



Thèse

2024

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Investigating p21-mediated dynamic regulation of CDK1 and cyclin B1

Poulain, Lina

How to cite

POULAIN, Lina. Investigating p21-mediated dynamic regulation of CDK1 and cyclin B1. Doctoral Thesis, 2024. doi: 10.13097/archive-ouverte/unige:176991

This publication URL: <https://archive-ouverte.unige.ch/unige:176991>

Publication DOI: [10.13097/archive-ouverte/unige:176991](https://doi.org/10.13097/archive-ouverte/unige:176991)

© The author(s). This work is licensed under a Creative Commons Public Domain (CC0)

<https://creativecommons.org/publicdomain/zero/1.0/>

UNIVERSITÉ DE GENÈVE

Section de Biologie
Département de Biologie Moléculaire et Cellulaire

FACULTÉ DES SCIENCES
Professeur Andreas Boland

Investigating p21-mediated dynamic regulation of CDK1 and cyclin B1

THÈSE

présentée aux Facultés de médecine et des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences en sciences de la vie,
mention Biosciences moléculaires

par

Lina POULAIN

de

Digne les bains, France

Thèse N° 279

GENÈVE

Nom de l'Atelier d'Impression

2024

DOCTORAT ÈS SCIENCES EN SCIENCES DE LA VIE DES
FACULTÉS DE MÉDECINE ET DES SCIENCES
MENTION BIOSCIENCES MOLÉCULAIRES

Thèse de Mme Lina Poulain

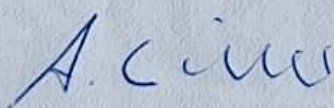
intitulée :

« Investigating p21-mediated dynamic regulation of CDK1 and Cyclin B1 »

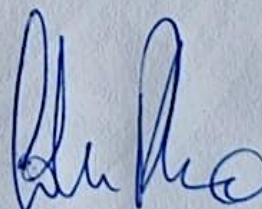
Les Facultés de médecine et des sciences, sur le préavis de Monsieur Andreas BOLAND, professeur assistant et directeur de thèse (Département de biologie moléculaire et cellulaire), Monsieur Paul GUICHARD, professeur associé (Département de biologie moléculaire et cellulaire) et Madame Alexis BARR, professeure (Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London, United Kingdom) autorisent l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 2 avril 2024

Thèse - 279 -



Le Doyen
Faculté de médecine



La Doyenne
Faculté des sciences

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

À mon Père

« Je n'ai point de prétention à la profondeur, mais je réclame le droit d'avoir un peu de bon sens »

David Copperfield, Charles Dickens

« Rien ni personne ne peut avoir tout à fait tort. Même une horloge arrêtée a raison deux fois par jour »

John Steinbeck

TABLE OF CONTENTS

TABLE OF CONTENTS.....	4
ABSTRACT	7
RESUME	8
ACKNOWLEDGEMENTS	9
INTRODUCTION.....	11
1. Introduction to the cell cycle.....	11
1.1. Cell division	11
1.2. Cell cycle	11
1.3. Mitosis in eukaryotes.....	12
2. Introduction to cell cycle regulation.....	15
2.1. Checkpoints.....	15
2.2. cyclins and CDKs.....	16
2.2.1. The central role of cyclins in CDKs activation.....	16
2.2.2. Structure of monomeric mammalian CDKs.....	17
2.2.3. Structure of monomeric mammalian cell cycle cyclins	18
2.2.4. CKS proteins	20
2.3. Cell Cycle regulation by CDK/cyclin complexes in mammals	21
2.3.1. During mitosis	21
2.3.1.1. G1 phase.....	21
2.3.1.2. S and G2 phases.....	22
2.3.1.3. M phase	22
2.3.2. During meiotic M phase	25
2.3.3. Role of the cyclin-dependent kinase inhibitor p21	26
2.3.4. P21 as a kinase regulator and the importance of its post-translational modifications	27
2.3.5. Role of p21 on cyclin B1/CDK1	30
3. Aim of the project.....	32
RESULTS	35
1. Protein analysis	35
1.1. The CCCP21 complex assembly.....	35
1.2. P21 interacts with the CCC complex even when cyclin B1 phosphate-binding pocket is mutated	37
1.3. CCC interacts strongly with the CDK1-recognition domains of p21.....	39
2. Usage of AlphaFold-generated structure predictions as a tool to understand the interactions between p21 and the CCC complex	41
3. Understanding the structure of CCCP21 with Electron Microscopy	44
3.1. Negative Staining of CCCP21.....	47
3.2. CryoEM of CCCP21	48
3.2.1. First dataset with tagged CCCP21	48
3.2.2. Datasets without the tag.....	50
3.2.3. Improving grid quality	57
3.2.3.1. Detergents	58
3.2.3.2. Salts.....	59
3.2.3.3. EM grid-types.....	61
4. HDX-MS studies	64
4.1. HDX-MS results upon p21 binding	65
4.2. HDX-MS results when the PBP is mutated	71
4.3. CCC EFE and TEM	74
5. Activity of CCCP21 WT, TEM and EFE.....	77

DISCUSSION	84
1. Structural information about CCCP21.....	84
2. The important role of dynamic disorder in p21-mediated regulation and its regulation by post-translational modifications	87
3. Cyclin B1 surface charges and hydrophobicity role on p21's binding and regulation	90
4. A potential relationship between the CDK1 T-loop and the cyclin B1 PBP	92
5. A proposed model for p21-mediated regulation of cyclin B1/CDK1 during mitosis	94
6. Perspectives and openings into the understanding of p21 degradation during mitosis	97
MATERIAL AND METHODS	98
1. List of reagents	98
2. List of protein sequences	98
3. Cloning, protein expression and purification	100
3.1. List of Polymerase Chain Reaction (PCR) protocols used during this study.....	100
3.2. Bacteria	100
3.3. Insect cells.....	101
4. Protein purification.....	101
4.1. Purification of the CC and CCC complexes.....	102
4.2. Purification of CKS1.....	102
4.3. Purification of cyclin B1.....	103
4.4. Purification of p21 1-35, 36-120 and 121-164	103
4.5. Purification of separase	104
4.6. Purification of p21 and p21 T57E and CCCP21 complex assembly	104
5. Pull-down assays	105
5.1. CC and separase.....	105
5.2. CCC and CDK1 and p21 1-35, 36-120 and 121-164	105
6. Radioactive kinase activity assays	106
7. AlphaFold predictions	106
8. Electron microscopy.....	106
8.1. List of EM grids used during this study	106
8.2. Graphene-oxide grid preparation	107
8.3. Negative staining	107
8.4. CryoEM of CCCP21 with tags	108
8.5. CryoEM of CCCP21 without tags	108
8.6. CryoEM of the first big CCCP21 dataset.....	109
8.7. CryoEM of the second big dataset	110
8.8. Cryo-EM of CCCP21 with the Center for Electron Microscopy	110
8.9. CryoEM to improve the quality of the particles.....	112
9. HDX-MS	113
10. LC-MS/MS	118
APPENDIX. Structural analysis of the Cell Migration-Inducing and hyaluronan-binding Protein (CEMIP)	120
1. Introduction	120
2. Description of the protein and its AlphaFold structure prediction	121
3. Attempts to determine the crystal structure of the first GG domain.....	123

3.1.	Expression and purification of GG domain 1	123
3.2.	Improving the protein crystallization chances	124
4.	Electron Microscopy	129
4.1.	Attempts to express and purify the full-length CEMIP	129
4.2.	Negative staining	130
4.3.	CryoEM.....	131
5.	Discussion.....	136
5.1.	Crystal structure of the first GG domain	136
5.2.	Cryo-EM analysis of the full-length CEMIP.....	136
5.3.	Potential future structural analysis of CEMIP interactions with HA and with other protein partners	137
6.	Material and methods	138
6.1.	AlphaFold predictions	139
6.2.	Cloning and protein expression	139
6.3.	Protein purification	140
6.4.	Protein crystallization and in-situ limited proteolysis.....	141
6.5.	Negative staining	142
6.6.	CryoEM.....	142
6.6.1.	Graphene oxide-coated EM grids preparation.....	142
6.6.2.	First small data collections	143
6.6.3.	Big data collection	143
REFERENCES		145

ABSTRACT

Cell cycle progression is a tightly regulated process that results in cell proliferation, senescence, or apoptosis. Faithful progression through the cell cycle is – among other proteins – controlled by a family of serine/threonine kinases whose activities strictly depend on a regulatory cyclin subunit and are therefore termed cyclin-dependent kinases (CDKs). CDK activity is further regulated through binding to CDK inhibitors (CKIs). CKIs prevent or limit the phosphorylation of substrates and are divided into two families: INK4 and CIP/KIP. The CIP/KIP (CDK-Interacting Protein/Kinase-Inhibitory Protein) family member p21(CIP1) (referred to as p21) is an intrinsically disordered protein that adopts a defined tertiary structure upon binding to target proteins. It has been proposed that p21 post-translational modifications, in particular phosphorylation of key residues, are important for CDK1 binding. Interestingly, a recent study from our group identified a novel phosphate-binding pocket in B-type cyclins that we speculate is important for p21 binding, once the inhibitor is phosphorylated at specific sites. This phospho-dependent regulation of p21-binding to the cyclin B/CDK1 complex would consequently represent a novel layer of regulation of CDKs. Furthermore, all existing structural studies have been conducted in the context of truncated inhibitor versions and it cannot be excluded that full-length proteins provide additional binding sites to promote high-affinity binding. During my thesis, I aimed at broadening our molecular understanding of CDK1 regulation by determining the structure of p21 in complex with the cyclin B1/CDK1/CKS1 complex and it reveals the dynamic binding mode of full-length p21 to cyclin B1/CDK using cryoEM. By using complementary biophysical and biochemical techniques such as Hydrogen/Deuterium exchange Mass Spectrometry (HDX-MS), I unveil the complexity of this dynamic regulation between p21 and the cyclin B/CDK1 complex. In addition, *in vitro* kinase activity assays show that p21-mediated regulation is organized into two separate but complementary steps during mitosis; a direct inhibition of CDK1 kinase activity and a conformational reorganization, mediated through the cyclin B1 phosphate-binding pocket, of the full complex allowing a slow transition to a fully active cyclin B1/CDK1 complex. Therefore, the results presented in my thesis provide novel insights into the molecular mechanisms that underly an ordered progression through the cell cycle.

RESUME

La progression à travers le cycle cellulaire est un processus hautement régulé favorisant la prolifération cellulaire, la sénescence ou encore l'apoptose. Le contrôle de ce processus est régulé, entre autres, par une famille de kinases à sérines/thréonines dont l'activité dépend strictement de leur association avec leur sous-unité régulatrice, les cyclines et sont ainsi nommées Kinases Dépendantes des cyclines (cyclin-Dependent Kinases, CDKs). L'activité des CDKs est de plus régulée par leur association avec des Inhibiteurs des CDK (CDK inhibitors, CKIs). Ces derniers préviennent ou limitent la phosphorylation de substrats par les CDKs et sont divisés en deux familles : les INK4 et les CIP/KIP. P21(CIP1) (ici nommée p21) est une protéine membre de la famille CIP/KIP (CDK-Interacting Protein/Kinase-Inhibitory Protein) intrinsèquement désordonnée adoptant une structure tertiaire spécifique selon la protéine avec laquelle elle interagit. Il a été proposé que les modifications post-traductionnelles apportées à p21, en particulier la phosphorylation de résidus clés, sont importantes pour la régulation de CDK1. Une récente étude publiée par notre groupe de recherche a identifié un nouveau site de reconnaissance de résidus phosphorylés chez les cyclines de type B et pouvant jouer un rôle important dans la régulation de CDK1 par p21. Cette régulation phospho-dépendante de la reconnaissance de p21 par le complexe cycline B/CDK1 pourrait représenter, de ce fait, un nouveau mécanisme régulateur de CDK1. De plus, toutes les recherches effectuées sur la structure de p21 ont été faites dans le cadre d'une protéine tronquée et il n'est pas impossible que la protéine entière ait plus de sites de reconnaissance que ceux démontrés jusqu'ici. Durant ma thèse, mon but était d'enrichir notre connaissance sur le mécanisme moléculaire de la régulation de CDK1 en déterminant la structure de p21 avec le complexe cycline B1/CDK1/CKS1 et ainsi révéler le dynamisme des modes de liaisons de p21 par cryoEM. Dans cette thèse est montré la complexité de cette régulation et grâce à des techniques de spectrométrie de masse HDX et de mesure de l'activité enzymatique de CDK1, j'ai pu montrer que p21 régulait CDK1 de deux manières différentes mais complémentaires. La première est par l'inhibition directe de l'activité kinase de CDK1 et la seconde par le réarrangement structural du complexe entier le menant doucement à son activation durant la mitose. Les résultats présentés dans cette thèse démontrent d'un nouveau processus moléculaire régissant la progression à travers le cycle cellulaire.

ACKNOWLEDGEMENTS

I would like to start by thanking my supervisor, **Prof. Dr. Andreas Boland**, for allowing me to perform my Ph.D. in his nascent research laboratory and for all the support during these years. I was the last member of the first generation of researchers to come in his group and we build everything from scratch all together. It gave me freedom to try a vast range of techniques and to think by myself about my research project. That was an important experience that enriched my independency skill and modeled how I perceive my future as a scientist.

I would like to thank my thesis jury members, **Prof. Dr. Alexis Barr** and **Prof. Dr. Paul Guichard** for the time taken to read and evaluate my thesis and my TAC committee, **Prof. Dr. Marko Kaksonen**, **Prof. Dr. Robbie Loewith** and **Prof. Dr. Paul Guichard**. Twice they accepted to guide me during my thesis and for that I am truly thankful. I thank also the representants of the Ph.D. life science program who listened to my concerns and guided me through those years, **Prof. Dr. Ivana Gasic**, **Prof. Dr. Florian Steiner** and **Prof. Dr. Marko Kaksonen**. Finally, I would like to thank all the other professors with who I felt safe to discuss my science and career and who offered me their advice and opinion, **Prof. Dr. Orsolya Barabas**, **Prof. Dr. Enrica Bordignon**, **Dr. Viriginie Hamel** and **Prof. Dr. Takuji Adaki**.

I would like to continue by thanking my scientific collaborators. First, **Dr. Oscar Vadas**, specialist in HDX-MS who performed experiments for me but also provided excellent explanations, written procedure, scientific support and advice. Second, **Prof. Dr. Leonardo Scapozza**, **Dr. Marco Prunotto** and **Dr. Sofia Spataro** who offered me a thrilling scientific collaboration but also taught me a lot about their field and their expertise. I also thank **Prof. Dr. Thomas Mayer** and his Ph.D. student **Rebecca Demming** for our collaboration on cyclin B3 and for sharing their expertise on cell cycle. I thank also the research group of Prof. Dr. Robbie Loewith, particularly **Dr. Lucas Tafur**, **Dr. Jennifer Kefauver** and **Dr. Luoming Zou** and the **entire research group** of Prof. Dr. Orsolya Barabas for sharing their scientific knowledge about electron microscopy and for helping me when no one else was.

My research lab mates are also to thank for their support during all those years. The first generation of colleagues, **Dr. Anna Höfler**, **Dr. Jun Yu** and **Dr. Pierre Raia**, it was far from being easy, but I am proud of what we did with what we had. **Dr. Margherita Botto**, you came to this team like sun comes after a long dark winter in the poles, I will never be grateful enough for everything you have done for me and what a great friend, colleague, scientist and woman you are. **Sophia Schmidt**, I believe in you, and you already grew so much in the last

two years spend in the lab, everything will be fine at the end. **Dr. XueFeng Zhang, Li Yao** and **Ilse Schaeffner**, I am happy to have met such nice people and to have shared your scientific path, I wish that you will accomplish great things and see you at your Nobel prize party Ilse. Finally, I would like to thank **Caroline Gabus**, technician in Prof. Dr. Robbie Loewith research group, for all the scientific but also personal support given during all my Ph.D. You are really the perfect technician, a brilliant scientist and fantastic teacher.

I would be lost without the most brilliant scientist and wonderful man I have ever met, **Dr. Yashar Sadian**. You taught me almost everything I learned during my Ph.D., EM, data processing, protein purification, data analysis and discussion. You were born to shine over us and to be a teacher. You put efforts in pushing my project, guided me to make clever decisions and made me a better scientist and person. I am very lucky to have you in my life, you shaped my future for the best and all you deserve is to be recognized forever for how smart and amazing you are.

I would like to thank all the great people I met in this university and who were there to help me when I was the most down. First, **Jelena Gajic**, you were the first one to reach for me during my darkest period and for this, I will always love you. Then comes my Ph.D. family, **Maria Capitaio, Umut Batman, Maxime Revel, Vincent Louvel, Marianna Ossorio, Elena Delfino, Dr. Joël Tuberosa, Marine Mauperin, Sofia Spataro, Mate Fulop** and at least a hundred more amazing people that I cannot cite all. The Ph.D. association **PhAGe** was very important for me during the two years I was part of it, and I will cherish forever our great memories together. I love you guys.

I would like to thank the people who accompany me in my personal life, my group of girls who became my sisters over time, **Dr. Marie Zilliox, Dr. Cécilia Luciani, Angélique Puget** and **Joanna de Chiara**. You are all four parts of the light that guides me through life. I also thank the wonderful scientists and friends I met during my master in Marseille and Stockholm who helped me during my Ph.D. Finally, I thank my sister **Julie Colignon** for her support and advice on my academic path and career, she knows this world even better than I do and I value her precious opinion.

To finish, I dedicate this work to my father, a man of letter and of science who taught me to always be driven by passion and compassion.

INTRODUCTION

1. Introduction to the cell cycle

1.1. Cell division

During the 17th century in Europe numerous eminent scientists lived including Antoni Van Leeuwenhoek and Robert Hooke. Their revolutionary work on developing microscopy techniques allowed them to observe for the first time what they described as the “cellula”, the basic element to form an organism^{1,2}, paving the way for many subsequent cellular studies. In the 19th century, Walter Flemming described in detail his observations of dividing cells^{3,4,5}, and published a comprehensive book named *Zellsubstanz, Kern und Zellteilung* (Cell substance, nucleus and cell division) where he described the different steps of cell division as we still know them today. Despite the limited equipment in comparison to modern light microscopes, he was able to describe the process of cell division and divided it into two distinct phases. First, a “progressive” phase as the appearance of the nuclear threads in the mother cell and their alignment in the center of the cell and a “degressive” phase as of the separation of the two groups of threads and the formation of two daughter nuclei.

By the end of the 19th century, Strasburger⁶ proposed terms to describe the different observed stages of cell division/mitosis that are still used today. In order of event, *Prophase* (apparition of a mitotic figure and the chromosomes become longitudinally split), *Metaphases* (alignment of the chromosomes in the equatorial plane of the spindles), *Anaphases* (separation of the chromatids on opposite poles of the spindle) and *Telophases* (recurrence of two nuclei and separation of the two daughter cells)⁷. Moreover, it was observed that two one diploid cell undergoes two cell divisions to form four haploid cells with the same number of chromosomes⁸. This process was termed meiosis. It is only when live imaging microscopy emerged at the beginning of the 20th century that researcher extensively studied the mechanics of mitosis and meiosis at their different stages.

