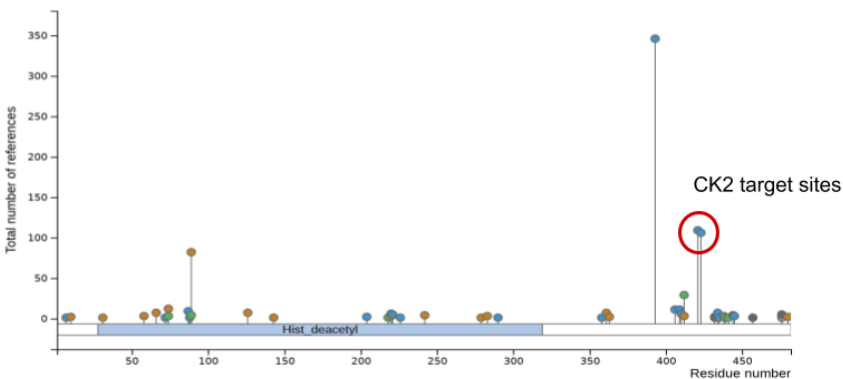
Sheet	Name	Content
1	Table 13	Network of tools and databases included in the strategies that are co-mentioned in the Scholarly literature.
Appendix A		Supplemental Material – Chapter 4

## Figures



**Supp. Figure 4.1.** Kinase-substrate prediction for CK2 sites. A) Number of sites found for CSNK2A1 and CSNK2A2 (overlap: 100%). Summary of networkin\_score. B) Number of sequences belonging to each subset above and below the cutoff. C) Sequence logos of sequences above and below the cut-off.

**PhosphoSitePlus (domain)**



**InterPro (domain)**



**PTMfunc**

PTMfunc					Documentation Statistics Contact			
22	421	S	phos	18	Regulation	Log2 Regulation	Condition	Pubmed ID
23	423	S	phos	19				
24	432	K	ace	9				
25	438	K	ace	9				
26	439	K	ace	9	Conservation	Species	Spectral Counts	
22	421	S	phos	18				
23	423	S	phos	19				
24	432	K	ace	9	M.musculus	214		

**PhosphoSitePlus (PTM effect)**

- Effects on Modified Protein**
- activity, induced
  - enzymatic activity, induced
  - molecular association, regulation
- Triggered By**
- S421-p
  - S421-p , S423-p
  - S421-p , S423-p
- Effects on Biological Processes**
- cell adhesion, induced
  - cell differentiation, induced
  - cytoskeletal reorganization
- Triggered By**
- S421-p
  - S421-p
  - S421-p

**Supp. Figure 4.2.** Motif and domain information for the CK2 targets sites in HDAC1, Ser421, and Ser423. PhosphoSitePlus (PSP) information shows no domain towards the C-terminal where the sites are located. InterPro analysis points to a 23.4% of disorder prediction for the whole protein occurring in the C-terminal. PTMfunc points to the regulation of this modification in cell cycle and in human embryonic stem cell differentiation. PSP information on PTM regulatory effects indicates the role of these sites in the induction of HDAC1 activity and the regulation of HDAC1 interaction.



## Chapter 5

### 5 Discussion

Understanding the signaling networks involved in controlling cell proliferation and survival in normal and cancer cells is necessary for devising novel and effective therapies. Recent studies following a systems biology approach have yielded important insights in this field, highlighting the involvement of the clinically relevant protein kinase CK2 [1,2] in hierarchical phosphorylation [3] and in modulating cell survival by attenuating caspase pathways through phosphorylation of residues adjacent to the cleavage sites [4–6]. These results suggest that CK2 target sites are a breeding ground for post-translational modifications (PTM) interplay somewhat like the histone tails. Contrary to most kinases, CK2 is catalytically competent independent of effectors or phosphorylation by other kinases and is constitutively expressed in cells [7]. In correspondence and considering the hundreds of hits in the human proteome matching the minimal CK2 consensus annotated phosphorylated in databases [8] or showing up in phosphoproteomic experiments [9] a better understanding of CK2 specificity is crucial for identifying true substrates.