1.2. Cell cycle

Until the mid-20th century, the mitotic cell cycle term was used to describe the division of two cells from one and was therefore including steps of the cell cycle that are not part of

the mitosis *per se*. It was, while studying extensively the interphases of beat root cells when researchers observed that phosphorus-32 is incorporated in cells during a certain window of time⁹ during the interphase, which was named S (for synthesis of DNA) phase. The periods of interphase before and after the S phase were named G1 (gap1) and G2 (gap2) phases dividing therefore the cell cycle into four distinct phases; G1, S, G2 and M¹⁰. The process of mitosis is defined as the separation of previously copied DNA molecules and the formation of two daughter cells from one mother cell. As described previously, mitosis is divided into four basic phases: prophase, metaphase, anaphase and telophase. Later it was discovered that cells can enter a “quiescence” state waiting for a signal to re-enter into the cell cycle¹¹, and this state was named G0 phase (Figure 1.).

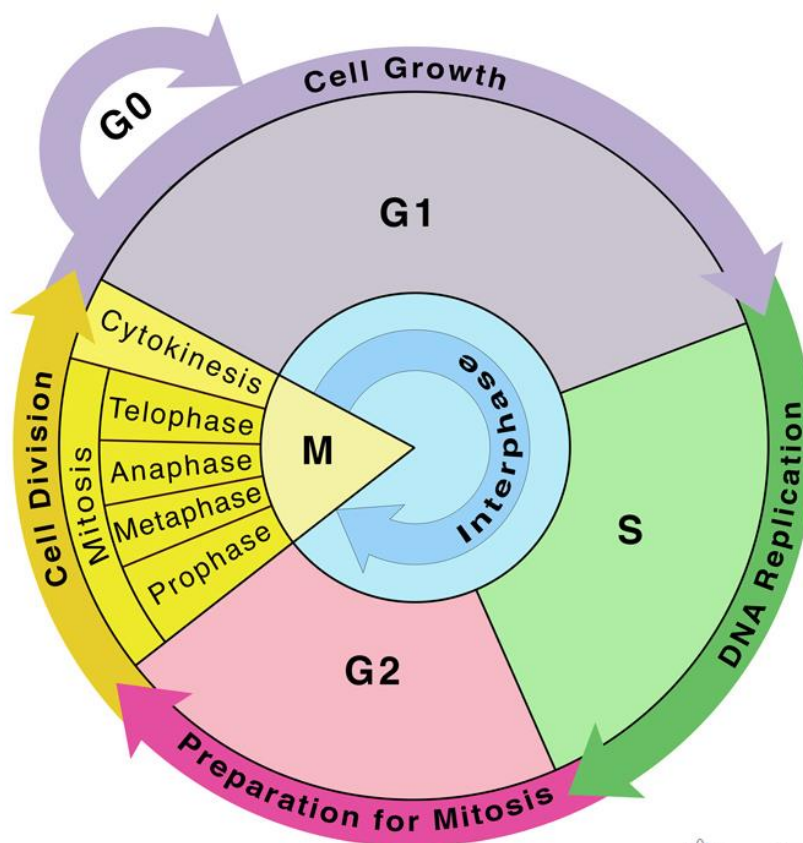


Figure 1. Diagram of the different phases of the cell cycle

Schematic representation of the four phases of cell cycle: G1, S, G2 and M. The M phase, also known as mitosis, is divided into five steps: prophase, metaphase, anaphase, telophase and cytokinesis, and is resulting in the formation of two daughter cells from one mother cell. G1, S and G2 phases are part of the interphase during which the cell grows, and its DNA is be replicated before entering M phase. An additional optional quiescence state phase G0 can occur at the end of M phase.

Taken from ScienceFacts.net, <https://www.sciencefacts.net/cell-cycle.html>

1.3. Mitosis in eukaryotes

The early mitotic events start at the end of the G2 phase, when the cell grew enough and produced all molecules needed to undergo mitosis (Figure 2.). The cell enters the prophase of mitosis, in which the duplicated DNA molecules (synthetized during S phase) condensate to form typical X-shaped chromosomes. At the same moment, the nuclear

envelop breaks down. Each chromosome is composed of two identical sister chromatids that are attached to one another at the centromere. The compaction and reorganization of chromosomes into their highly compact mitotic structure is catalyzed by the condensin protein complex, recruited to the chromatin¹². Cohesion of sister chromatids is mediated by a ring-shaped protein complex, named cohesin, present ultimately at the centromere only¹³. During prophase, in parallel to the condensation of the chromosomes, the mitotic spindle starts to polymerize from the two centrosomes, located at each pole of the cell¹⁴. The emerging polymerizing microtubules are responsible for the chromosome movement during mitosis^{15,16}. The disappearance of the interphase cell cytoskeleton facilitates the rearrangement of a round cell shape essential for the ultimate cytokinesis¹⁷ but also reduces the cytoplasm's viscosity, therefore facilitating protein trafficking within the cell.

The metaphase can be divided into two different sub-phases, pro-metaphase and metaphase. The end of prophase and beginning of pro-metaphase varies between species, however in eukaryotes, it is established that pro-metaphase starts when the nuclear envelop has entirely disappeared. At this time, the spindle attaches to the chromatids by a large protein structure present on centromeres, the so-called kinetochores. In humans, kinetochores are composed of over a hundred of different proteins and are formed by fibrous proteins and motor enzymes that generate the motile forces needed for chromosome movement and microtubules dynamics (cycles of microtubules polymerization and depolymerization) forming a gigantic structure anchoring 10 to 15 microtubules^{18,19}. The spindle therefore attaches in a very organized structure to the centromeres of each sister chromatid. Improper attachment is corrected during pro-metaphase²⁰ and separation of sister chromatids is delayed by the spindle assembly checkpoint, until correct microtubule attachment to the centromeres has occurred. In addition, correct alignment of the chromosomes at the metaphase plate takes place during the second step of metaphase. The metaphase plate is the equatorial plane on the spindle and defines the starting point of sister chromatids segregation during anaphase²¹. Because the chromosomes are not static on the metaphase plate but have the tendency to oscillate due to the pushing and pulling forces induced by the microtubules, a brief "sharpening" of the metaphase plate is visible before the gradual proteolysis of a number of regulatory proteins occurs and the onset of anaphase is induced.

The transition between metaphase to anaphase can be described as "a point of no return" because the sister chromatids will be irremediably separated by the cleavage of the cohesion subunit SCC1 which leads to an opening of the ring structure by the cysteine

endopeptidase separase²². Following the separation of chromosomes, two subsequent events will occur; Anaphase A, in which the distance between the chromosomes and the pole they face will decrease and Anaphase B, in which the separation between the poles will increase. These movements include rapid changes in length of microtubules. In Anaphase A, the Kinetochore microtubules (KMTs) will depolymerize and eventually disappear during their approach to the poles²³. In contrast, during Anaphase B, the non-attached microtubules to the kinetochores slide along each other forming the inter-polar spindle²⁴. The increasing number of the microtubules at the midplane of the separating chromosomes will form a bundle called “midbody”. This formation ultimately disappears as a consequence of numerous proteolytic events.

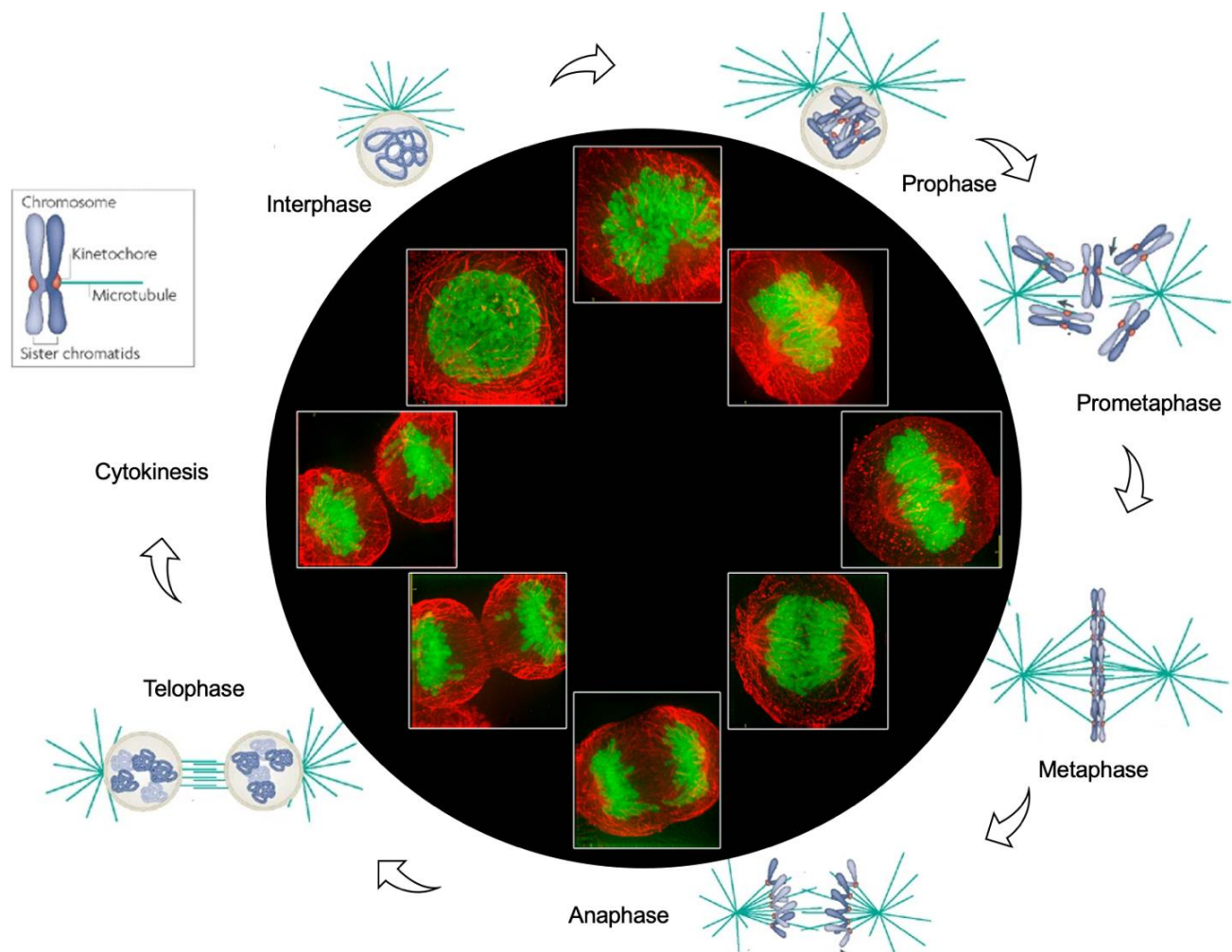


Figure 2. Diagram representing the differences phases of mitosis

In the center are represented the different mitotic steps in human HeLa cancer cells from the top, clockwise, prophase, prometaphase, metaphase, anaphase, telophase and cytokinesis. Drawings of each stage have been implanted to represent each stage. The chromosomes are stained in green and the mitotic spindle in red.

Different stages of mitosis in HeLa cancer cells, William J Moor, University of Dundee

Adapted from Cheeseman and Desai, *Nature Reviews Molecular Cell Biology*, 2008.

At the end of anaphase, the activity of the mitotic kinases drops preparing the two newly formed daughter cells to go back to interphase by regulating the activity of the condensin complex allowing imbedding of DNA into the newly forming nuclear envelopes. By the end of nuclear envelop and nuclear pore reconstructions, the chromatin decondenses allowing the DNA transcription, and the exchanges between nucleoplasm and cytoplasm to resume. The ultimate mitotic phase, the telophase, consists into the formation of two cells by cytokinesis and mitotic exit.

2. Introduction to cell cycle regulation

2.1. Checkpoints

The cell cycle is universal to all cells of all organisms and has evolved into a complex, tightly regulated process. The regulation of cell division is necessary to avoid uncontrolled cell duplication and appearance of neoplasms. During each transition from one phase to another phase of the cell cycle, checkpoints need to be passed in order to continue. Many proteins are involved in each checkpoint control allowing an organized regulation of the DNA replication, cell growth and mitosis. There are three checkpoints that need to be passed: One at the G1/S and G2/M transitions and one at the metaphase-to-anaphase transition (Spindle Assembly Checkpoint (SAC)) (Figure 3.). Central components of cell cycle regulation and in particular for the validation of the checkpoints are cyclin-dependent kinases. A dysregulated cell cycle and perturbation of CDK activity is considered as a hallmark of cancer²⁵. The concentration of these kinases does not vary along the cell cycle, however, their kinase activity requires binding to their activating subunits called cyclins, for which the concentrations are oscillating. I will describe in more detail here the roles of CDKs and cyclins during the different phases of the cell cycle, and the progression through the checkpoints with a focus on the entrance and exit of mitosis.

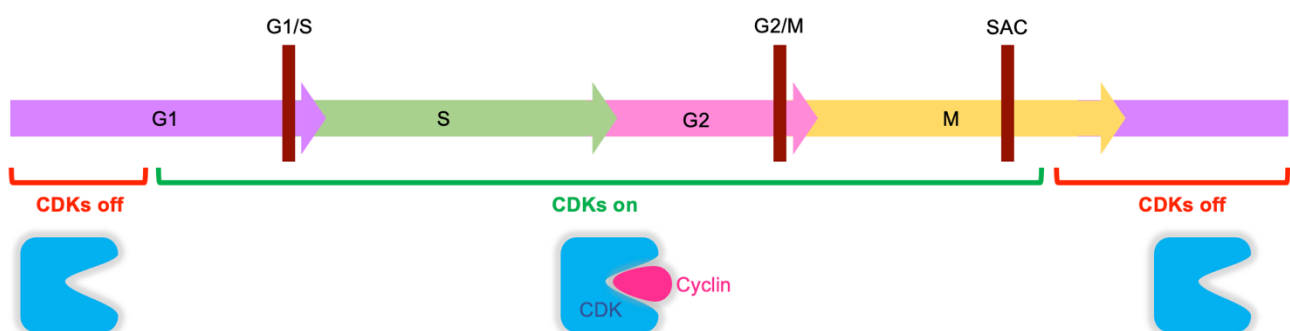


Figure 3. Schematic representation of the different checkpoints during cell cycle progression.

The activity of CDKs is dependent on binding to their activating partner, the cyclins. CDK activation regulates the passages through the three main cell cycle checkpoints: the G1/S, G2/M and the SAC.

Adapted from David O. Morgan, *The Cell Cycle*, 2007.

2.2. cyclins and CDKs

2.2.1. The central role of cyclins in CDKs activation

In human, there are 20 genes encoding for 11 Cyclin-Dependent Kinases (CDKs) and 9 CDK-like proteins, and more than 30 genes have been identified encoding for cyclins²⁶. CDK proteins share a well conserved primary structure and are part of a large serine/threonine protein kinase family. The best characterized CDKs are those that regulate the cell cycle (CDK1, 2, 4, 6 and 7)²⁷ and DNA transcription regulation (CDK7, 8, 9, 12 and 13)²⁸. The rest of the family members have more diverse functions from promoting cell proliferation²⁹ to viral RNA splicing³⁰. Like CDKs, not all cyclins are direct cell cycle regulators, but they all share a key role; binding and activating inactive CDKs. Eleven cyclins (A1-2, B1-3, D1-3, E1-2 and H) interact with five cell cycle-regulating CDKs and can be classified into four groups, the G1/S (cyclin E), S (cyclin A), M (cyclin A and B) and G1 (cyclin D) cyclins, depending on their role during cell cycle progression. Cyclins can have different functions depending on the cell type and the expression timing, but also have different functions during the distinct phases of the cell cycle.

The first CDK family members to be described were CDC28 in *Saccharomyces cerevisiae* and CDC2 in *Saccharomyces pombe*. During the 1970s, the *cdc28* and *cdc2*³¹ genes were discovered in *S. cerevisiae* and *S. pombe*, and it was demonstrated that these genes are necessary for ordered progression through the G1 phase³² and G2/M transition³¹ respectively. The function of the encoded CDC2 protein is modified by its phosphorylation^{33,34}, resulting in fluctuating kinase activity during the different phases of the cell cycle, reaching its highest peak during mitosis³⁵. The human homologue of CDC2³⁶ was originally named CDC2 until 1991 when a consortium of scientists decided to adopt the general term Cyclin-Dependent Kinases and renamed it CDK1. The characterization of different CDK1 homologues carried out by several research teams^{37,38}, provided enough evidence to postulate that there is a universal mechanism regulating eukaryotic mitosis³⁹, and that this mechanism is regulated by CDK1. At the same time, a new homologue of CDK1,

later named CDK2⁴⁰, was discovered in human cells, described as a mitotic CDK. In the following years, driven by the development of PCR technology, new CDK-family members were identified, such as the cell cycle regulators CDK4⁴¹, CDK6⁴² and CDK7⁴³, member of the CDK-activating complex (CAK) with its activating subunits cyclin H and MAT1⁴⁴.

In 1983, Tim Hunt's team discovered, while studying sea urchins, what they called "periodic proteins" or cyclins⁴⁵. They observed that cyclins are synthesized and degraded periodically during each cell cycle at specific time points. While studying the role and identification of the *cdc2* encoded Maturation-Promoting Factor (MPF) in *S. cerevisiae*, it was demonstrated that the appearance of cyclins is sufficient to induce mitosis in *Xenopus laevis*⁴⁶, and that if the degradation of cyclins is blocked, the cells were stuck in mitosis. This results clearly demonstrated that the synthesis and degradation of those proteins are necessary for entry and exit of mitosis⁴⁷. By the end of the 1980s, cyclin A⁴⁸ and B⁴⁹ were identified and cloned by different research teams and were shown to associate and activate the MPF (*cdc2* kinases CDK1 and CDK2)⁵⁰, and to be degraded by the ubiquitin pathway⁵¹ at a specific time, thereby allowing the cell cycle to resume. Shortly after the discovery of mitotic cyclins, the focus was placed on G1 phase regulating cyclins, namely E and D-types cyclins. These cyclins are key activators of CDK4 and CDK6, and ensure a timely progression through cell cycle and DNA replication⁵².

2.2.2. Structure of monomeric mammalian CDKs

Mammalian CDKs regulating the cell cycle are very similar in their polypeptide length, ranging from 297 amino acids for CDK1 to 346 amino acids for CDK7⁵³. They also share similarities in their structural organization (Figure 4.). All CDKs have an N-terminal lobe mainly composed of β -sheets and a C-terminal lobe organized mainly in α -helices. In between the two lobes an ATP molecule binds to the active site, forming a cleft that recognizes a protein substrate to which the terminal γ -phosphate of the ATP can be transferred (phosphorylation) once the CDK is activated. As a monomer, a CDK cannot phosphorylate a substrate, due to two structural features. First, CDKs have a large activation loop, also called T-loop, that blocks substrate-binding to the catalytic cleft. Upon binding to its activating cyclin, this T-loop needs to be phosphorylated by the CAK complex^{54,55}, and is pushed out of the cleft to face the cyclin co-factor when the complex is active⁵⁶. Second, binding to their activating cyclin also induces structural reorganization of some key residues, such as the glycine-rich region in the N-terminal lobe, to allow ATP to enter the active site

and attach properly to the CDK^{57,58}. Two common CDK α -helices are particularly important for CDK activation. The PSTAIRE helix (α 1) is located close to the N-terminal lobe and interacts with the cyclin to rearrange and reorient the residues interacting with ATP molecules. The L12 helix, located just before the T-loop forms a β sheet when the cyclin binds, reinforcing the T-loop rearrangement upon activation of the complex.

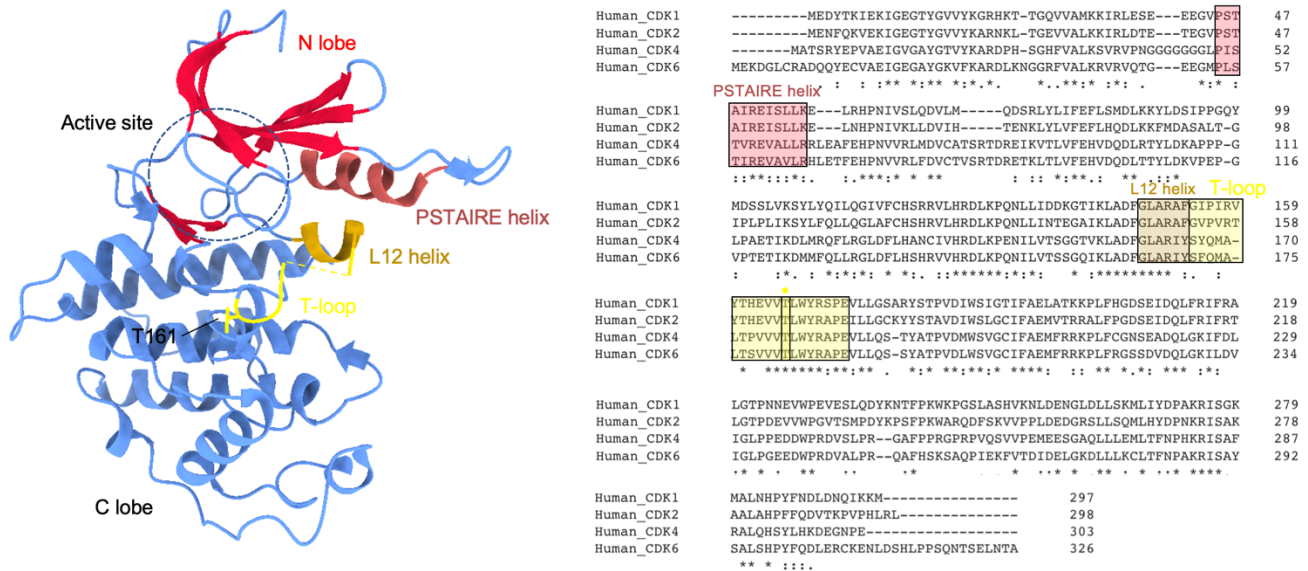


Figure 4. Description of CDK1's structure and conserved domains amongst CDKs

The left panel shows the structure of inactive CDK1 determined by x-ray crystallography⁵⁹ (PDB: 4YC6). The N-terminal lobe shown in red is mainly composed of β -sheets and in which the active site is located, where ATP molecules bind. The PSTAIRE helix is depicted in brown and necessary for binding to cyclin B1, B2 and A2. Located right below is the L12 helix (dark yellow), followed by the T-loop (light yellow) with Thr161 highlighted. Thr161 becomes phosphorylated by the CAK upon binding to a cyclin, resulting in an active complex. The right panel shows a multiple sequences alignment of different cell cycle CDKs with common features annotated using the same color code as in CDK1's structure. The yellow star highlights the highly conserved activating threonine common to many CDKs. This sequence alignment has been performed with Clustal Omega. Stars (*) mark invariant residues in all sequences. Colons (:) label highly similar residues. Dots (.) mark partially conserved residues.

2.2.3. Structure of monomeric mammalian cell cycle cyclins

Cyclins are quite different from another in their composition in amino acids but a region of 100-150 residues called the cyclin box can be found in all cyclin members and is responsible for the binding to their CDK partner and their subsequent activation⁶⁰ (Figure 5.). This cyclin box domain (CBD) can be divided into N terminal and C terminal subdomains common to all cyclins described in this thesis, except cyclin H in which the C-terminal cyclin box domain (or C-CBD) differs. Another common feature found in all cyclins is the presence of a lysine and a glutamate residue in the CBD (Lys257 and Glu286 in cyclin B1). These

residues are important for the structural integrity of the cyclin and to stabilize the interface with the CDK. The CBD is the folded part of the protein, whereas the N- and C-terminal regions of the protein are unstructured, show low sequence conservation, and vary in length between the different cyclins. This is notably striking for B-type cyclins that share very similar sequence identity and length of their CBDs, however while cyclin B1 and B2 have a similar polypeptide length, cyclin B3 possesses an roughly 8-time longer N-terminal unstructured region with still unknown functions⁶¹. The N- and C-terminal CBDs for A, B, D and E-type cyclins are relatively similar with a higher similarity between the mitotic A- and B-type cyclins. D-type cyclins differ to A- and B-type cyclins in their C-terminal CBD, which is more extended. The similarities between all those cyclin CBDs can be partially explained by their ability for some of them to bind the same CDKs, for example, CDK1 can be activated by A- and B-type cyclins, CDK2 by A-, B- and E-type cyclins and CDK3 by A- and E-type cyclins. The overall structure of a cyclin is organized into five α -helices. A-, B-, D- and E-type cyclins contain in their first helix (H1) a hydrophobic patch, also called the MRAIL motif, that is necessary to bind substrates bearing an RxL (Arg-x-Leu) consensus motif⁶². Finally, cyclins contain a destruction box found in the unstructured N-terminal part of the protein recognized by the E3 ubiquitin ligases responsible for their ubiquitination and ultimately their destruction by the proteasome.

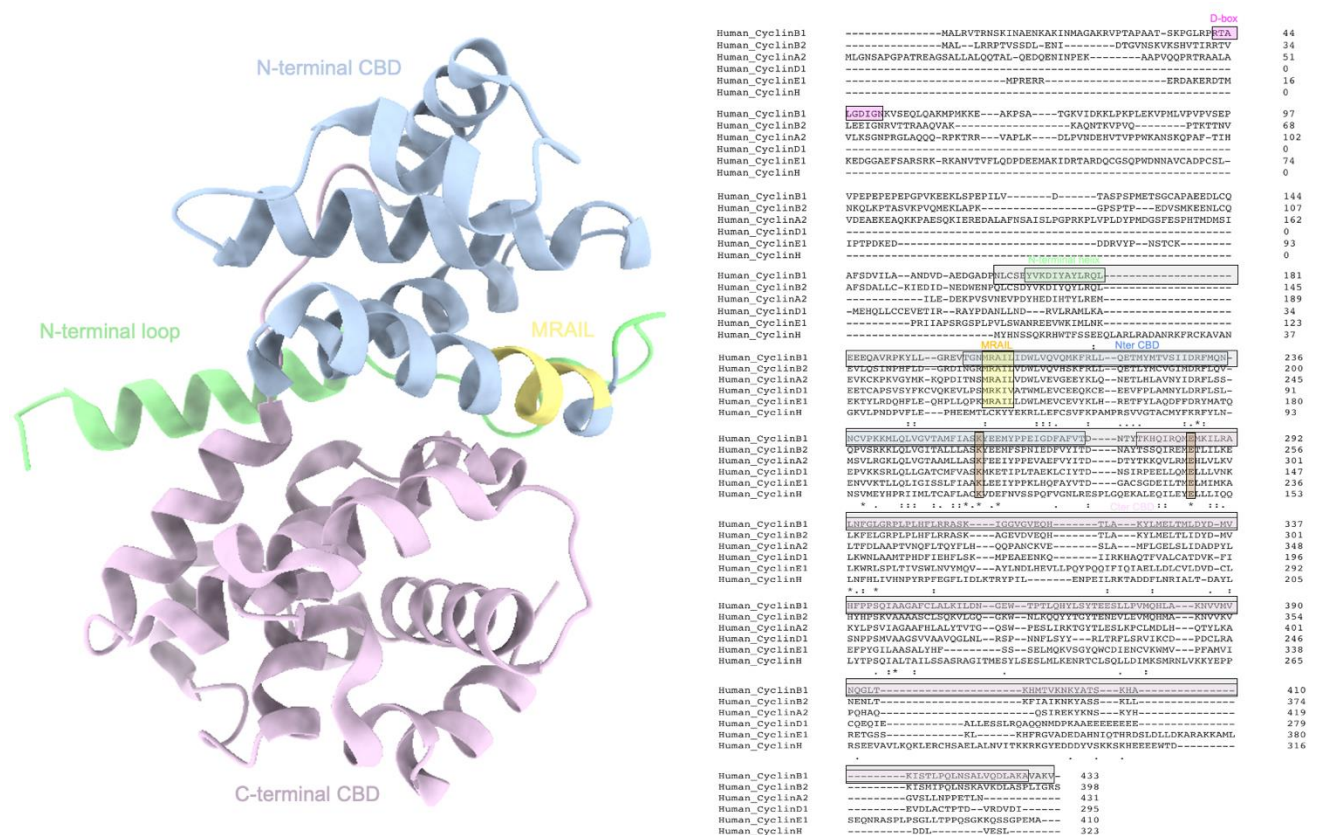


Figure 5. Description of the cyclin B1 structure and conserved domains amongst cyclins

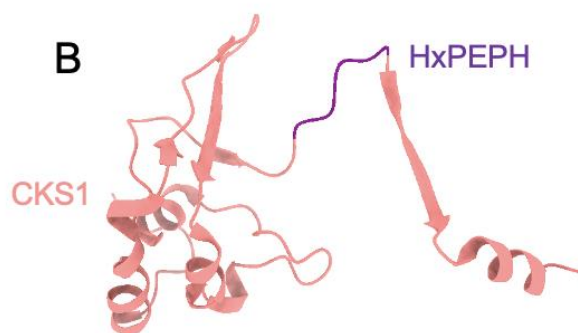
The left panel shows the structure of human cyclin B1 (164-433) determined by x-ray crystallography⁵⁹ (PDB: 4YC3). The N-terminal loop of cyclin B1 is shown in green, the N-terminal CBD composed of five α helices is shown in light blue, and the C-terminal CBD is colored in light pink. The helix contains the MRAIL motif is highlighted in yellow, the region necessary to bind RxL motif-containing interactors of cyclins. The right panel shows a multiple sequence alignment of different cell cycle cyclins, with the same color code as in cyclin B1's structure. The MRAIL motif is common to most cyclins, except cyclin H. All other annotations are specific to cyclin B1. Similar residues can be found in the closely related cyclin B2 but not in cyclin D1 or E1. Greyish-shaded residues of cyclin B1 highlight the structured part of the protein that can be seen in the left panel. The invariant lysine and glutamate residues present in all cyclins are labelled in brown. This sequence alignment has been performed with Clustal Omega. Stars (*) mark invariant residues in all sequences. Colons (:) label highly similar residues. Dots (.) mark partially conserved residues.

2.2.4. CKS proteins

A

Human_CKS1	MSHKQIYYSD	KYDDEEF	YRHVMLPKDIAKLVPKTHLMSESEWRNLG	VQC	SGWVHYMIH	60
Human_CKS2	MAHKQIYYSD	KYFDEHYEY	RHVMLPRELSKQVPKTHLMSEEEWRRLG	VQC	SLGWVHYMIH	60
	*:***** **.:*****:::* *****.***.***** *****					
Human_CKS1	EPEPH	ILLFR	RPLPKPKK	79		
Human_CKS2	EPEPH	ILLFR	RPLPKDQOK	79		
	*****. :*					

B



C

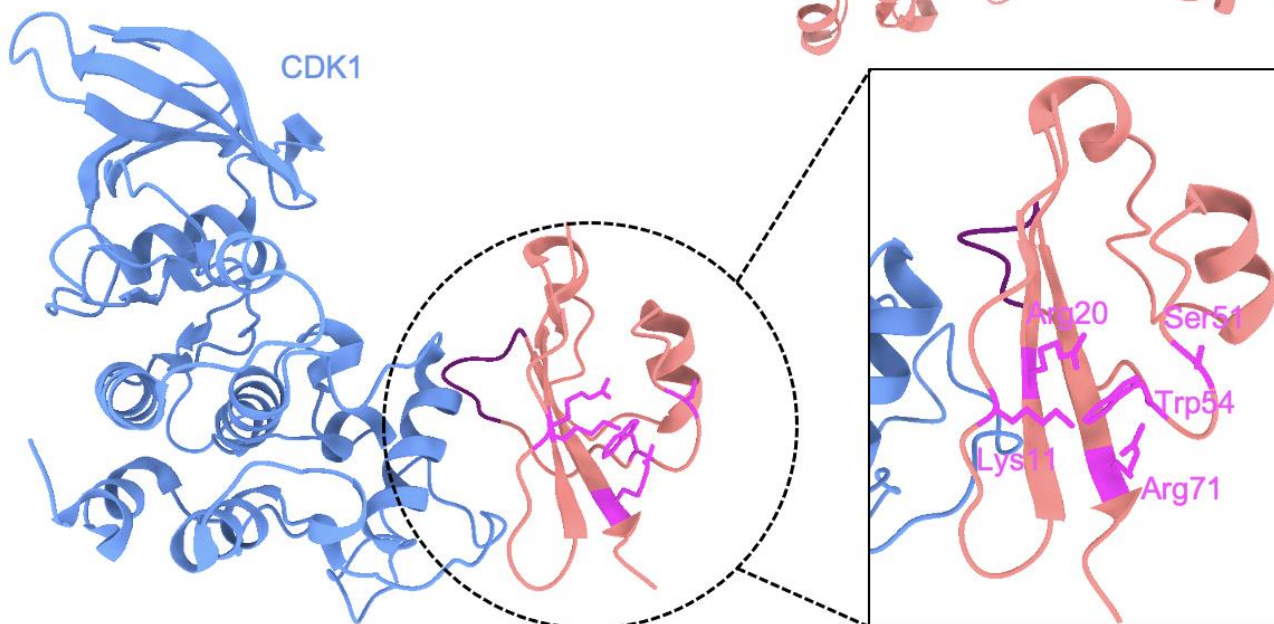


Figure 6. Structure of human CKS1 as a monomer and bound to CDK1

A, Sequence alignment of human CKS1 and CKS2 with annotated common features. In pink are the conserved residues (Lys11, Arg20, Ser51, Trp54 and Arg71) forming the phosphate-binding pocket of CKS1 and CKS2. The β -hinge is shaded in grey, with the conserved HxPEPH motif highlighted in purple. The structures are identically color-coded. The sequences alignment has been generated using Clustal Omega. Stars (*) mark

invariant residues in all sequences. Colons (:) label highly similar residues. Dots (.) mark partially conserved residues. **B**, Structure of the monomeric *S. cerevisiae* CKS1 resolved by X-ray crystallography⁶³ (PDB: 1QB3). **C**, Structure of CKS1 in complex with CDK1 resolved by cryoEM⁶⁴ (PDB:7NJ0), with a zoom on CKS1. The β -hinge containing the HxPEPH motif in purple refolds into a barrel-like structure allowing CKS1 to bind CDK1, but also liberating residues that form the phosphate-binding pocket shown in pink.

Cyclin-dependent Kinase Subunits (CKS) are small proteins that bind to a subset of CDKs and are existing in mammals in two paralogues, CKS1 and CKS2 (Figure 6. A). For a long time after their first observation at the end of the 1980s, the role of CKS proteins remained elusive. However, in the 1990s, it was shown that in *S. pombe*, *S. cerevisiae* and *X. laevis*, CKS1 is an important factor to initiate and promote mitosis by binding cyclin B1/CDK1⁶⁵. The structures of the human CKS1 as a monomer (Figure 6. B) and bound to CDK1 (Figure 6.C) show a protein composed of four anti-parallel β -sheets with a conserved β -hinge region comprising the residues HxPEPH, and two α -helices with an overall structure that resembles the N-terminal lobe of CDKs. Upon binding to CDK1, the β -hinge folds into a barrel-like structure, an important rearrangement allowing CKS1 to bind to the active cyclin/CDK complex and to promote CDK substrate specificity⁶⁶. Indeed, rearrangement of the β -hinge exposes five conserved residues amongst CKS protein (Lys11, Arg20, Ser51, Trp54 and Arg71 in CKS1) forming a phosphate-binding pocket⁶⁷ that binds to phosphorylated threonines of cell cycle regulators, thus enhancing the multisite-phosphorylation capacity of CDKs⁶⁸.

2.3. Cell Cycle regulation by CDK/cyclin complexes in mammals

2.3.1. During mitosis

2.3.1.1. G1 phase

D-type cyclins (cyclin D1, D2 and D3) and E-type cyclins (cyclin E1 and E2) are considered as the mammalian G1 cyclins. D-type cyclins bind and activate their catalytic partners CDK4 and CDK6, whereas E-type cyclins bind to CDK2 (Figure 7.). Cyclin D1, D2 and D3 in complex with CDK4 and CDK6 are mainly active during early G1 phase where they phosphorylate their substrates in the nucleus⁵². Their activity remains until the start of DNA replication, before their ubiquitination by the CLR4^{AMBRA1} E3 ubiquitin ligase and subsequent degradation by the proteasome^{69,70,71}. D-type cyclins also activate the expression of S phase inhibitors, mainly members of the Rb and Cyclin-Dependent Kinase Inhibitors (CKIs) families.

The activity of CDK2 in complex with cyclin E1 or E2 is maximal during the transition between G1 to S phase, also known as the G1/S checkpoint. These complexes are responsible for the stimulation of the expression and activity of G1 inhibitors like Rb and p27 in an analogous way to D-type cyclins with CDK4 and CDK6. Once entering S phase, cyclin E1 and E2 are ubiquitinated by the SCF^{Fbw7/SKP2} E3 ubiquitin ligase and subsequently degraded by the proteasome⁷².