CK2 target sites conform to the minimal acidophilic motif [ST]XX[DE] [8], where the negative charge at position +3 downstream of the phosphorylatable residue can also be mimicked by phosphorylated Ser or Tyr residues [8,10]. Furthermore, Pro and positive amino acids are thought to be negative determinants at +1 and +1-+3 positions, respectively [10]. Additional specificity determinants likely exist in the cellular context [7] for instance regulation of specificity at the substrate level by the CK2 regulatory subunit [11] as the minimal consensus sequence can be considered degenerate [12]. Additionally, post-translational modifications occurring in the vicinity of the target sequence like lysine acetylation can change the local electrostatic properties and modulate the effect of lysine as a negative determinant. Thus, the goal of this thesis is to expand this view and explore the role of lysine and acetyllysine as determinants of CK2-dependent phosphorylation at these positions and to identify potential candidate substrates containing either determinant. The role of CK2 in chromatin organization is also summarized considering the involvement of acetylation and phosphorylation

networks in this process [13] and the regulatory role of CK2 on lysine acetylation erasers [14]. In addition, the functional analysis steps performed were documented, integrating the analytical workflow into an R Shiny data-driven application where the information generated can be accessed, reused, and generated *de novo*; the latter is ideal for applying the same logic to other enzymes where interplay at the substrate level is suspected.

## 5.1 Summary of findings and research impact

### 5.1.1 Intricate relationships of CK2 with lysine acetylation networks

Here we determined that several PTM types can target the vicinity of known CK2 target sites including sites modified by lysine acetylation [15], a few of which were predicted to interplay with the CK2 sites. This hinted to a functional relevance for the interplay between CK2-dependent phosphorylation and lysine acetylation given that acetylation removes a positive charge from lysine changing the local properties of the target sequences [16] which then could be more permissive to CK2. Follow-up experimental and *in silico* analysis performed by us showed for the first time that lysine and acetyllysine are tolerated by CK2 at position +2 downstream the phosphorylatable residue in the context of the linear pattern [ST][DE]X[DE]. This position is more accommodating to the presence of non-acidic residues [8] and when acetyllysine is present *in vitro* phosphorylation was rescued to 70 % of that of the CK2 peptide substrate used as control compared to 25 % when lysine is present. These findings provide a benchmark for assessing and exploring CK2 specificity and the interplay between CK2 and acetyllysine sites in future experiments but also in the human proteome, the data annotated in PTM databases, the literature, and in published or accessible phosphoproteomic and acetylome data sets. After approaching the latter, we effectively observed that for curated CK2 target sites, sites phosphorylated *in vitro* by CK2, and/or sites modulated by treatment with the ATP-competitive CK2 inhibitor CX-4945 phosphorylation by CK2 may occur at [ST][DE]X[DE] when lysine is present at position +2. Phosphorylation by CK2 when lysine is present either at position +1 or +3 was observed when at least four negatively charged amino acids were present. This allowed us to generate a list of *in vitro* and *in vivo* CK2 candidates containing lysine at position +2. As expected, based on our initial analysis of CK2 substrates several of the candidate

sites were also observed to be targeted by lysine acetylation at positions +1-+3. This is in accordance with our motif discovery findings in which *in vivo* acetylation sites identified in [17] were observed to target negative motifs in proteins. More importantly, we determined that treatment with pan-lysine deacetylase inhibitors, i.e., inhibitors of acetylation erasers that increase protein acetylation in the cell [18], modulate CK2-dependent signaling. Next by looking *in vivo* at the fraction of the phosphoproteome that is acetylated by performing dual enrichment followed by mass spectrometry we observed for the first time a number of phosphosites matching the CK2 consensus containing an acetyllysine residue in their vicinity which in several instances was found to target a lysine residue at position +2.