2.3.1.2. S and G2 phases

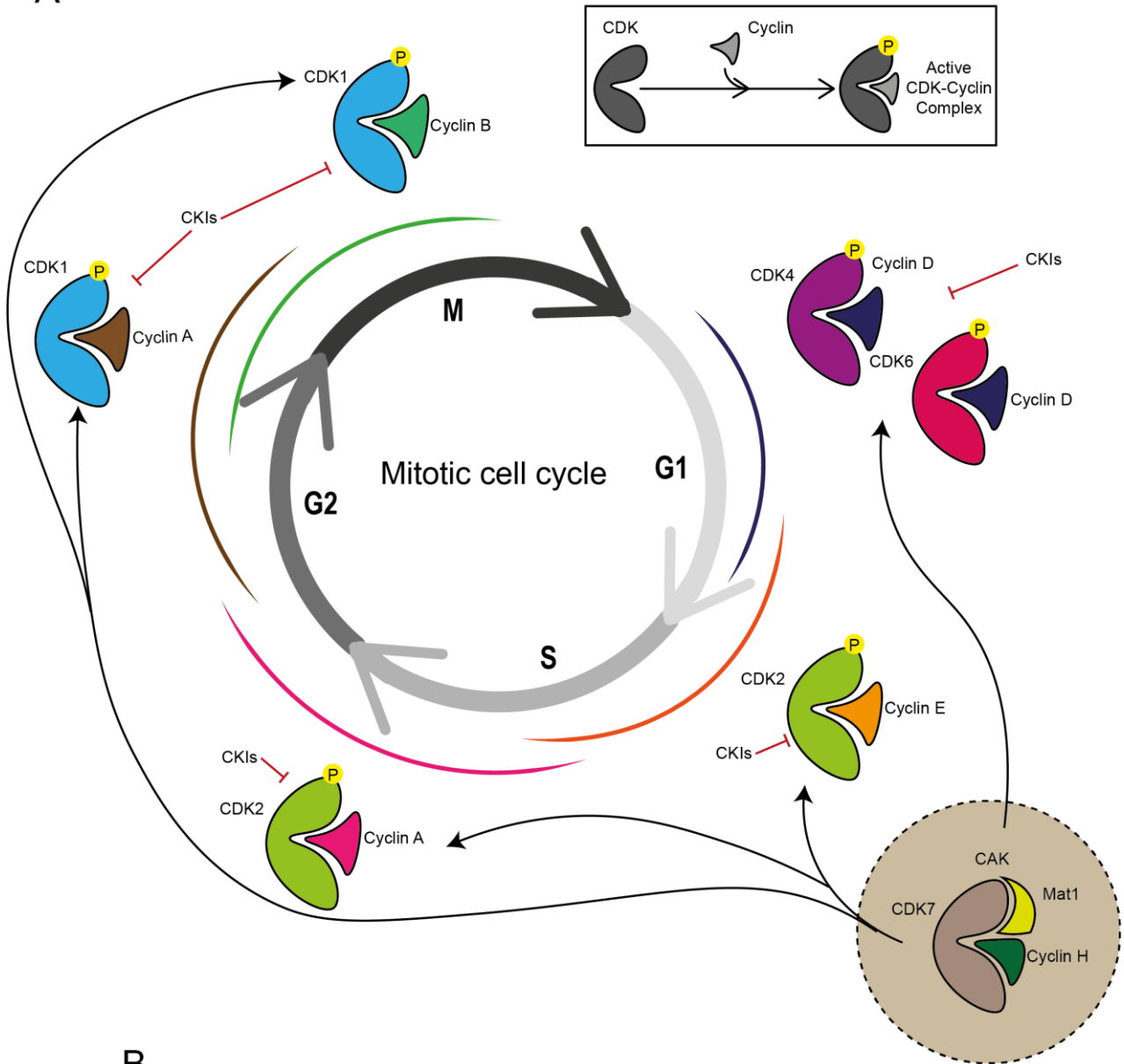
After dissociating from cyclin E1 and cyclin E2, CDK2 is activated by cyclin A2 in S phase (Figure 7.). Slow expression of cyclin A2 starts in G1 phase, reaching its peak during S phase and remains high until M phase. Cyclin A2/CDK2 complexes phosphorylate substrates involved in DNA replication, expression of mitotic cyclins and activation of mitotic CDKs⁵³. Once DNA replication is completed, with the help of the CKIs, the activity of cyclin A2/CDK2 is tamed while the formation of cyclin A2/CDK1 complexes starts in G2. Contemporaneously, cyclin B1 and B2 are expressed and associate with CDK1 (Figure 7.). Cyclin B1/CDK1 and cyclin B2/CDK1 complexes are the key regulators of M phase and therefore need to be quickly inactivated during the G2/M checkpoint by the kinases Wee1 and Myt1, through phosphorylation of Tyr14 and Tyr15 of CDK1, thereby keeping the complexes in an inhibited form²⁷. The CKI family member p27 inhibits cyclin A2/CDK1 during G2 avoiding an early entrance in M phase and is ultimately ubiquitinated by the SCF^{Skp2} E3 ubiquitin ligase complex⁷³, leading to a release of the kinase activity of CDK1. Cyclin A2/CDK1 is considered as a mitotic entry regulator by being part of the cascade of protein phosphorylations that lead to the activation of the phosphatase Cdc25 by the Polo kinase 1 (Plk1). Plk1 is responsible for the removal of the inhibitory phosphorylation on cyclin B1/CDK1 and therefore the start of mitosis⁷⁴. Cyclin A2/CDK1 and cyclin B1/CDK1 complexes remain active after the G2/M checkpoint.

2.3.1.3. M phase

CDK1 is the main regulator of mitosis and is activated by B-type cyclins (Figure 7.). Cyclin B1/CDK1 and cyclin B2/CDK1, in part due to their remarkable high number of substrates (>70 in mammalian cells) are involved in multiple cellular processes, such as the fragmentation of the Golgi network, the breakdown of the nuclear envelop⁷⁵, phosphorylation

of microtubule-binding proteins⁷⁶, replication and translation⁷⁷, and also in the regulation of proteins involved in mitotic exit, like Cdc20 and Cdh1⁷⁸, coactivators of the Anaphase-Promoting Complex/Cyclosome (or APC/C) E3 ubiquitin ligase, and regulation of the proteolytic activity of separase⁷⁹. Cyclin B2 co-localizes with the Golgi apparatus and has specific roles. Moreover, completion of mitosis is possible in the absence of cyclin B2 but not cyclin B1⁸⁰, demonstrating that the CDK1 substrate specificity is partly due to the localization of their activating cyclin⁸¹. In G2 phase formed cyclin A2/CDK1 complexes remain active during the first steps of mitosis and are degraded upon their ubiquitination by the APC/C^{cdc20} in prometaphase^{82,83}. Cyclin B1 expression level remains stable until their degradation after the SAC completion in anaphase also initiated by the APC/C. The degradation of cyclin A2/CDK1 is necessary for the progression from prometaphase to metaphase because cyclin A2/CDK1 complexes inhibit the stable attachment between the kinetochores and the microtubules⁸⁴. During interphase, cyclin B1/CDK1 complexes shuffle between the cytoplasm and the nucleus. In prophase, the nuclear localization signal in cyclin B1 allows the translocation of these complexes into the nucleus where they associate with the centrosomes and promote their separation and chromosomes condensation. From prophase until metaphase, an activating feedback loop led by Cdc25 increases the number of active cyclin B1/CDK1 complexes in the dividing cell, thus accelerating their nuclear translocation. Additionally and at the same time in order to be fully active, the conserved T-loop residue Thr161 in CDK1 is phosphorylated by the CAK complex⁸⁵.

A



B

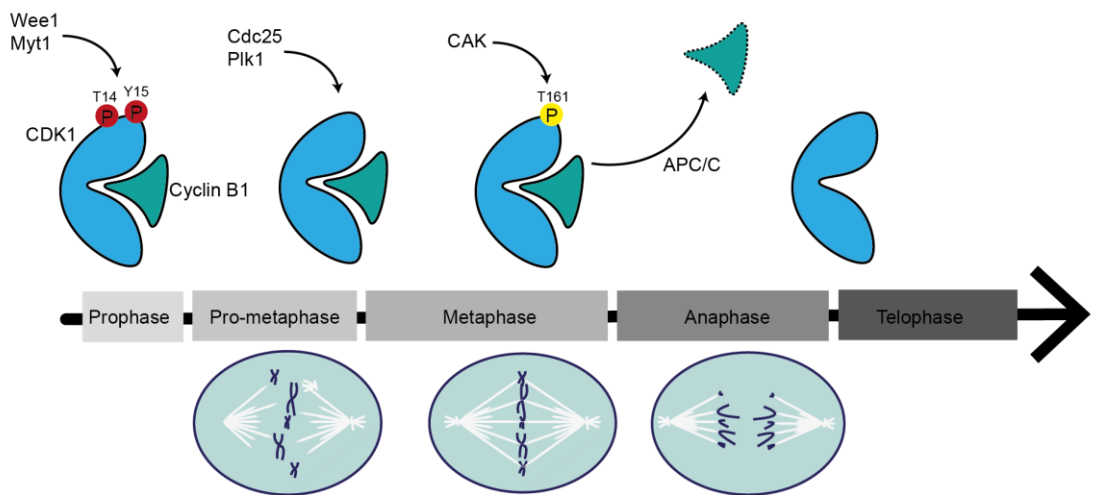


Figure 7. Oscillations of Cyclin/CDK activity throughout the cell cycle

A, Cartoon representation of the oscillating formation of distinct cyclin/CDK complexes in different phases of the cell cycle. Colored lines next to the picture of the cell cycle represent the timing when the same-colored cyclin and its CDK partner is expressed and active. CKIs (p21, p27 and p57) inhibit and regulate the cyclin/CDK complexes during the entire cycle. The CAK complex activates CDKs by phosphorylating their conserved threonine residues at multiple time points during the cell cycle. **B**, Focus on cyclin B1/CDK1 during mitosis and its highly regulated progressive activation through all the mitotic phases until the degradation of cyclin B1 during anaphase by the APC/C.

In prometaphase, while cyclin A2 is degraded by the APC/C, cyclin B1 is stabilized by the SAC. The CKIs-mediated inhibition of cyclin B1/CDK1 slows down and assures a complete G2/M transition until their ubiquitination by APC/C^{Cdc20} in prometaphase and subsequent degradation^{86,87}. While cyclin B1/CDK1 is the main catalyzer for the cell entrance into M phase, its degradation is equally important to exit from mitosis. In early mitosis, cyclin B1/CDK1 and cyclin B2/CDK1 complexes exhibit kinase activity, but also bind and inhibit separase^{64,79}, keeping separase enzymatically inactive and thus unable to separate sister chromatids⁸⁸. The second universal inhibitor of separase is the protein securin⁸⁹ and inactivation and degradation of both securin and cyclin B1/B2 is essential for the proteolysis of cohesin by separase. Cohesin, a ring-like protein complex, entraps the duplicated sister chromatids and opening of the cohesin ring is crucial to promote sister chromatid segregation. The dephosphorylation of many CDK1 substrates at the entrance of anaphase and consequently the degradation of cyclin B1 and B2 by the release of APC/C^{Cdc20/Cdh1} from the SAC continues until the end of the mitosis⁹⁰.

2.3.2. During meiotic M phase

During meiosis, cyclin A1, A2, B1, B2, B3 and O have specific functional roles. An important meiotic cyclin, which I worked on during my thesis, is cyclin B3. Cyclin B3 was identified for the first time in 1994⁹¹ in chicken and it is localized in the nucleus. It is expressed in somatic cells in a pattern similar to other B-type cyclins, and shares homology with A- and B-type cyclins and binds to CDK1. In mammals, cyclin B3 is about three times larger than its orthologues in non-mammalian vertebrates and this is due to the presence of a noticeable extension of exon 8⁹². In *D. melanogaster* and *C. elegans*, cyclin B3 is expressed in both meiotic and mitotic cells during embryogenesis. Cyclin B3 deficient mouse mutants show normal viability and fertility in males but females are infertile with degenerating foeti despite a normal ovary development⁹³. In addition, meiosis of *cyclin B3*^{-/-} oocytes arrest at metaphase with high levels of CDK1 and inseparable homologous chromosomes.

2.3.3. Role of the cyclin-dependent kinase inhibitor p21

The protein p21 (also known as Cip1, WAF1, Sdi1 or CAP20) was identified in the 1990s⁹⁴ as a polypeptide that migrates in an SDS-PAGE experiment at a molecular mass of approximately of 21 kDa and it was shown to be able to associate with and inhibit multiple cyclin/CDK complexes in cells^{95,96}. It is commonly accepted that p21 is a universal inhibitor of cyclin/CDK complexes (cyclin D/CDK4, cyclin E/CDK2, cyclin A/CDK2 and cyclin B/CDK1)⁹⁷. While originally described being associated in a complex consisting of a cyclin, a CDK and the Proliferating Cell Nuclear Antigen (PCNA) DNA clamp, p21 can also form a complex with each cyclin/CDK pair in the absence of PCNA and still act as a kinase inhibitor. The protein p21 directly regulates p53 and its associated pathway⁹⁸. The literature is very diverse regarding p21 and it has been shown that the protein is also involved in other processes⁹⁹ than regulating CDKs during the cell cycle, such as DNA replication and DNA repair¹⁰⁰, transcription¹⁰¹, apoptosis¹⁰² and even cell motility functions¹⁰³.

P21 homologues have been identified in several other organisms like the protein XIC1 in *X. laevis*¹⁰⁴, dacapo in *Drosophila melanogaster*¹⁰⁵, CKI-1 in *Caenorhabditis elegans*¹⁰⁶ or in *Arabidopsis thaliana*¹⁰⁷. In mice, the expression of p21 has been detected in various tissues like the spleen, thymus, lungs, testes, brain and in a particularly high concentration in the small intestine¹⁰⁸. Interestingly, p21 is less expressed in the heart, liver, skeletal muscles and the kidneys. These expression levels are independent of the expression of p53, except in the spleen. According to the Human Protein Atlas, in human, p21 expression has a low tissue specificity as well. RNA expression is found in all tissues, whereas the protein expression is very low or absent in the brain and muscles tissues.

Development of spontaneous tumors is observed in p21-depleted mice¹⁰⁹, however in human cancers, a total loss of function of p21 is almost never observed but rather a reduced protein expression, indicating that p21 on its own is not sufficient to act as a tumor suppressor. The anti-apoptotic and oncogenic activities of p21 are mostly described in a cytoplasmic protein localization context¹¹⁰, and are linked to aggressive malignancies with poor prognosis⁹⁹. This is mainly caused by promoting the bypass of G1/S arrest by active CDK4/6 that is associated with D-type cyclin family members¹¹¹. Few studies exist on the oncogenicity of nuclear p21 and the specific interactions with nuclear and mitotic cyclin/CDK complexes and specific p21 mutations reported in that context are rare. Most of them relate to the loss of controlled degradation of p21 or its inability to translocate into the nucleus. A

single point mutation in p21 in breast carcinomas can alter the inhibitory effect of p21 on a broad range of CDKs bound to their cyclin¹¹², however this is probably caused by the inability of p21 to bind to its target proteins, rather than a specific loss of inhibitory function on the cyclin/CDK complexes.

Because p21 is involved in the regulation of the cell cycle and DNA damage repair and replication, this protein is highly regulated by the signals that control these processes. It has been reported that p21 has a typical half-life of less than an hour before its degradation by the proteasome¹¹³. The protein p21 can be ubiquitinated by three different E3 ubiquitin ligase complexes: the SCF^{Skp2} and CRL4^{Cdt2} complexes^{114,115} in S phase and the APC/C^{CDC20} in M phase⁸⁶. A complete degradation of p21 in prometaphase by APC/C^{CDC20} is necessary for the full activation of CDK1 and the progression of mitosis.

2.3.4. P21 as a kinase regulator and the importance of its post-translational modifications

P21 inhibits the serine/threonine kinase activity of CDKs *in vitro* and *in vivo*⁹⁵. The protein has a stronger affinity and thus inhibitory effect on cyclin/CDK pairs involved in G1 and S phases (cyclin E/CDK2, cyclin A/CDK2 and cyclin D/CDK4 and CDK6)¹¹⁶, and interacts with lower affinity with CDK1¹¹⁶. P21 has also been reported to inhibit the kinase activity of the cyclin H/CDK7 CDK-Activating Complex (CAK), but no direct interaction has been detected between p21 and this complex¹¹⁷. Mainly described as a kinase inhibitor, p21 can also act as a positive regulator of CDK4 and CDK6 by promoting the complex formation with D-type cyclins¹¹⁸.

The protein p21 is encoded by the *cdkn1a* gene located on chromosome 6 in humans. It is a 164 amino-acid polypeptide (19 kDa) lacking any secondary and tertiary structure¹¹⁹, and 14% of its residues can be phosphorylated¹²⁰ (Figure 8. A and B). The first biophysical experiments performed on p21 revealed that the protein lacks stable structured domains and is highly disordered and flexible in solution, however capable of adopting a defined three-dimensional fold when bound to its target. P21 belongs to a group of protein called Intrinsically Disordered (or Unstructured) Proteins (IDP or IUP). Bioinformatic studies demonstrated that 35% to 51% of eukaryotic proteins contain disordered domains of at least 40 consecutive residues¹²¹. IDPs are known to play broad roles in signaling and are often involved in regulating different protein targets. CKIs members that share homology with p21, such as

p27 and p57, are also considered as IDPs. Because of the high flexibility and binding diversity that this class of IDP exhibit, they universally regulate all cyclins and CDKs. Despite the high sequence conservation between CDKs, and partly also cyclins, the flexibility of p21 allows subtle changes in the binding mode of p21 to the targets, resulting in slightly different complexes formation. P21 shares high homology with p27 (30%), however differs in its functions in cells. The bulk of conserved residues lies in the N-terminal part of these proteins, responsible for binding to the cyclins and CDKs.

The cyclin-dependent kinase regulators p21, p27 and p57 share a similar N-terminal Kinase Inhibitory Domain (KID)¹²² (Figure 8. A and B). P21's KID (residues 17-78) and p27's KID (residues 28-89), for which structural data are available, can be divided into three subdomains: cyclin-binding subdomain 1 (D1), long helix linker subdomain (LH) and cyclin-binding subdomain 2 (D2). Structural and biophysical analysis of p27's KID bound to cyclin A/CDK2 showed that this domain is necessary for a sequential binding of full-length p27 to the complex^{123,124}. Because p27 binds preferentially to cyclins alone rather than CDKs, it is proposed that the D1 subdomain binds first to the cyclin of a preformed cyclin/CDK complex, thereby facilitating the binding of the D2 subdomain to the CDK. The central LH subdomain will, upon binding, adopt a partial α -helical conformation and play a structural role in connecting the D1 and D2 subdomains. In p21's case, it was shown that the LH subdomain of its KID is less rigid but rather more dynamic¹²⁵. Binding to cyclin A/CDK2, results in an extended conformation of the p21 LH domain, a structural feature critical for the activity of the complex. In contrast, if the LH domain is shortened, the inhibitory activity of p21 on cyclin B1/CDK1 is affected. By modulating its length, the p21 LH domain is the key determinant of the specific recognition of the cyclin/CDK complex to which p21 binds to.

Two cyclin-binding motifs can be found in p21¹²⁶, Cy1 (residues 17 to 24) and Cy2 (residues 155-158) (Figure 8. A). Both motifs contain a well-conserved RxL (Arg-x-Leu) cyclin-interacting motif that can also be found in other cyclin/CDK interacting proteins, and that is essential for p21 to regulate the CDK kinase activity¹²⁷. The CDK-binding motif K (residues 53 to 58) contains a threonine residue (Thr57) that, once phosphorylated by CDK2 during the G2/M transition phase, binds to CDK1 during M phase¹²⁸. P21 also contains a so-called 3₁₀ helix (residues 74 to 79) that binds and acts in conjunction with the K domain on CDKs. Based existing structural information such as the crystal structure of p21 bound to cyclin D1/CDK4¹²⁹, as well as the crystal structure and the cryo-EM structures of p27 bound to cyclin A/CDK2^{124,130}, the CDK N-terminal lobe and more specifically its ATP-binding site is

sterically blocked by the well conserved Tyrosine 77 (Tyr, Y77) of the 3₁₀ helix domain of p21 (Tyr, Y88 in p27).