### 5.1.2 CK2-dependent regulation and chromatin organization

Among the candidate CK2 sites containing lysine at position +2 and the co-occurring phospho-acetyl sites proteins involved in chromatin remodeling were observed. Correspondingly, the role of CK2 in this process was investigated and summarized with known and novel associations retrieved. For example, a known regulatory role in histone H3 and H4 phosphorylation [19,20], heterochromatin DNA methylation patterns [21], chromatin-remodeling [22], and transcriptional repression [23] and a potential novel role in regulating the function of the repressor complex PRC2 and the lysine acetyltransferase KAT7. This provides a biological context in which widely used tools for exploring chromatin and its associations with signaling pathways can be applied to understand the regulatory role of CK2. Beyond their association to CK2 a wealth of other phosphorylation sites occurring in the vicinity of lysine acetylation sites or vice versa was observed. This challenges to a point the assumption that PTMs are mutually exclusive [24] and opens the opportunity for further validation of the occurrence of PTM interplay at specific residues of attractive molecular targets such as HIF1A [25] and several histone methyltransferases, kinases, and lysine-acetylated histone binders. It also points to the importance of developing quality antibodies and proteomic-based mass spectrometry methods to account for the presence of multiple modification sites at the region of interest of the protein.

### 5.1.3 Proteome-wide distribution of CK2 sites containing lysine residues

Finally, based on scans of the human proteome against linear patterns for CK2 sites containing lysine residues at positions +1-+3 it was determined that there are likely other functionally relevant sites not included in the candidate list, but additional information is required. This is illustrated by the cellular and boundary properties [26] of the proteome hits observed for positions +1-+3. For example, several of the candidate sites were identified conserved across different species in metazoan and orthologs, the residues matching the motifs were more conserved than the flanks, and/or they are targeted by mutations. The observation of mutations targeting either the negative residues at positions +1 and +3 or the lysine residue at position +2 suggests that in addition to lysine acetylation this can be another “switch” to modulate CK2 specificity at the substrate level; highlighting the intricacy of the genetic and modification networks at play at these motifs in the cell.

### 5.1.4 visualRepo web application implementation for data processing and analysis

Our study involved the use of different logical steps to process and analyze large data sets including the integration of several bioinformatic tools, the literature, and biological databases [27]. Here we provide a web application implementation using R Shiny of the analytical workflow followed i.e., visualRepo v2. This application allows the user to perform main functional analysis strategies using R and CRAN and Bioconductor [28] packages for the functional analysis of lists of modified sites and proteins. These strategies were summarized and exemplified in the context of CK2 data sets. In addition, visualRepo v2 allows for access to several data sets of the CK2 candidates containing lysine at positions +1-+3 and of the co-occurring phospho-acetyl sites. Some of the functionalities provided by visualRepo v2 are enrichment analysis (such as GO, pathways, CK2 manipulation data, kinase), kinase activity prediction, GO classification, and data set comparison. It also allows to find in a list of modified sites the known modification sites, the known regulatory sites, the known enzymes-substrate associations, the known modulation by CK2 manipulation, the known PTM interplay, the position of

sites of interest inside proteins, and the formatting to sequence window for weblogo analysis. For sites quantification data visualRepo v2 performs hypothesis testing, outlier detection, and effect size analysis to identify differentially modulated sites. Importantly, visualRepo v2 adds the capability to filter the dataset before hypothesis testing by a keyword of interest such as acetylation of histone modification, this allows for adjusting only the p values in the functional category of interest. Regarding protein lists, visualRepo v2 finds known proteins that change upon CK2 modulation, the complexes in which they are involved, and their ranking according to their functional similarity to CK2. This last functionality utilizes different biological data sources such as kinase-substrate relationships, molecular interaction, mentions in the literature, and co-regulation allowing fast identification/ranking of proteins of interest in large lists that have previous functional information linking them to CK2. Several text and graphical reports can be downloaded from visualRepo v2. Because visualRepo v2 was dockerized it can be run using a docker container with all required R dependencies. Overall, visualRepo v2 provides a way to easily perform valuable functional analysis essential to uncover PTM interplay and enzyme-substrate functional associations not only for CK2 but for other kinase/enzyme or a list of sites of interest.