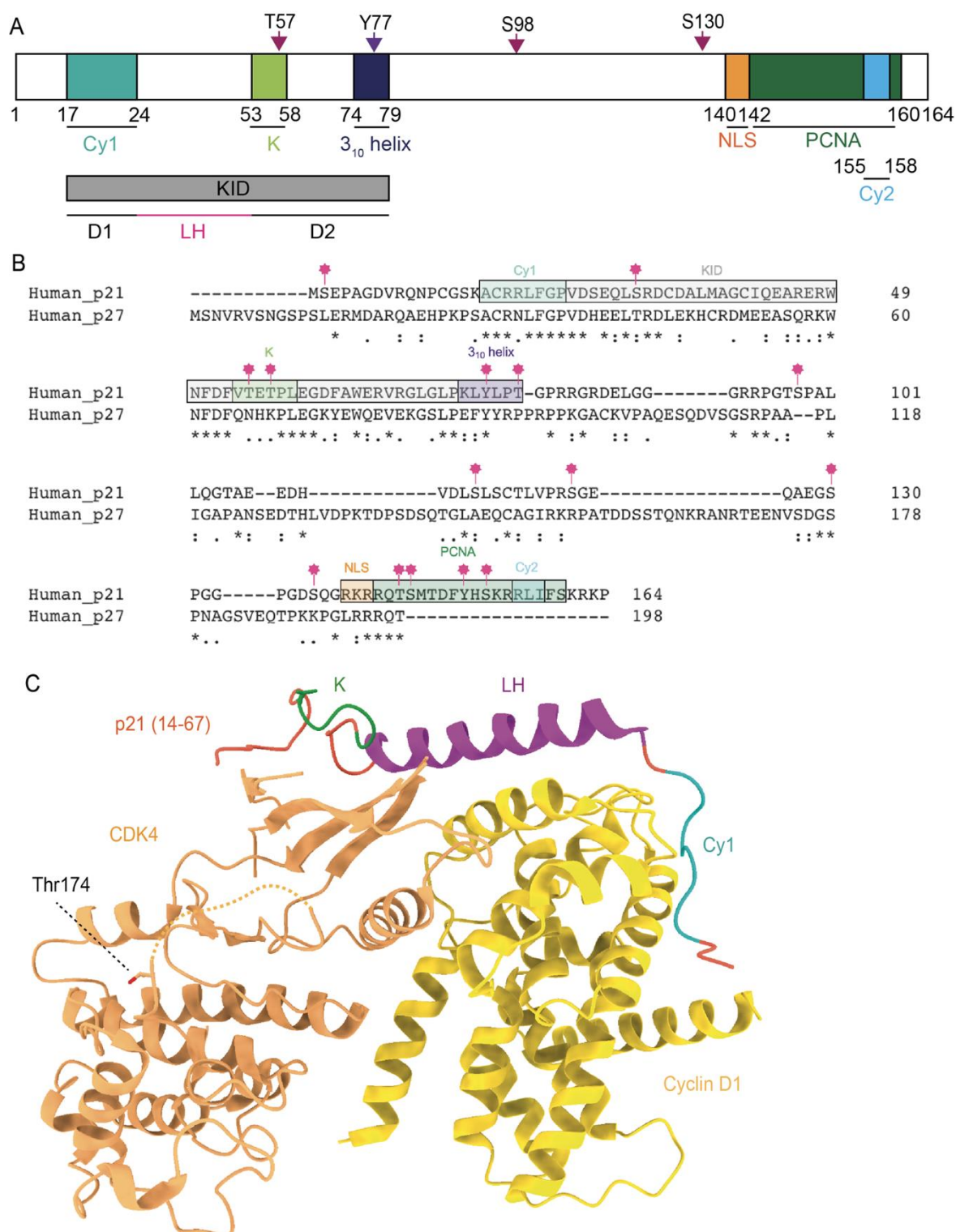


Figure 8. Amino-acid sequence alignment of human p21 and p27

A, Schematic representation of human p21 with the different domains annotated like in A; as well as the KID domain and its three sub-domains. **B**, This sequence alignment has been generated using Clustal Omega. Stars (*) mark invariant residues in all sequences. Colons (:) label highly similar residues. Dots (.) mark partially conserved residues. The different domains of p21 are annotated as follows: Cy1 (17-24, KID (17-79), K (53-58),

3₁₀ helix (74-79), NLS (140-142), PCNA (142-160) and Cy2 (155-158). Known p21 phospho-sites are marked with pink stars. **C**, Crystal structure of cyclin D1/CDK4 bound to a fragment of p21¹³¹ (PDB: 6P8H). The resolved segment of the KID domain includes the Cy1 domain (in blue), as well as the K domain (in green) that follows the LH (in purple), however the 3₁₀ helix is missing. Moreover, structural information about CDK4 T-loop was not resolved due to loop flexibility, and the well conserved Thr174 is resolved unphosphorylated.

The C-terminal region of p21 represents the most enigmatic part of the protein, with only little information about its involvement in cyclin/CDK complexes activity regulation known. The only existing structure of p21 bound to cyclin D1/CDK4 (Figure 8. C) lacks the information about its C-terminus or cannot be linked to a function regarding cyclin/CDK complex regulation¹³². In addition to the second RxL motif, a Nuclear Localization Signal (NLS) motif RKR (Arg-Lys-Arg) (residues 140 to 142) resides in the C-terminal region of p21¹³³. Importantly, a serine residue (Ser130) located in the C-terminal region is phosphorylated by cyclin E/CDK2 and was shown to promote SCF^{SKP2} E3 ubiquitin ligase-mediated degradation of p21¹³⁴. To date, no structural information exists on p21 associated with any ubiquitin ligase complex.

2.3.5. Role of p21 on cyclin B1/CDK1

The unstructured nature of p21 ensures a high degree of flexibility, which likely allows the protein to bind to several distinct cyclin/CDK complexes. In addition, post-translational modifications contribute to target binding. P21 was first described as a kinase inhibitor but it is now known that this protein also possesses many other regulatory roles. Moreover, the fact that p21 can regulate all cyclin/CDK complexes does not mean that it regulates all complexes identically. During my thesis, I focused on understanding the regulation of cyclin B1/CDK1 by full length p21.

Cyclin B1 and CDK1 are involved in the progression through M phase of the cell cycle. P21 was shown to interact with cyclin B1/CDK1 at the onset of M phase, blocking the cells in G2¹³⁵. The delay of mitosis caused by the DNA damage-induced G2 checkpoint can still occur in the absence of p21 and p53, as the protein kinase Chk1 inhibits Cdc25 and CDK1 Thr14 and Tyr15 dephosphorylation^{136,137}, but their presence is necessary for a sustained G2 arrest once cdc25 is active. Indeed, p21 does not inhibit dephosphorylation of Thr14 and Tyr15, but is more likely to inhibit transiently the phosphorylation of Thr161 by the CAK complex in the nucleus (Figure 9.). Thr161 phosphorylation is strictly needed for CDK1 kinase activity⁸⁷.

As reported before, p21 binds with lower affinity to cyclin B1/CDK1 than to other complexes¹¹⁶ and less protein complex stoichiometry is observed *in vivo*⁸⁷ revealing that the

known inhibiting pathways of cyclin B1/CDK1 are not the only mechanisms responsible for p21-mediated kinase inhibition. During M phase, p21 is gradually phosphorylated on different residues by cyclin B1/CDK1 and its associated kinases as these phosphorylation disappear once cyclin B1 is degraded at the end of the mitosis and do not appear in its absence¹³⁸. The CDK1 consensus (S/T-P-X-K/R) site can be found three times in p21¹³⁹, namely Thr57, Ser98 and Ser130 (Figure 8. A). Ser130 was shown to be directly phosphorylated *in vitro* by cyclin B1/CDK1¹³⁸. Phosphorylation of Ser130 does not influence CDK1 kinase activity, but instead p21 and cyclin B1 form a less stable complex that is rapidly degraded. Interestingly, phosphorylation of Ser130^{p21} by the cyclin D/CDK4-6 complex was shown to be required for CDK7-mediated phosphorylation of the bound CDK on their threonine 172 (Thr172), the equivalent of Thr161 for CDK1¹⁴⁰. Phosphorylated p21 on Thr57 enhance the formation of cyclin B1/CDK1 complexes and promotes its activity. Moreover, p21 presence is necessary for the arising activity of the complex. When cells are p21 depleted, the kinase activity of cyclin B1/CDK1 is significantly delayed even though the levels of cyclin B1 and CDK1 remain the same. What is rather observed it a reduced associated CDK1 and specially phosphorylated Thr161 CDK1¹²⁸ indicating that p21 could act as a temporary glue between cyclin B1 and CDK1 until the complex is stabilized by the phosphorylation on Thr161.

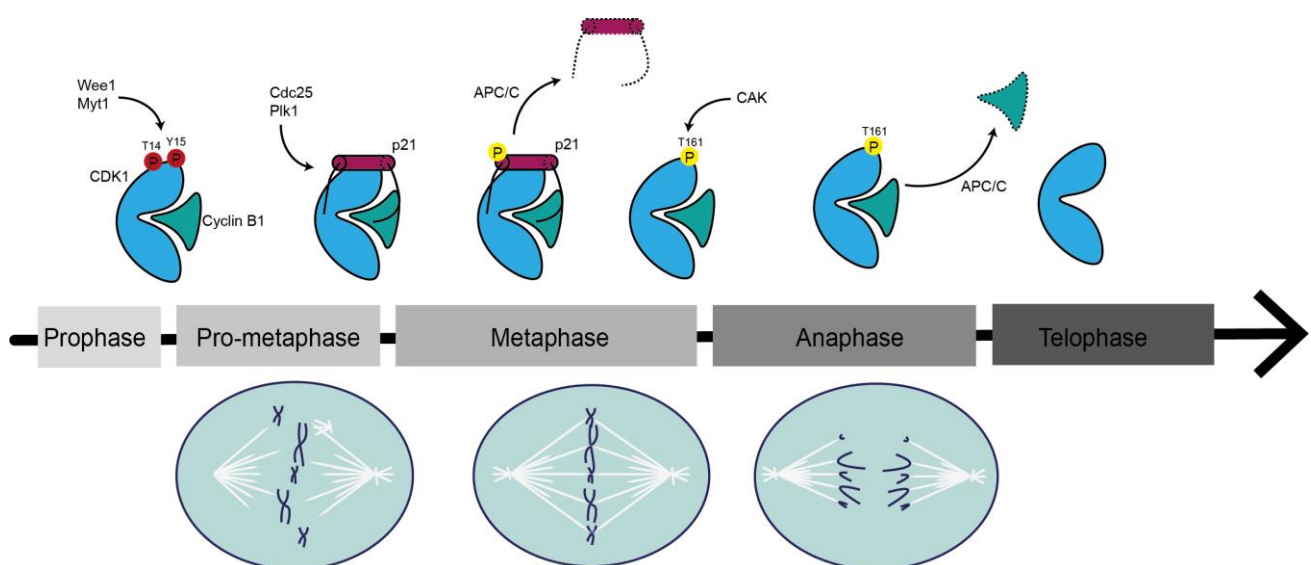


Figure 9. Regulation of cyclin B1/CDK1 during mitosis with more details on p21's role

What is known and proposed about p21's role during mitotic events on CDK1 and cyclin B1. The residues here are annotated with one letter where T stand for threonine and Y stands for tyrosine. The yellow P represents a phosphorylated protein. P21 binds inactive cyclin B1/CDK1 complex during G2/M checkpoint without interfering the dephosphorylation of CDK1 Thr14 and Tyr15 by Cdc25 and Plk1. The slowly incorporated post-translational modifications to p21 induces its detachment from the complex during metaphase and its degradation by the proteasome mediated by the E3 ubiquitin ligase APC/C. Consequently, cyclin B1/CDK1 is fully active during anaphase before the degradation of cyclin B1 once sister chromatids are segregated.

3. Aim of the project

Since its discovery, p21 has been extensively studied by numerous research groups. Like the other CKIs family members p27 and p57, p21 plays a central role during cell cycle progression by regulating the activity of several active cyclin/CDK pairs. P21 inhibits the kinase activity of CDKs, but also promotes the formation of active complexes. As described in the introduction, p21 (unlike p27 and p57) has a ubiquitous role on all cell cycle-regulating cyclin/CDK complexes, by directly binding to them or by occluding substrate binding to CDKs. Despite the astonishing diversity of discoveries on p21, very little information on the structure of p21, e.g., in complex with a cyclin/CDK pair exist up to date. This lack of information is caused by the intrinsic disordered nature of p21. So far, only the LH region has been shown to adopt a helical conformation upon binding to cyclins and CDKs, bridging the two proteins. Often, IDPs are unstable proteins in isolation prone to degradation by proteases. As a consequence, IDPs are studied in the presence of a protein partner *in vitro*. Structural studies by X-ray crystallography cannot be conducted in the context of full-length p21, because the inherit flexible nature of p21 does not allow the formation of protein crystals. Therefore, only the structured regions of p21 (14-81) has been resolved by crystallography in complex with cyclin D1/CDK4¹³¹, highlighting the limitations of X-ray crystallography to solve unstructured regions of p21. Another explanation is that p21 is a highly phosphorylated protein with at least 15 reported phosphorylated sites in the literature. Each of those phosphorylation site has a unique purpose during distinct steps of the cell cycle and phosphorylation of the sites depends on many different factors including the localization of p21 in the cell, to which target complex p21 is bound and by which kinase it has been phosphorylated¹⁴¹. Moreover, it has been shown that dynamic rearrangements of p27, when bound to cyclin A2/CDK2, was driven by specific phosphorylations, ultimately leading to the reactivation of CDK2 kinase activity^{142,143,144}. Similar mechanisms have been hypothesized for p21 but recent molecular dynamic simulations showed that Tyr77 of p21, when phosphorylated, unlike Tyr88 of p27, does not facilitates the ejection of the 3₁₀ helix out of the CDK2 active site but rather enhances the formation of a stable ternary complex together with cyclin A2¹⁴⁵. These results are contrary to previous NMR observations describing an increased CDK2 activity once Tyr77 is phosphorylated and the ejection of residues 70 to 83, corresponding to the 3₁₀ helix, from the ATP-binding site has occurred¹⁴⁶. To overcome these contradictions present in current literature and to better understand p21-mediated regulation of cyclin/CDK complexes,

structural and functional data on p21 dynamic binding modes to CDKs and cyclins are required.

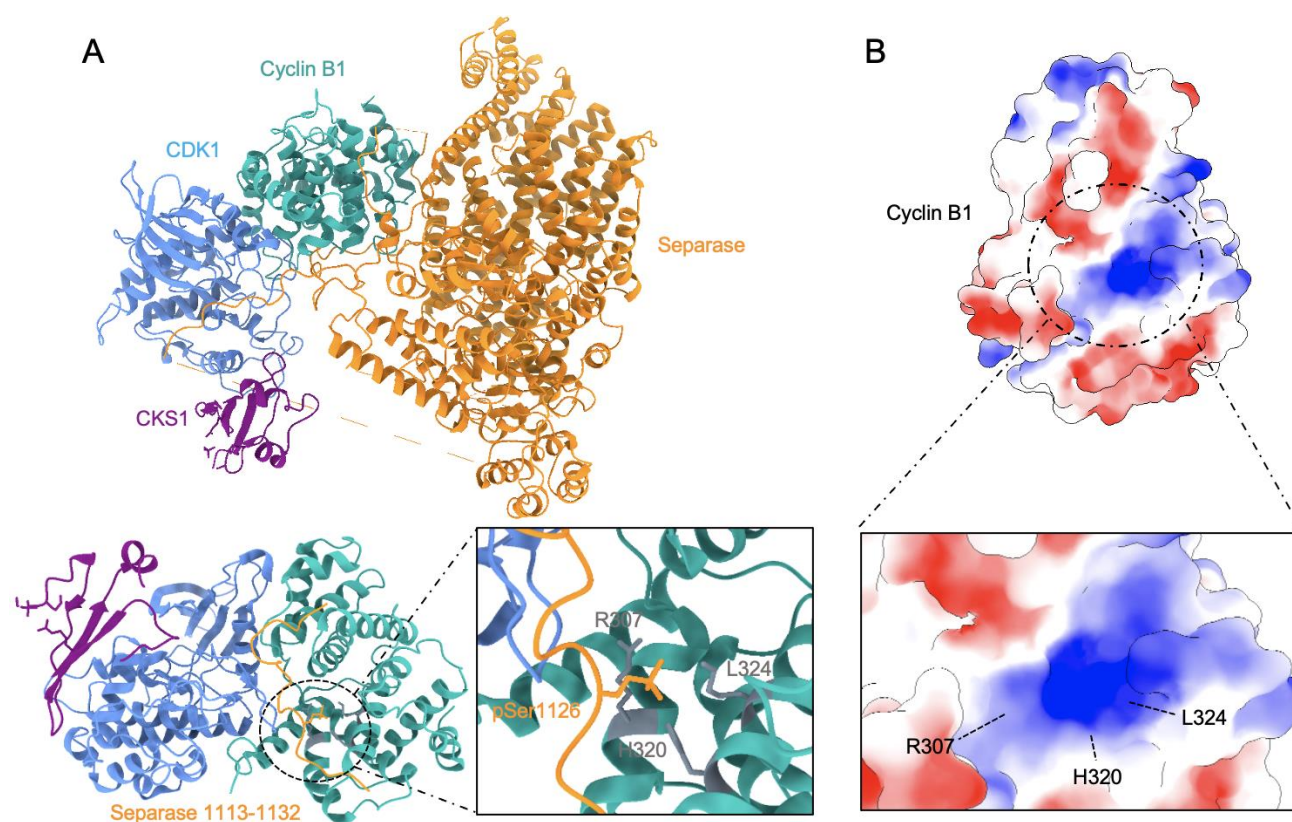


Figure 10. Identification of the cyclin B1 Phosphate-Binding Pocket (PBP)

A, Structure of the human separase in complex with cyclin B1/CDK1/CKS1⁶⁴ (PDB:7NJ0) resolved by cryoEM, with a zoom on the cyclin B1 PBP and the specific recognition of the phosphorylated separase Ser 1126 by the key residues Arg307, His320 and Lys324. **B**, Representation of the cyclin B1 surface charges (PDB:7NJ0) with a zoom on the PBP. Residues involved in creating a positively charged interaction region containing PBP key residues.