### 5.1.5 Limitations

The limitations associated with our study relates to how mass spectrometry-based proteomic techniques and computational proteomic methods assess the identification of multiply-modified peptides. In the future the costs of the experiments performed may be lower and the identification methods less stringent with the identification scores and the false discovery rate testing. In addition, antibodies that recognize multiple modified epitopes are not common or are more specific for phosphorylation. Nevertheless, here we bypassed some of the limitations by incorporating the processing and analysis of a wealth of data from diverse sources and used only the sites that have been previously identified and annotated in the databases or the literature. One can argue that the databases are often incomplete; therefore, we also included PTM information from relevant studies in the literature or obtained by our group to provide a biological and experimental context to a subset of the sites analyzed.

## 5.2 Future directions

Several *in silico* and wet experiments can be derived from the findings presented here in the future. Most importantly, we should continue to monitor the literature and the phosphoproteomic data sets generated by our group to expand and refine the list of substrates containing lysine at positions +1-+3. In this regard, it will be of interest to profile the phosphoproteome represented by a mixture of phosphopeptide generated with other proteases besides Trypsin such as Asp-N and their combination. Another variable that can be changed is the cell type used for phosphoproteomic profiling which can be expanded from U2OS and HeLa cells to leukemia, lung, and prostate cancer cells where CK2 is an attractive molecular candidate [29]. This can be achieved in part by monitoring the scholarly literature for new phosphoproteomic datasets on the modulation of CK2-dependent signaling and documenting the experimental set up. Once processed this information can be added to visualRepo v2 specially to be used with the CK2 manipulation enrichment functionality. Likewise, libraries complementary to the ones already implemented in visualRepo should be monitored and included if possible, to extend the functionalities; any step that is performed manually should be automated whenever a new library created by others or our group should be made available. Additional efforts can be done to generate antibodies against a few of the candidate sites which can be of use for immunoprecipitation and immunoblot experiments. In this case our group has been actively developing CK2 specific antibodies and proteins like BAZ1B are candidates. Finally, dual enrichment experiments should be further optimized to increase the yield of modified peptides and expand to assess other lysine deacetylase inhibitors or dual CK2-lysine deacetylase inhibitors across different cell types. The obtained raw files should also be searched using up-to-date computational proteomic strategies as this is a rapidly evolving area of bioinformatics research.

## 5.3 Concluding remarks

Overall, this thesis approaches the study of the modulation of protein kinase CK2 specificity by neighboring PTMs expressly lysine acetylation. It demonstrates that the overlap of phosphosites and acetylation networks in the cell are extensive and that they converge at several sites including known and tentative CK2 candidate sites *in vivo*. This

highlights the importance of considering kinase consensus sequences as breeding grounds for interplay and the potential of missing biologically relevant substrates if we fail to expand our study to include multiple modified proteoforms in signaling networks. To perform this intricate analysis requires the integration of many large data sets and thus here we exemplify the importance of documenting the analytical workflow and developing it into an application that can be used to maintain reproducibility and facilitate reuse of the data. By combining experimental and *in silico* analysis we also demonstrated that some technical limitations can be overcome by performing prediction and/or reuse existing data. Altogether, our results bring us closer to understanding the regulatory role of CK2 and its overlap with lysine acetylation networks at the substrate level but also to a potential role in chromatin organization. This is of great importance to the community as the first steps are being made to device dual CK2 and lysine deacetylase inhibitors and chromatin organizations is a cornerstone biological process.

## 5.4 References

1. Benavent Acero, F.; Capobianco, C.S.; Garona, J.; Cirigliano, S.M.; Perera, Y.; Urtreger, A.J.; Perea, S.E.; Alonso, D.F.; Farina, H.G. CIGB-300, an anti-CK2 peptide, inhibits angiogenesis, tumor cell invasion and metastasis in lung cancer models. *Lung Cancer* **2016**, doi:10.1016/j.lungcan.2016.05.026.
2. Ampofo, E.; Rudzitis-Auth, J.; Dahmke, I.N.; Rössler, O.G.; Thiel, G.; Montenarh, M.; Menger, M.D.; Laschke, M.W. Inhibition of protein kinase CK2 suppresses tumor necrosis factor (TNF)- $\alpha$ -induced leukocyte-endothelial cell interaction. *Biochim. Biophys. Acta* **2015**, *1852*, 2123–36, doi:10.1016/j.bbadis.2015.07.013.
3. St-Denis, N.; Gabriel, M.; Turowec, J.P.; Gloor, G.B.; Li, S.S.C.; Gingras, A.C.; Litchfield, D.W. Systematic investigation of hierarchical phosphorylation by protein kinase CK2. *J. Proteomics* **2015**, *118*, 49–62, doi:10.1016/j.jprot.2014.10.020.
4. Turowec, J.P.; Vilck, G.; Gabriel, M.; Litchfield, D.W. Characterizing the convergence of protein kinase CK2 and caspase-3 reveals isoform-specific phosphorylation of caspase-3 by CK2 $\alpha$ : implications for pathological roles of CK2 in promoting cancer cell survival. *Oncotarget* **2013**, *4*, 560–71, doi:10.18632/oncotarget.948.
5. Turowec, J.P.; Duncan, J.S.; Gloor, G.B.; Litchfield, D.W. Regulation of caspase pathways by protein kinase CK2: identification of proteins with overlapping CK2 and caspase consensus motifs. *Mol. Cell. Biochem.* **2011**, *356*, 159–67, doi:10.1007/s11010-011-0972-5.
6. Duncan, J.S.; Turowec, J.P.; Duncan, K.E.; Vilck, G.; Wu, C.; Lüscher, B.; Li, S.S.-C.; Gloor, G.B.; Litchfield, D.W. A peptide-based target screen implicates the protein kinase CK2 in the global regulation of caspase signaling. *Sci. Signal.* **2011**,