Recently, our research group provided a structural rationale for the previously observed inhibition of the protein separase by the cyclin B1/CDK1/CKS1 complex (CCC)^{64,79} (Figure 10. A). Separase, when phosphorylated at Ser1126, is recognized by cyclin B1 through a newly identified phosphate-binding pocket (PBP). Since then, another study also showed the importance of the PBP for the interaction and activity of cyclin B1/CDK1 with other substrates¹⁴⁷. Interestingly, previous structural studies, focusing on the interaction between cyclin B1/CDK1 with p107, already proposed that, by comparing the surface charges of cyclin A2/CDK2 with cyclin B1/CDK2, a positively charged region located in the first α helix of cyclin B1, and different to the one in cyclin E and cyclin A2, could be a potential new interaction site⁵⁹ specific to cyclin B (Figure 10. B). Due to the lack of a phosphorylated substrate bound to this positive patch, partly attributed to limitations in X-ray crystallography,

it could only be hypothesized that this region might contribute to the recognition of specific substrates. While Arg307 was mentioned other residues forming this PBP were not yet identified. Here, the cryoEM structure of separase bound to the CCC complex recently determined led to the identification of this novel specificity site formed by three crucial amino acids: Arg307, His320 and Lys324 (Figure 10. A and B). These three examples of the involvement of the PBP into the binding and regulation of proteins by cyclin B1 led us to hypothesis if the cyclin B1 PBP had a broader range of interacting proteins and if its integrity was necessary for the activity of CDK1 underlying a never described new CDK1 regulatory pathway.

During G2/M transition, and lasting until metaphase, phosphorylation levels of p21 rise responding to DNA damage control and the metaphase-to-anaphase transition checkpoint. At the same time, the levels of phosphorylated cyclin B1 and active CDK1 also increase drastically. It has been established since a long time that p21 is an inhibitor of CDK1, but some research groups also described the discussed CDK-releasing role of p21 phosphorylations. In the light of the newly described cyclin B1 PBP and its potential to interact with many phosphorylated residues and the existing literature on p21 phosphorylations and their key role in CDK1 activity¹²⁸, I hypothesized that cyclin B1 could be recognized in a non-inhibitory way by p21 through its PBP and furthermore that the pocket is essential for kinase activity regulation and correct advancement into mitotic phases.

To answer these questions, structural information on the binding mode of p21 to the CCC (CCCP21) is desired and I chose to answer this by using cryoEM, in combination with mass-spectrometry and functional analysis using kinase activity assays. Therefore, this work provides an important structural and functional framework on p21-mediated regulation of cyclin/CDK activity.

RESULTS

1. Protein analysis

1.1. The CCCP21 complex assembly

The first step in this project was to express, purify, and establish an *in vitro* reconstruction of the cyclin B1/CDK1/CKS1/P21 complex or CCCP21 complex. The expression and purification of CCC was performed in insect cells in order to get highly active complexes. Indeed, previously to the work performed in our group, assembling active cyclin B1/CDK1 was a tedious task. It involved the separate expression of each protein followed by an *in vitro* phosphorylation of the CDK1 T-loop Thr161 by the CAK complex and, once the reaction was completed, followed by the removal of the CAK complex before being able to use active cyclin B1/CDK1. What is known and established now is that insect cells naturally express a CAK complex similar enough to the human one to be able to phosphorylate CDK1 *in vivo*^{64,148}. Because p21 is a CDK inhibitor and CDKs show high sequence similarity in eukaryotes, expressing p21 in excess in insect cells would lead to a pause in mitosis and eventually result in cell death. Therefore, we decided to express p21 in bacteria along fused to a solubility tag that facilitates the solubility of the recombinantly expressed protein in isolation before mixing the protein with purified CCC complex (Figure 11. A). It is important to note here that the purified p21 protein used to form the complex is a phosphomimetic mutant replacing Thr57 with a glutamate residue (T57E). This mutation located in the K domain allows stronger binding of p21 to CDK1 while an alanine substitution (T57A) abolishes this interaction¹²⁸. First, we decided to keep the Maltose-Binding Protein (MBP) tag on p21 in order to increase the molecular weight of the recombinant protein, potentially also enhancing the chances to be visualized and analyzed with cryoEM. The CCCP21 complex elutes on an analytical size exclusion chromatography at the expected elution volume, showing a mono-disperse elution profile that indicates that one molecule of p21 binds to one CCC complex to form a stable quaternary complex (Figure 11. B). To extend our studies and remove the possibility of non-specific interactions with the tag, we also prepared complexes in which the tag was removed. Another round of size-exclusion chromatography run was performed in order to remove the cleaved tag, however we noticed that the complex became slightly more unstable, and a fraction of the complex formed soluble aggregates that

run in the void volume of the SEC run (Figure 11. C). Nevertheless, we could get stable CCCP21 complex and used it for most of the subsequent studies.

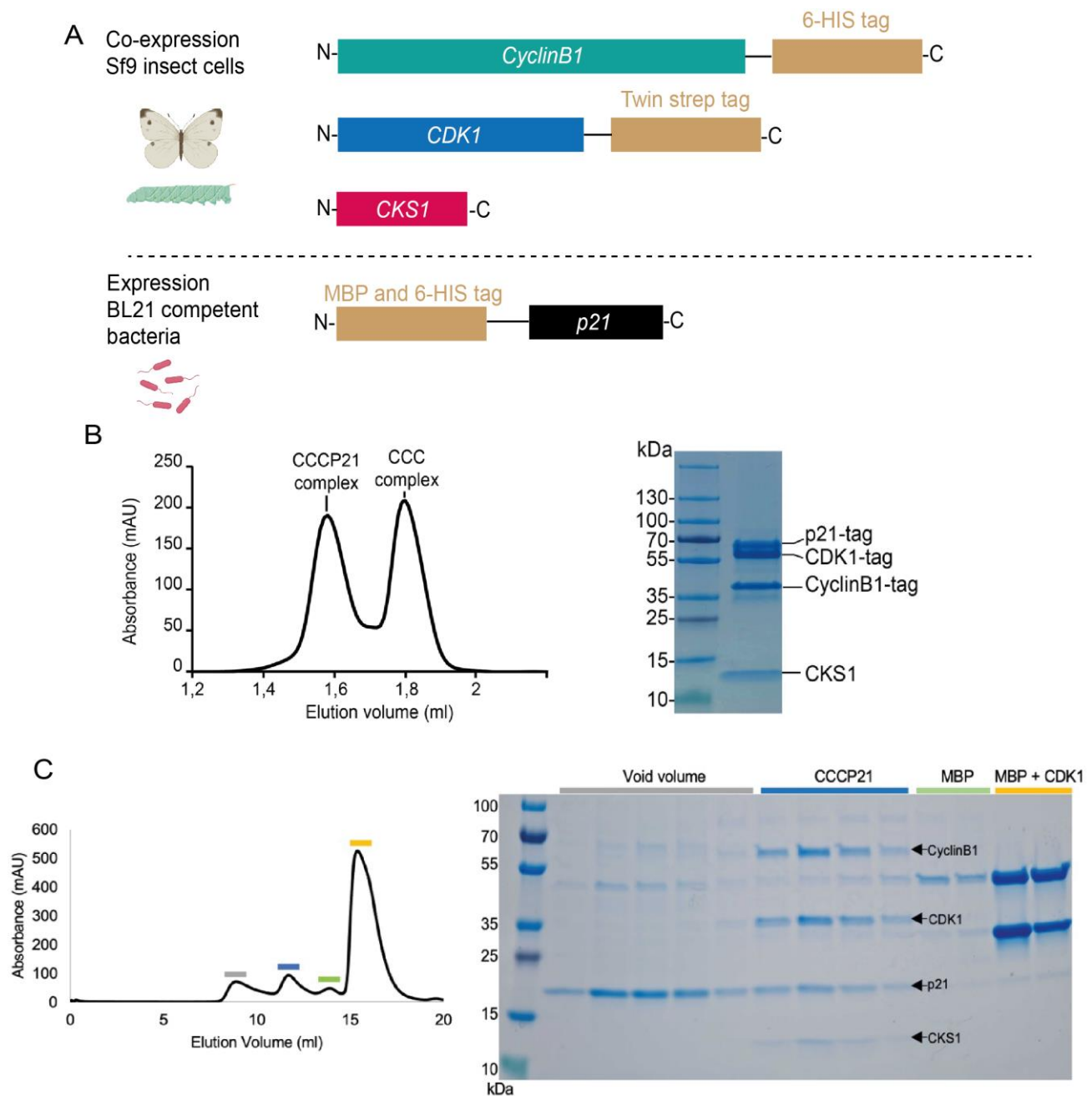


Figure 11. Expression, purification, and *in vitro* reconstruction of a CCCP21 complex.

A, cyclin B1, CDK1 and CKS1 are co-expressed in insect cells. *In vivo* they assemble into an active complex that can be purified. The protein p21 is recombinantly expressed in bacteria fused to an N-terminal MBP-6xHIS tag. **B**, Elution profile of an analytical size exclusion chromatography run of purified CCC mixed with p21 (with tags), shown on the left panel. Two main elution peaks can be observed. The stable CCCP21 complex elutes at around 1.6 mL and is visualized on an SDS-PAGE Coomassie gel shown on the right panel. **C**, Elution profile of a size exclusion chromatography run of the CCCP21 complex after tags are removed (left panel). Each peak, annotated in different colors, was loaded on a SDS-PAGE Coomassie gel visible on the right. A stable CCCP21 complex is eluting at around 11.5 mL.

1.2. P21 interacts with the CCC complex even when cyclin B1 phosphate-binding pocket is mutated

To understand the potential role of the cyclin B1 PBP for p21 binding, the interaction between p21 and the CCC with a mutated PBP was tested. When comparing the amino acid sequences of different cyclins by sequence alignments, it becomes apparent that the PBP key residues RHK (Arg307, His320 and Lys324) are unique to cyclin B1 (Figure 12. A). Comparing cyclin B1 with cyclin A2, both mitotic cyclins capable of binding CDK1, illustrates that the critical residues of the cyclin B1 PBP, namely Arg307, His320 and Lys324 are respectively changed into a threonine, a glutamine and a methionine (TEM) in cyclin A. Therefore, residues in cyclin A are negatively charged compared to the positively charged PBP residues in cyclin B1. Sequence comparison between the cyclin B1 PBP to other non-mitotic p21-targeted cyclins reveals that residues and charge drastically changes. It seems therefore that the residues and the charge of the cyclin B1 PBP are specific to this protein. Next, we wanted to test the implication of those residues on separase inhibition. To this end, we generated cyclin B1 PBP mutants, which included a cyclin B3-like PBP (Arg-Leu-Arg, RLR), a cyclin A2-like PBP (Thr-Glu-Met, TEM) and a mutant where the charges are completely inverted (Glu-Phe-Glu, EFE) and tested if these mutant are able to interact with a catalytic dead mutant of separase (Cys2029Ser) fused to securin that lacks the first 141 residues (securin fusion to the N-terminus of separase stabilizes the protein complex without diminishing the protease activity of the caspase). I first confirmed that all cyclin B1 mutants were soluble by expressing the CBD of the protein without its partner CDK1. Here, no difference in the expression and purification of the different protein versions could be observed. After, we co-expressed the full-length cyclin B1 mutants with CDK1 and tested their interaction with separase (Figure 12. B). To our surprise, none of the mutants could interact with separase, except the intact cyclin B1 PBP (WT). This result prompted the question if such cyclin B1 versions are capable of binding p21. Therefore, we generated CCC mutants (TEM and EFE) and performed SEC runs, in order to test if the CCCP21 complex was forming a stable complex in the presence of those mutations (Figure 12. C). Experiments were performed in the context of purified cyclin B1/CDK1 complexes, because it is known that p21 has a low affinity to monomeric CDK1 and cyclin B1¹⁴⁹. Contrarily to separase, p21 can bind TEM and EFE cyclin B1 mutants in a similar way to the WT protein. The protein p21 has other contact sites with cyclin B1 and CDK1 that are described in the introduction, and it

seems that mutating the PBP of cyclin B1 does not interfere with those. However, these results do not give information about the activity of the complex when the pocket is mutated, which will be investigated in later parts of this thesis.

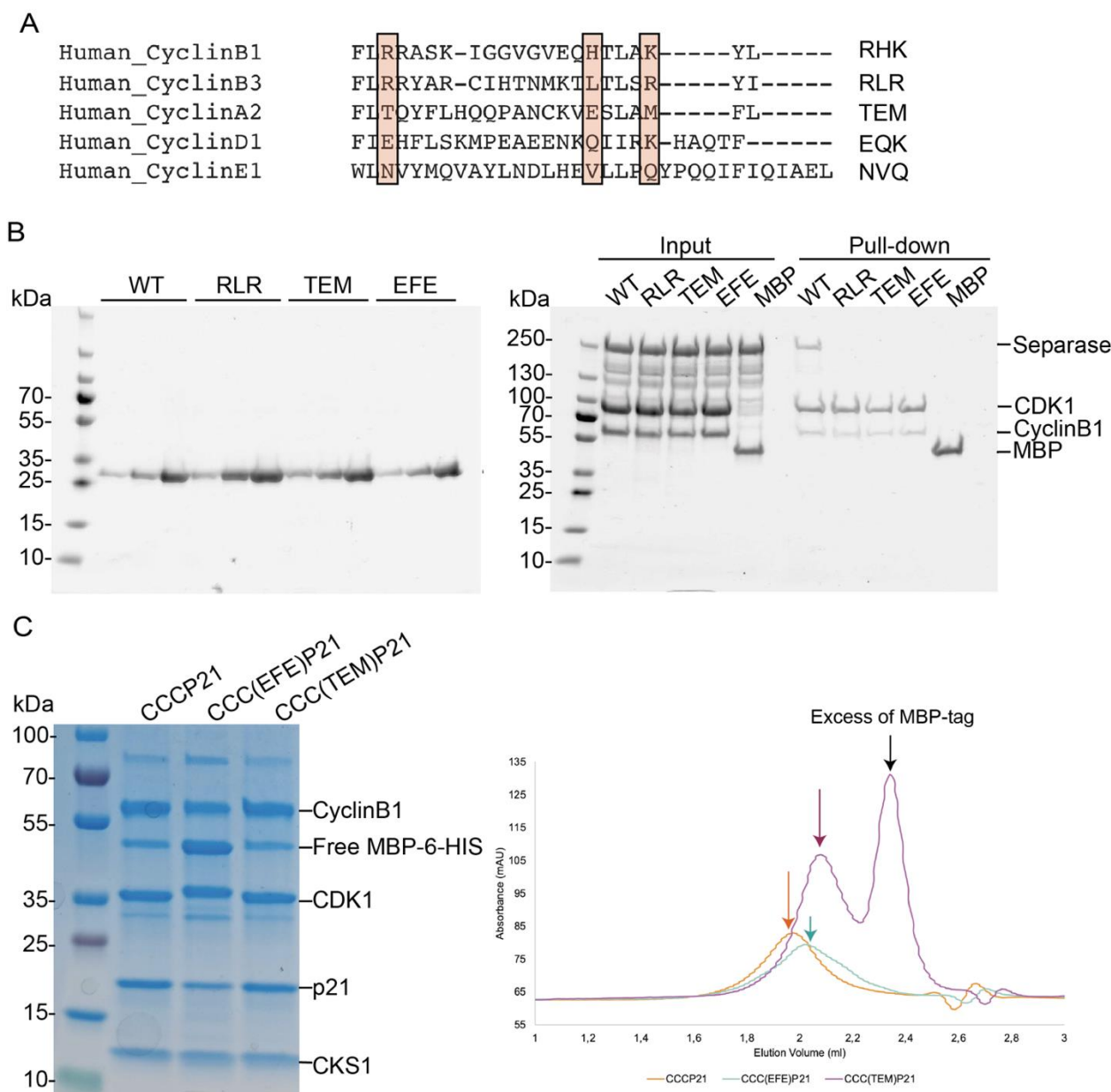


Figure 12. The cyclin B1 phosphate-binding pocket is necessary to bind separase but not p21

A, Sequence alignment of different human cyclins. Residues corresponding to cyclin B1's phosphate-binding pocket are shaded in orange in the MSA and summarized on the side. This sequence alignment has been performed with Clustal Omega. **B**, The left panel shows an SDS-PAGE of the elution fractions of purified cyclin B1 (144-433) with an intact PBP (WT) or a mutated PBP (RLR, TEM and EFE). For the pull-down experiment shown in the right panel, WT, RLR, TEM and EFE cyclin B1 was co-expressed and co-purified with CDK1 fused to an N-terminal MBP tag that was used to pull-down human separase (C2029D, catalytic dead mutant) fused to securin (lacking the first 141 residues). Input and pulled-down experiments were loaded on an SDS-PAGE gel. **C**, cyclin B1 EFE and TEM mutants were co-expressed with CDK1 and CKS1 and incubated with p21 to form a complex, and the results of the purification was loaded on an SDS-PAGE gel on the left panel. The purified complexes were analyzed using an analytical size exclusion chromatography (right panel). CCC(EFE)P21 and CCC(TEM)P21 elute at the same elution volume than CCCP21 (around 2 ml).

1.3. CCC interacts strongly with the CDK1-recognition domains of p21

To further investigate the intricacies of p21's interaction with the CCC complex, we decided to divide p21 into three fragments based on its domain interaction with its targets (Figure 13. A). The first fragment (1-35) contains the first Cy domain that interacts with cyclin B1. The second fragment (36-120) contains the motifs interacting with CDK1 and responsible for its inhibition. Finally, the third fragment is the C-terminal part of p21 (121-164) containing the second Cy domain. The three fragments of p21 were separately cloned into an *E. coli* expression vector that contains an N-terminal MBP tag. The three recombinant proteins were well expressed, and large quantities of soluble proteins could be purified (Figure 13. B) and used for the binding studies. To test whether the three p21 constructs are able to interact with the wild-type CCC complex as well as the TEM and EFE mutants we performed pull-down experiments (Figure 13. C). Surprisingly, neither the first p21 fragment (aa 1-35), nor the second p21 fragment (aa 121-164) could pull down CCC WT, CCC TEM or CCC EFE. Only p21(aa 36-120) could pull down all three complexes in quantities. These results show that binding of p21 to the CCC complex is likely not mediated by the PBP. Further, I performed size exclusion chromatography experiments to analyze binding of the three fragments of p21 to the CCC complex and compared those runs to a preformed CCCP21 complex (Figure 13. D). The SEC runs show that only CCCP21(36-120) shifts to an earlier elution volume similar to the preformed CCCP21 complex and the corresponding SDS-PAGE analysis confirmed the pull-down assays (data not shown). Finally, I also performed pull-down experiments using CDK1 in isolation. It has been reported that p21 has a low affinity for monomeric CDK1 and, consistent with this observation, I could not pull it down using full-length p21. Interestingly, p21(36-120) can pull-down CDK1 in a quite equimolar way (Figure 13. E) to which we can conclude that this fragment of p21 has a stronger binding to CDK1 than the full length p21. These results show, that only the p21 construct containing the interacting domains with CDK1 and the helix LH pulls down the CCC complex efficiently, and that this interaction is independent of the PBP. The N-terminal and C-terminal fragments contain only unstructured regions of p21, and it is likely that binding to the CCC complex is impaired because it is lacking the main structured interaction with CDK1. We hypothesize from these results that the interaction of the unstructured part of p21 is at least in part dependent on the docking of the cyclin B1/CDK1-bridging helix. In accordance with this speculation, it was shown *in vitro* that altering the length of p21 LH segment abolishes the capacity of p21 to bind to certain cyclin/CDK complexes and impairs cell cycle progression *in vivo*¹²⁵. The LH sub-domain

seems therefore to act like a glue on cyclin B1/CDK1 allowing the rest of the protein to bind, thereby regulating cyclin/CDK complexes.