- 4, ra30, doi:10.1126/scisignal.2001682.
7. Litchfield, D.W. Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochem. J.* **2003**, *369*, 1–15, doi:10.1042/BJ20021469.
  8. Salvi, M.; Sarno, S.; Cesaro, L.; Nakamura, H.; Pinna, L.A. Extraordinary pleiotropy of protein kinase CK2 revealed by weblogo phosphoproteome analysis. *Biochim. Biophys. Acta - Mol. Cell Res.* **2009**, *1793*, 847–859, doi:10.1016/j.bbamcr.2009.01.013.
  9. Rusin, S.F.; Adamo, M.E.; Kettenbach, A.N. Identification of candidate casein kinase 2 substrates in mitosis by quantitative phosphoproteomics. *Front. Cell Dev. Biol.* **2017**, *5*, doi:10.3389/fcell.2017.00097.
  10. Meggio, F.; Pinna, L.A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **2003**, *17*, 349–68, doi:10.1096/fj.02-0473rev.
  11. Nuñez de Villavicencio-Díaz, T.; Ramos Gómez, Y.; Oliva Argüelles, B.; Fernández Masso, J.R.; Rodríguez-Ulloa, A.; Cruz García, Y.; Guirola-Cruz, O.; Perez-Riverol, Y.; Javier González, L.; Tiscornia, I.; et al. Data for comparative proteomics analysis of the antitumor effect of CIGB-552 peptide in HT-29 colon adenocarcinoma cells. *Data Br.* **2015**, *4*, doi:10.1016/j.dib.2015.06.024.
  12. Gouw, M.; Sámano-Sánchez, H.; Van Roey, K.; Diella, F.; Gibson, T.J.; Dinkel, H. Exploring Short Linear Motifs Using the ELM Database and Tools. *Curr. Protoc. Bioinforma.* **2017**, *58*, 8.22.1–8.22.35, doi:10.1002/cpbi.26.
  13. Liu, F.; Wang, L.; Perna, F.; Nimer, S.D. Beyond transcription factors: how oncogenic signalling reshapes the epigenetic landscape. *Nat. Rev. Cancer* **2016**, *16*, 359–372, doi:10.1038/nrc.2016.41.
  14. Khan, D.H.; He, S.; Yu, J.; Winter, S.; Cao, W.; Seiser, C.; Davie, J.R. Protein Kinase CK2 Regulates the Dimerization of Histone Deacetylase 1 (HDAC1) and HDAC2 during Mitosis. *J. Biol. Chem.* **2013**, *288*, 16518–16528, doi:10.1074/jbc.M112.440446.
  15. de Villavicencio-Díaz, T.N.; Rabalski, A.J.; Litchfield, D.W. Protein kinase CK2: Intricate relationships within regulatory cellular networks. *Pharmaceuticals* **2017**.
  16. Kim, S.C.; Sprung, R.; Chen, Y.; Xu, Y.; Ball, H.; Pei, J.; Cheng, T.; Kho, Y.; Xiao, H.; Xiao, L.; et al. Substrate and Functional Diversity of Lysine Acetylation Revealed by a Proteomics Survey. *Mol. Cell* **2006**, *23*, 607–618, doi:10.1016/j.molcel.2006.06.026.
  17. Schölz, C.; Weinert, B.T.; Wagner, S.A.; Beli, P.; Miyake, Y.; Qi, J.; Jensen, L.J.; Streicher, W.; McCarthy, A.R.; Westwood, N.J.; et al. Acetylation site specificities of lysine deacetylase inhibitors in human cells. *Nat. Biotechnol.* **2015**, *33*, 415–423, doi:10.1038/nbt.3130.
  18. Cole, P.A. Chemical probes for histone-modifying enzymes. *Nat. Chem. Biol.* **2008**, *4*, 590–597, doi:10.1038/nchembio.111.
  19. Oh, S.; Suganuma, T.; Gogol, M.M.; Workman, J.L. Histone H3 threonine 11 phosphorylation by Sch9 and CK2 regulates chronological lifespan by controlling the nutritional stress response. *Elife* **2018**, *7*, doi:10.7554/eLife.36157.
  20. Cheung, W.L.; Turner, F.B.; Krishnamoorthy, T.; Wolner, B.; Ahn, S.-H.; Foley, M.; Dorsey, J.A.; Peterson, C.L.; Berger, S.L.; Allis, C.D. Phosphorylation of histone H4 serine 1 during DNA damage requires casein kinase II in *S. cerevisiae*. *Curr. Biol.* **2005**, *15*, 656–660, doi:10.1016/j.cub.2005.02.049.



21. Deplus, R.; Blanchon, L.; Rajavelu, A.; Boukaba, A.; Defrance, M.; Luciani, J.; Rothé, F.; Dedeurwaerder, S.; Denis, H.; Brinkman, A.B.; et al. Regulation of DNA methylation patterns by CK2-mediated phosphorylation of Dnmt3a. *Cell Rep.* **2014**, *8*, 743–753, doi:10.1016/j.celrep.2014.06.048.
22. Khan, D.H.; He, S.; Yu, J.; Winter, S.; Cao, W.; Seiser, C.; Davie, J.R. Protein kinase CK2 regulates the dimerization of histone deacetylase 1 (HDAC1) and HDAC2 during mitosis. *J. Biol. Chem.* **2013**, *288*, 16518–28, doi:10.1074/jbc.M112.440446.
23. Zhou, Y.; Gross, W.; Hong, S.H.; Privalsky, M.L. The SMRT corepressor is a target of phosphorylation by protein kinase CK2 (casein kinase II). *Mol. Cell. Biochem.* **2001**, *220*, 1–13, doi:10.1023/a:1011087910699.
24. Grimes, M.; Hall, B.; Foltz, L.; Levy, T.; Rikova, K.; Gaiser, J.; Cook, W.; Smirnova, E.; Wheeler, T.; Clark, N.R.; et al. Integration of protein phosphorylation, acetylation, and methylation data sets to outline lung cancer signaling networks. *Sci. Signal.* **2018**, *11*, doi:10.1126/scisignal.aag1087.
25. Kalousi, A.; Mylonis, I.; Politou, A.S.; Chachami, G.; Paraskeva, E.; Simos, G. Casein kinase 1 regulates human hypoxia-inducible factor HIF-1. *J. Cell Sci.* **2010**, *123*, 2976–86, doi:10.1242/jcs.068122.
26. Krystkowiak, I.; Davey, N.E. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res.* **2017**, *45*, W464–W469, doi:10.1093/nar/gkx238.
27. Villavicencio-Diaz, T.N.; Rodriguez-Ulloa, A.; Guirola-Cruz, O.; Perez-Riverol, Y. Bioinformatics tools for the functional interpretation of quantitative proteomics results. *Curr. Top. Med. Chem.* **2014**, *14*, 435–49.
28. Huber, W.; Carey, V.J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B.S.; Bravo, H.C.; Davis, S.; Gatto, L.; Girke, T.; et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **2015**, *12*, 115–121, doi:10.1038/nmeth.3252.
29. Rabalski, A.J.; Gyenis, L.; Litchfield, D.W. Molecular Pathways: Emergence of Protein Kinase CK2 (CSNK2) as a Potential Target to Inhibit Survival and DNA Damage Response and Repair Pathways in Cancer Cells. *Clin. Cancer Res.* **2016**, *22*, 2840–2847, doi:10.1158/1078-0432.CCR-15-1314.

## Appendices

### Appendix A. Supplemental Materials.

Chapter	Content
Chapter 1	Supplemental Tables. High-resolution images
Chapter 2	Supplemental Tables
Chapter 3	Supplemental Tables
Chapter 4	Supplemental Tables

### Appendix B. LC-MS runs and protein/peptide identification files.

Chapter	Content
Chapter 2	MaxQuant identification
Chapter 2	SearchGUI, PeptideShaker identification

### Appendix C. Scripts and computational resources.

Chapter	Content
Chapter 2	visualRepo application
Chapter 3	
Chapter 4	

## Curriculum Vitae

**Name:** Teresa Nunez de Villavicencio Diaz

**Post-secondary Education and Degrees:** University of Havana, Faculty of biology  
Havana, Cuba  
2006-2011 Lic. Biochemistry

Center for Genetic Engineering and Biotechnology  
Havana, Cuba  
2012-2013 MSc. Biotechnology.

The University of Western Ontario  
London, Ontario, Canada  
2016-2020 Ph.D.

**Honours and Awards:** Summa cum laude,  
2006-2011

Dean's Doctoral Scholarship from the Schulich School of  
Medicine & Dentistry,  
2016-2020

**Related Work Experience** Teaching Assistant  
The University of Western Ontario  
2017-2017

Researcher, Systems Biology  
Center for Genetic Engineering and Biotechnology  
Havana, Cuba  
2011-2015

**Publications:**

T Nuñez de Villavicencio-Diaz, AJ Rabalski, DW Litchfield  
Pharmaceuticals 10 (1), 27. (2017)

R Martínez, TN de Villavicencio-Díaz, A Sánchez, Y Ramos, JN Ferro, ...  
Biochemistry and biophysics reports 5, 379-387 (2016)

TN de Villavicencio-Díaz, YR Gómez, BO Argüelles, JRF Masso, ...  
Journal of proteomics 126, 163-171 (2015)

TN de Villavicencio-Díaz, YR Gómez, BO Argüelles, JRF Masso, ...  
Data in brief, 468-473 (2015)

TN de Villavicencio-Díaz, Y Mazola, YP Negrín, YC García, OG Cruz, ...  
Biochemistry and biophysics reports 4, 20-27 (2015)

T Nunez Villavicencio-Diaz, A Rodríguez-Ulloa, O Guirola-Cruz, ...  
Current topics in medicinal chemistry 14 (3), 435-449 (2014)