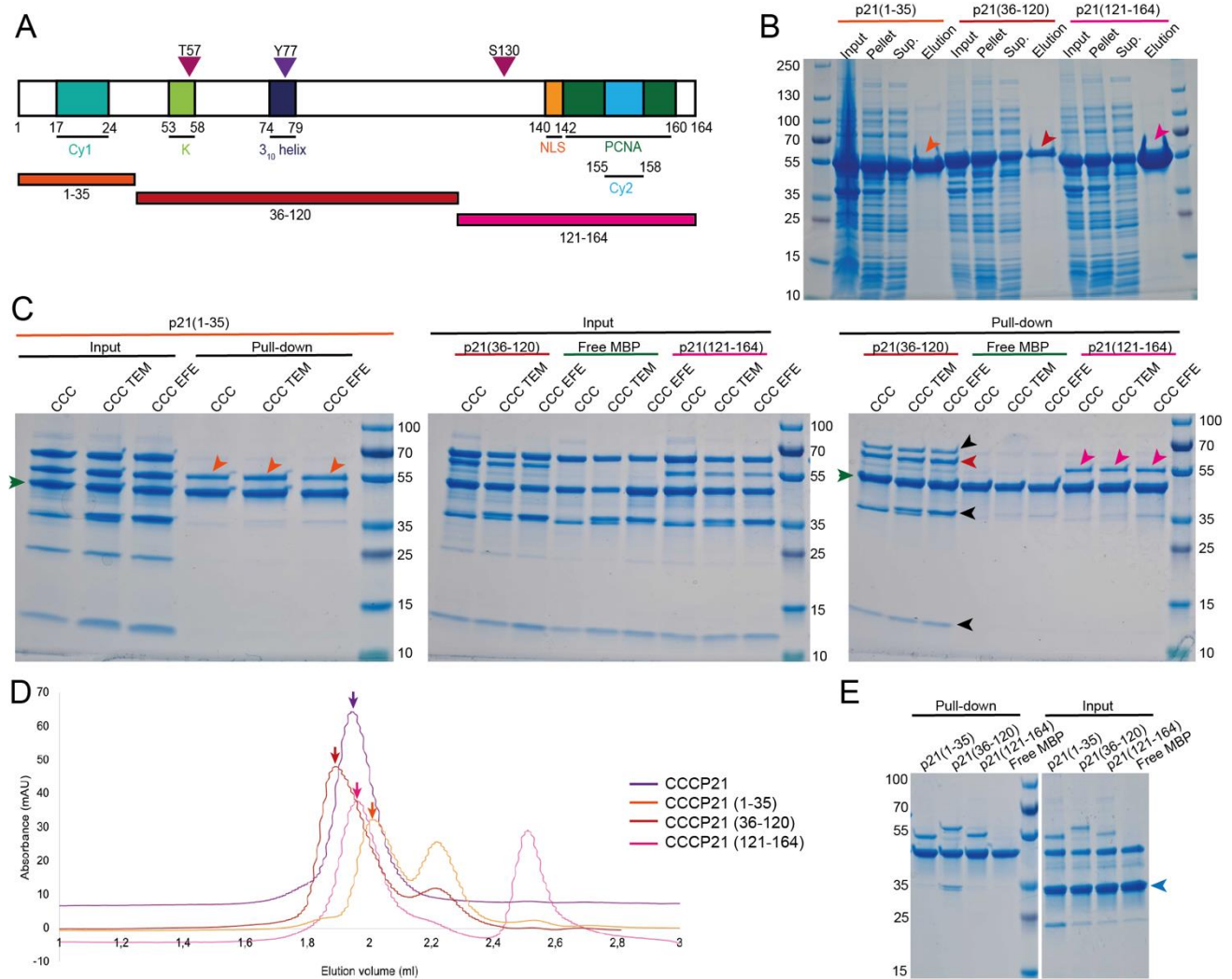


Figure 13. P21 residues 36-120 efficiently bind to CCC WT, TEM and EFE.

A, Schematic representation of the full-length human p21 with the domains annotated in different colors. Cy1 (cyan) and Cy2 (blue), K (green), 3₁₀ helix (dark blue), NLS (orange) and PCNA (green). The phospho-sites (CDK1 sites) Thr57 (T57), Ser130 (S130) and the well-conserved Tyr77 (Y77) are highlighted by triangles. This scheme is used to visualize the different fragments of p21, expressed recombinantly with an N-terminal MBP tag: 1-35 in orange, 36-120 in red and 121-164 in pink. **B**, p21 1-35, 36-120 and 121-164 constructs were recombinantly expressed with an MBP tag in *E. coli* and the different steps of the purification of these proteins (input, pellet, supernatant and elution) were loaded on an SDS-PAGE gel. The small colored arrows mark the purified soluble proteins used for further experiments. **C**, Purified proteins were used to perform a pull-down experiment with the CCC WT, CCC TEM and CCC EFE complexes. The green arrow indicates free MBP protein, and the black arrows indicate CCC complexes. Orange, red and pink arrows represent p21 1-35, 36-120 and 121-164, respectively. **D**, MBP-tagged p21 1-35, 36-120 and 121-164 protein constructs were incubated with CCC complexes and analyzed by analytical size exclusion chromatography and compared to preformed CCCP21 complex, serving as reference. CCCP21 elutes at around 1,9 ml (dark purple). P21 1-35 (in orange) and 121-164 (in pink) did not elute with CCC, indicated by two separate main peaks, one eluting at around 2 ml representing the CCC complex and a second peak, representing p21 versions. For p21 36-120, a shift is observed with an elution of CCCP21 (36-120), in red, at around 1,8ml. **E**, Equivalent to C, p21 recombinant constructs were used to perform a pull-down assay with purified CDK1 (marked by a blue arrow). The inputs and elutions were loaded on an SDS-PAGE gel.

2. Usage of AlphaFold-generated structure predictions as a tool to understand the interactions between p21 and the CCC complex

During the second year of my thesis, AlphaFold 2, an artificial intelligence-based structure prediction tool¹⁵⁰ became publicly available and helped us in our understanding of p21's interaction with the CCC complex. Computational methods to predict protein 3D structures from their primary sequence are not new to 2021, but AlphaFold became a widely used tool than any other structure prediction software capable of predicting the structural organization of a vast variety of proteins with near experimental accuracy. In general, pre-existing structural information deposited in the Protein Data Bank (PDB), together with Multiple Sequences Alignment (MSA) information on a target protein, allow a more accurate prediction by AF2¹⁵¹. With no existing structural information of a protein homologue, the accuracy of AlphaFold structure predictions is generally lower compared to highly conserved proteins for which structures of homologues has been determined. Similarly, the interaction with a protein partner can be more accurately predicted if MSA and structural homologues exist. This is exactly what we observed when we used AF2 to predict the binding of p21 to the CCC complex. Because p21 resembles p27 and a structure of the KID domain exists for both¹³¹, AF2 could predict how that domain interacts with the CCC complex but failed to predict the rest of the protein's folding (Figure 14. A). If we remove the aberrantly predicted parts of p21 (1-10 and 92-164) and the unstructured regions of cyclin B1(1-135), we can make conclusions and hypothesis out of the predicted CCCP21 complex (Figure 14. B). First, we can see that AF2 does not predict p21 to interact with CKS1. In this setup, p21 is unphosphorylated (except for the presence of the phosphomimic Glu instead of the Thr57) and we cannot exclude that a phosphorylated residues could interact with the PBP of CKS1. Second, AF2 predicted the formation of an α helix from the Ser26 to Trp48 corresponding to the LH subdomain of the KID. If we zoom on the interactions between cyclin B1 and p21 (Figure 14. C) we can first see that AF2 predicts p21 to bind to cyclin B1 only through its first Cy domain (RRL motif, residues Arg18, Arg19 and Leu20) and did not predict interactions with the Cy2 domain located in the C-terminal part of the protein. We can see that the RRL motif faces the MRAIL motif of cyclin B1 and p21 Arg18 is nicely recognized by cyclin B1 Gln211 which has been shown before⁵⁹. Moreover, p21 Lys20 nicely engages with the adjacent cyclin B1 hydrophobic patch which is the main interaction of the RRL motif to cyclin B1. If we now focus on the interaction of p21 with CDK1 (Figure 14. D), AF2 predicted the K

domain right on top of the N-lobe of CDK1, folding this loop in a very similar way to the existing p21 structure with cyclin D1/CDK4¹³¹ (Figure 8. C) packing and rearranging the structure of CDK1 N-lobe allowing the conserved CIP/KIP aromatic residues Trp49, Phe51, Phe53, Phe63 and Trp66 to burry themselves against CDK1 β sheets. This interaction is believed to be stabilized by a hydrogen bond between Phe53 and CDK1 Tyr77¹³⁰. AF2 also predicts p21 to adopt a β sheet conformation from Phe62 to Val67 interacting directly with CDK1 and stabilized by hydrogen bonds between p21 Glu66 and Val68 and CDK1 Lys20 and Val18 respectively. The first 17 residues of CDK1 were predicted to be unstructured upon binding to p21 but we expected a β sheet from residues 4 to 13. The 3_{10} helix is predicted to enter the catalytic cleft of CDK1 and its ATP-binding site. The key role of the well conserved Tyr77 is to assure this binding and the inhibition by p21 by supposedly forming hydrogen bonds with CDK1 Glu81 and Leu83. Moreover, it is expected that the 3_{10} helix will be locked between CDK1 residues Lys33, Phe80 and Ala144. To this point, the AF2 predictions were going in the same direction as the published literature. This is mainly due to the highly conserved residues between p27 and p21 and similar structural binding and inhibition performed by their KID on cyclin/CDK complexes but where it was shown that p27 forms a hydrogen bond with CDK2 Gln131, p21 was not predicted to form the same interaction with CDK1 Gln132. The rest of p21 was not predicted to interact with cyclin B1/CDK1 with high certitude and this is where an experimentally generated structure becomes necessary to understand furthermore this regulation.

We can conclude from those results that AF2 does a remarkably good prediction of the interaction between p21 (11-91) and the CCC complex and these predictions go in the same direction as already published data and existing structures of p27 with other cyclin/CDK complex but also p21 with cyclin D1/CDK4. Nevertheless, AF2 fails to find interactions with the rest of the protein and, more importantly with the C-terminal part of the protein. These first observations of what could be the structure of p21 are very valuable to design the future of the project but need to be validated by experimentation and, most importantly, structural determination.

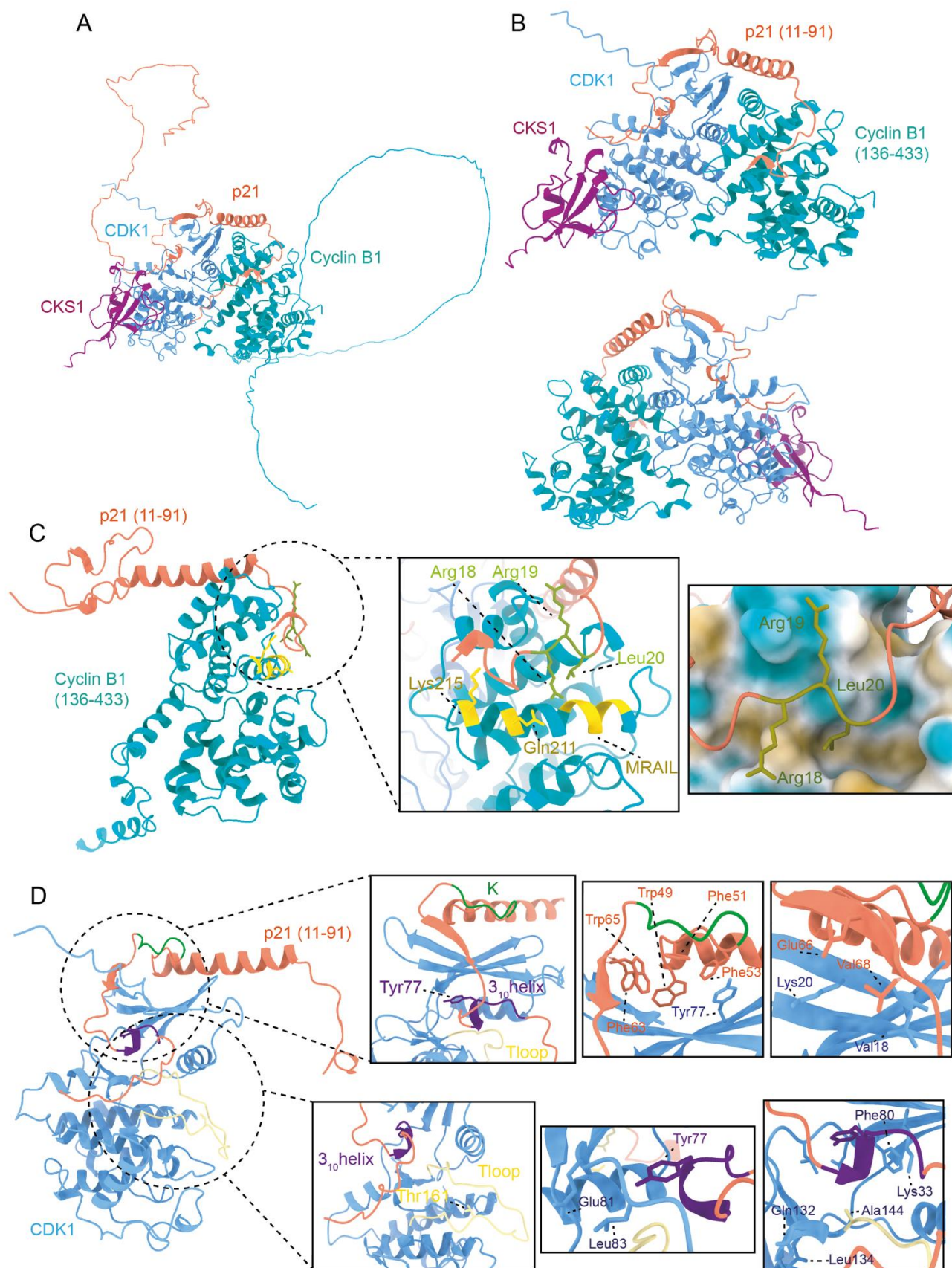


Figure 14. AlphaFold2 prediction of the CCCP21 structural organization

A, Predicted model of the structure of full-length CCCP21. A p21 sequence mimicking threonine 57 phosphorylation was used as input (Thr57Glu, T57E). The different proteins are shown in different colors, with CDK1 in blue, cyclin B1 in cyan, CKS1 in purple and p21 in orange. Large regions of cyclin B1 and p21 are

predicted to be unstructured with no predicted interactions with other CCCP21 complex members. **B**, Complex structure with the first 135 residues of cyclin B1 and residues 1-10 and 91-164 of p21 not shown. The front and back views of CCCP21 are displayed. **C**, Zoom on the predicted interactions between p21 and cyclin B1. Residues of the p21 RxL motif (Arg18, Arg19 and Leu20) are shown in green. The MRAIL hydrophobic docking site in cyclin B, that is known to be important for substrate recognition, is displayed in yellow. Gln211 of cyclin B1 is predicted to interact with the first arginine (Arg18) of the p21 RxL motif, while Lys215 is part of a positively charged patch. The RxL motif nicely engages with cyclin B1's hydrophobic patch of this region, with the surface being colored in brown. **D**, Zoom on the predicted interactions between p21 and CDK1. In green is colored p21 K domain, in purple p21 3_{10} helix domain and in yellow CDK1 activation segment, the T-loop. This predicted interaction region can be divided into two. The first one is the part where p21 interact with the N-lobe of CDK1 in which can be found the K domain. Right after the connecting LH domain and the K domain, a beta sheet is predicted to form lying on CDK1 beta sheets exposing residues which will hydrogen bond with CDK1. This rearrangement allows the packing of p21 on CDK1 and multiple hydrogen bonds between CIP/KIP conserved aromatic residues and CDK1. The second predicted interaction region is at CDK1 active site. The 3_{10} helix will enter the ATP binding site allowing the well conserved Tyr77 to hydrogen bond with CDK1 and the locking of this domain into the active site.

3. Understanding the structure of CCCP21 with Electron Microscopy

Electron microscopy (EM) is a powerful tool to determine protein structures and ideally link these structures to their functions in the cell. Transmission electron microscopes (TEM) use an electron beam that goes through and interacts with a specimen before the electrons are recorded by a detector to form an image of the sample. Unlike X-rays or neutrons, electrons can accelerate, scatter and being focused by an electromagnetic field, a magnetic lens, to form a high-resolution image of the studied specimen. Traditionally two methods of TEM microscopy have been used in the past decades: Negative-staining microscopy and cryogenic electron microscopy. Negative-stain TEM is a technique that was already developed in the early 1950s, however still proves very useful to date. In negative-staining microscopy, a protein sample is applied onto a support layer, typically a thin sheet of amorphous carbon that is covering the EM grid. Once the protein sample is attached to the support layer, excess of the protein solution is removed, and staining solution is applied. The staining solution typically consists of salts of heavy metals like uranium, tungsten or molybdate. This technique allows a rapid visualization of the protein shape or conformation (in case of large macromolecular machines) at room temperature but is limited to a final resolution of around 20 Å (due to the grain size of the staining solution). Due to the low resolution, this technique is mainly used as a quality check tool, visualizing the integrity of the purified sample. Nevertheless, negative stain TEM is still used to determine protein structures and arrangements of larger macromolecules like viruses and the information collected can still drive research discoveries¹⁵².

Cryogenic electron microscopy (cryoEM) uses a focused electron beam to irradiate an EM grid on which a protein sample has been applied and immobilized in cryogenic conditions.

To prevent electrons from being scattered by other gas molecules or dust, it is important to impose a high vacuum on the electron microscope. Further, to avoid any damage on the protein sample that could be induced by the cryogenic conditions or the high vacuum in the microscope, a method termed plunge-freezing or vitrification¹⁵³ was developed. Vitrification of biological macromolecules, deposited on an EM grid, is initiated by the removal of excess of water (blotting). When the blotting is done, the grid is rapidly plunged into liquid ethane cooled by liquid nitrogen. As a result, a thin layer of amorphous ice is formed in which proteins are embedded in their native state. During this fast vitrification or freezing procedure, the physical and chemical properties of water molecules remain similar to liquid water and as a consequence the protein structure remains intact and comparable to that in solution¹⁵⁴ (Figure 15. A). Another problem linked to this technique is the damage of the sample induced by electron radiation. Applying a high dose on the sample will induce radiation damage and therefore limit the final achievable resolution of the sample, whereas a low electron dose reduces the amount of detected protein information (signal) compared to the background (noise), also known as signal-to-noise ratio. A low signal-to-noise ratio will result in a poorly defined electron map and therefore drastically reduces the chances to reconstitute a high-resolution 3D model of the imaged protein. Reducing the temperature by flash-freezing the sample helps reducing the damage. However, key to overcome this problem was to Fourier transform low-electron dose produced images. This principle was demonstrated by recording electron diffraction images of 2D protein crystals that can be calculated into a high-resolution 3D reconstruction map of the protein¹⁵⁵. The same principle can be applied to non-crystallized proteins embedded in vitreous ice, if enough images (micrographs) with protein particles are collected to improve the SNR. In addition, as mentioned above, many protein particles are needed to also obtain different orientations (views) of a protein. Structural heterogeneity, including different conformations and shapes of the protein target might further require extensive data collection to obtain high-resolution information¹⁵⁶. Imaging as close as possible to the focus is advantageous to obtain high-resolution information, but at the same time this reduces dramatically the contrast of protein particles in ice when a low electron dose is applied. The Contrast Transfer Function (CTF) measures how much out of focus a micrograph is recorded and provides information on how the image is distorted due to the applied defocus.

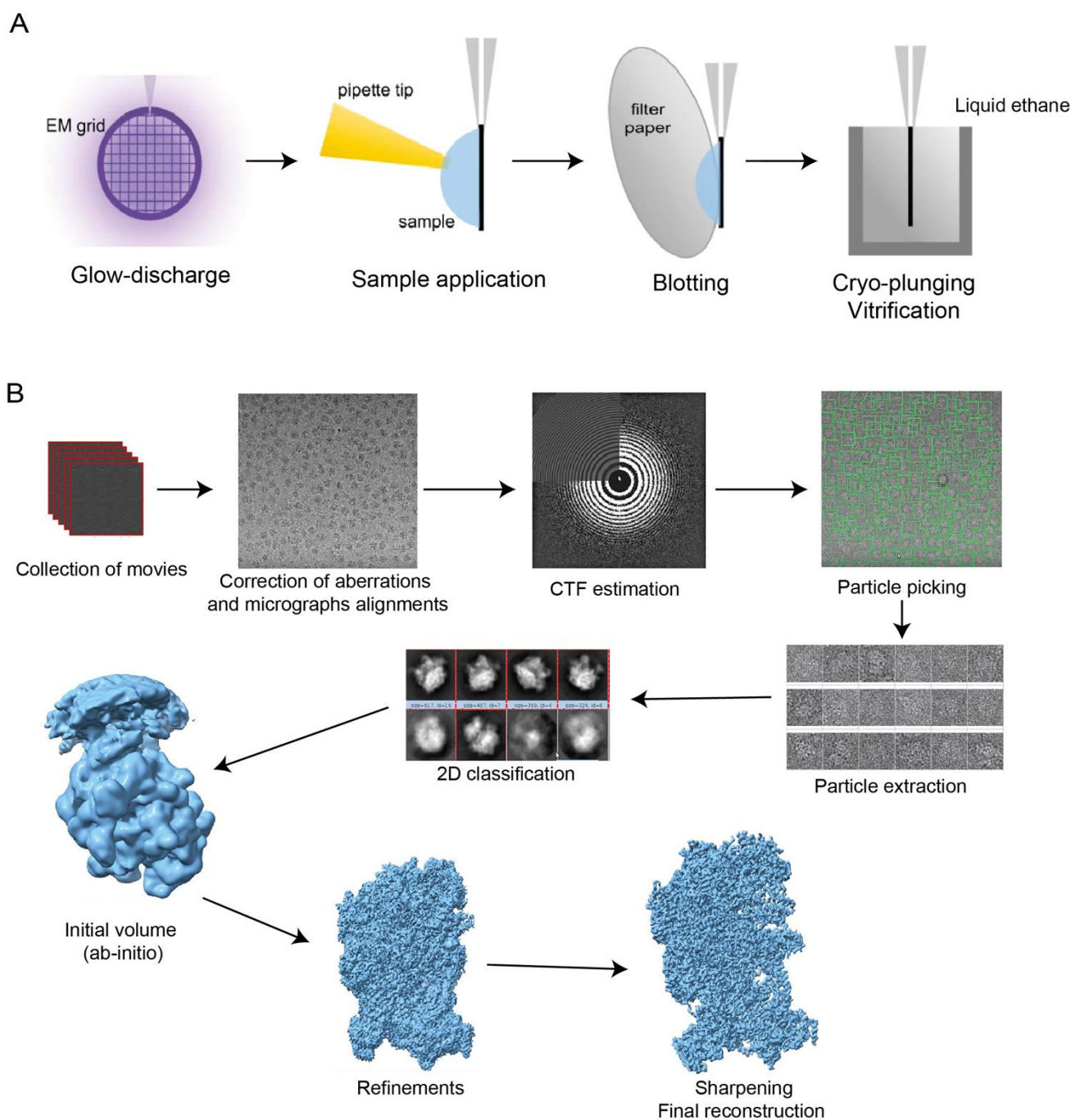


Figure 15. Typical workflow for cryoEM sample preparation and image processing

A, Modified from Koning *et al.*, *Nat. Com.*, 2022¹⁵⁷. Typically, EM grids are first glow discharged to render the grids hydrophilic. This step facilitates sample application onto the grid. A drop of sample containing the protein of interest is applied to the grid before the excess of sample is being removed by blotting with filter paper, resulting in a thin layer of buffer on the grid. Next, the grid is plunged into liquid ethane vitrifying the sample and the protein in a native state. The grid is ready to be analyzed in a microscope. Specific plunging conditions for the grids prepared during this thesis are described in the methods section.

B, Modified from Vilas *et al.*, *Chem. Rev.*, 2022¹⁵⁸. Movies are collected on an electron microscope before being motion corrected and aligned, and the CTF of each micrograph is estimated. From each micrograph, particles are picked and extracted before being classified into 2D classes. Selected particles are used to generate a first 3D density map (volume), also known as *ab-initio* volume. During advanced processing steps, this volume is refined and sharpened to improve the final resolution and solve the structure of the studied protein. To see the specific processing workflow used during my thesis, see the methods section.

The so-called “resolution revolution” in cryoEM was mainly driven by two technical developments: the invention of ultrafast direct electron detectors (that in contrast to older CCD camera record electrons directly as the name implies) and the development of novel computing software. These developments allow to record movies (that can be divided into frames), and therefore record and correct the motion of protein particles in between the frames. This motion correction procedure enables the formation of a final, high-resolution averaged micrograph¹⁵⁹ (Figure 15. B). Ideally, the protein particles that can be observed best on the averaged micrograph, are adopting many different orientations, so that the protein particles can be observed from different angles. Particles of a similar confirmation or protein composition (in case of protein complexes) that are frozen in similar orientations in ice are grouped into distinct classes, the so-called 2D classes. These 2D classes give already information on the integrity of the protein (complex), and if the signal is strong enough, features such as α -helices can be observed. A common problem in cryoEM however is, that particles like to adopt a preferred orientation in ice, a phenomenon named preferential orientation and that often leads to a compromised reconstruction of a 3D model. Modern cryoEM software allows to use 2D classifications to select for the particles that provide most information (intact particles with high contrast) on each micrograph. Once selected, thousands of these particles are then used to generate a density map, in which a final model is built.

3.1. Negative Staining of CCCP21

CCCP21 is a relatively small complex with a molecular weight of around 110kDa (cyclin B1 is 48kDa, CDK1 is 34kDa, CKS1 is 9,6kDa and p21 is 18kDa). Negative stain is a technique that does not allow to obtain high resolution information, and similar to cryoEM, protein targets of 150 kDa or smaller are difficult to analyze by negative stain EM. The smaller and less globular the protein is, the less visible it is on a micrograph. In our initial studies we decided to keep the N-terminal MBP-tag because we speculated that a flexible tag likely doesn't affect the 3D reconstruction and, in addition, if the MBP-tag is bound to the complex in a specific conformation it might even contribute to particle alignment. We applied the CCCP21 sample on an EM grid directly after the size-exclusion chromatography step and were able to observe nicely distributed protein particles on each micrograph (Figure 16. A). A small, 200 micrograph-containing dataset was collected and subsequently processed. I observed that the obtained 2D classes were blurry but always exhibit a triangular shape which

was expected, using existing structures as reference (Figure 16. B, C). As discussed above, the MBP-tag is attached to p21 by a long, flexible linker. We therefore assumed that the MBP-tag is likely flexible and will not affect the fold of the protein or binding of p21 to the CCC complex (Figure 16. C).

While we could not obtain a well-defined low-resolution map of the CCCP21 complex by negative stain, these results nevertheless demonstrated that the complex is likely monomeric and does not aggregate using such buffer conditions. Therefore, these results encouraged me to use this sample for future cryoEM studies.

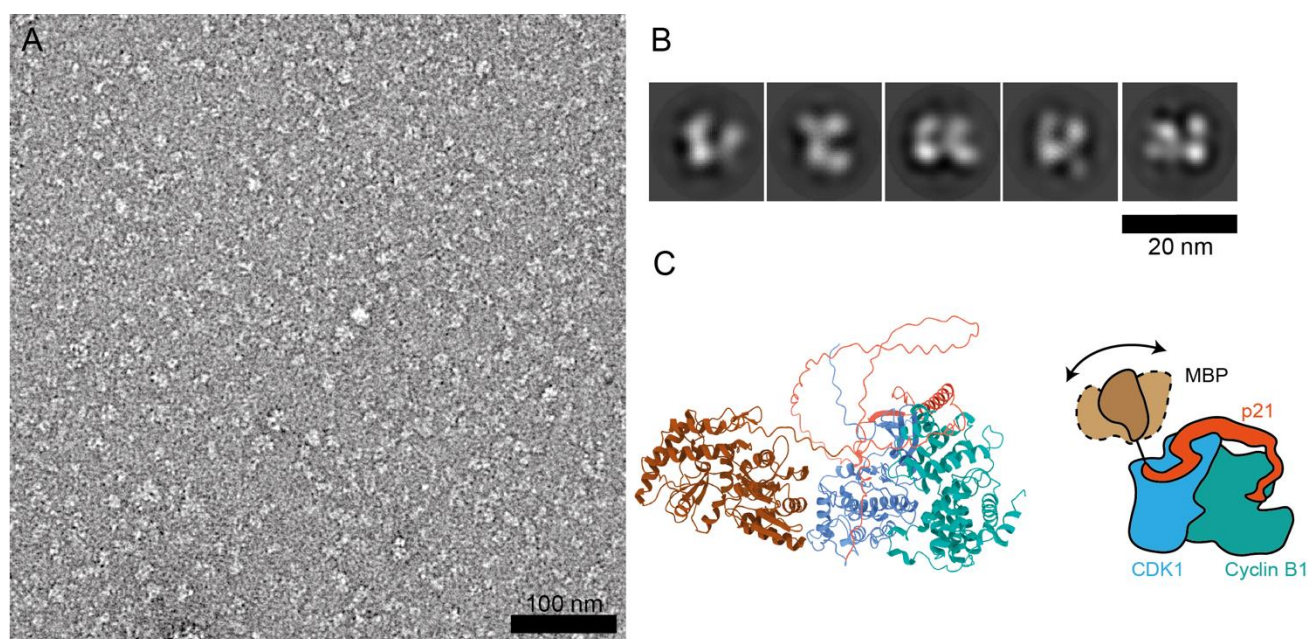


Figure 16. Negative staining of tagged CCCP21

A, Micrograph of tagged CCCP21 on an EM grid stained with uranyl acetate. The particles are visible in white as the background in stained. 230 micrographs were collected and analyzed. **B**, Selected 2 D classes of picked particles. Using these 230 collected micrographs, resulted in 71.930 picked particles that were used to generate 2D classes. **C**, On the left is an AlphaFold2 prediction of the tagged CCCP21 complex (CKS1 is not displayed). The MBP tag fused to p21 is shown in brown, p21 in orange, cyclin B1 in cyan and CDK1 in blue. On the right is a cartoon representation of the complex with an emphasize on MBP's potential movement ability on that complex compromising the quality of the averaged particles.

3.2. CryoEM of CCCP21

3.2.1. First dataset with tagged CCCP21

As a first try, we decided to use CCCP21 with the MBP-tag fused to p21. Analyzing small particles with cryoEM is a challenging task that has been repeatedly reported in the literature¹⁶⁰. Originally, single particle cryoEM was used only for big proteins or protein complexes (usually bigger than 200 kDa) because of the low SNR inherent to small proteins.

Recent developments in grid-preparation techniques, image acquisition but also image processing helped overcome this limitation and it is now established that “small” particles can be also analyzed by cryoEM to resolutions that allow the modelling of atomic coordinates. Nevertheless, small proteins or protein complexes remain challenging targets to be analyzed by cryoEM.

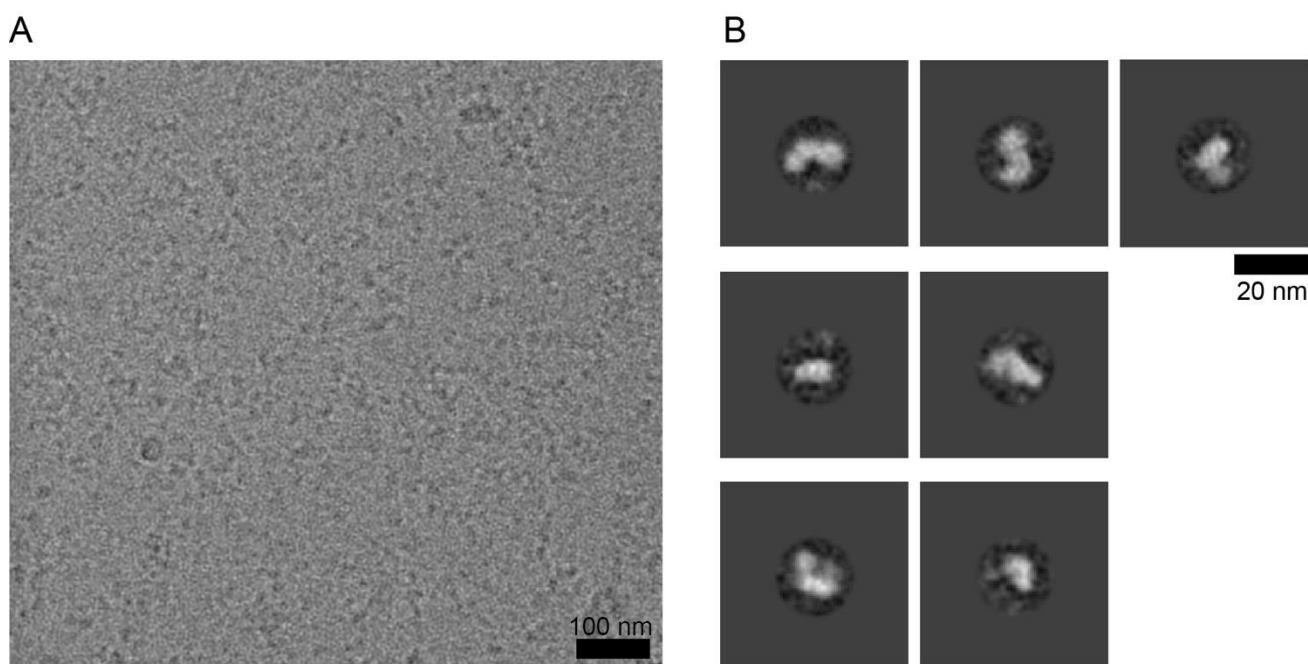


Figure 17. First cryo-EM experiments using tagged CCCP21

A, Representative micrograph of tagged CCCP21 particles embedded in vitreous ice and visualized by electron microscopy. 432 movies were collected and analyzed. **B**, cryoEM 2D classes of tagged CCCP21. A 20 nm circular mask is applied, reducing the noise from neighboring particles.

Here, the ideal sample preparation parameter (e.g. small proteins often adopt preferential orientations in ice during the vitrification procedure) and/or microscope parameters (e.g. the usage of a phase plate or an electron filter) need to be determined for successful structure determination. To facilitate structure determination by EM, we first imaged the sample with the highest molecular mass as possible, and therefore included the MBP tag. We managed to get nice micrographs that showed clearly separated single particles and were able to pick those particles (Figure 17. A). Unfortunately, we did not obtain 2D-classes that showed strong protein features, (Figure 17. B) but instead, only obtained poorly aligned particles with no visible features (this often results in streaky lines around the particles, because protein alignments are not very accurate). Moreover, the aligned particles seem to represent a very heterogeneous sample, possibly induced by the flexible N-terminal MBP tag. We tried to improve the quality of the picked particles by changing different

parameters during grid preparation but also to change the type of EM grid used (Table 1.), but we did not observe any major difference between these experiments. We also explored the possibility to use grids coated with an additional support layer, the naturally hydrophilic graphene oxide¹⁶¹. In an ideal case, particles are randomly embedded in the thin ice layer, but empirical studies show that particles, especially small protein particles, are often drawn to the air-water interphase. Interactions at the air-water-interface can damage the integrity of protein particles by also induce preferential particle orientation in the ice. Additional support layers were developed to overcome this problem in cryoEM^{162,163,164}, by attracting particles into the ice layer which helps to randomize their orientation and showed success in the past in our research team⁶⁴. Even though graphene oxide is known to have low background noise, we did not see CCCP21 particles in the ice and decided to not continue with it. Therefore, we decided to proteolytically remove the tag and continue with a 110kDa CCCP21 without any tag.

EM-grid type	Protein concentration	Blotting time (seconds)	Additional move	Glow-discharge (seconds)	Temperature/% humidity
Graphene-oxide 1.2/1.3	100µM 200µM	10s delay + 2s blotting	1mM	40s	10°C/85%
Graphene-oxide 2/2	100µM 200µM	10s delay + 2s blotting	1mM	40s	10°C/85%
Quantifoil carbon film on gold 1.2/1.3	0,1 mg/ml 0,2 mg/ml	2s delay + 1,5s blotting	1mM	40s	10°C/85%
Quantifoil carbon film on gold 2/2	0,1 mg/ml 0,2 mg/ml	2s delay + 1,5s blotting	1mM	40s	10°C/85%
Lacey carbon film on copper	0,1 mg/ml 0,2 mg/ml	2s delay + 1,5s blotting	1mM	40s	10°C/85%

Table 1. Different conditions to improve the quality of cryoEM grids with tagged CCCP21
Different grid types and blotting conditions were screened in order to improve the quality of the collected datasets.

3.2.2. Datasets without the tag

We started by preparing grids using different protein concentrations and noticed that the optimal concentration to obtain micrographs that were evenly covered with proteins was around 0.1 mg/ml of CCCP21. This condition was used for all subsequent experiments. The first thing we noticed when analyzing our grids were the strong differences in ice thickness inside the holes used for data collection (Figure 18. A). Thick ice is usually observed as

greyish, while very thin ice or empty holes appear to be white. Differences in ice thickness can typically also be observed within a hole, with thicker ice on the side and thinner ice towards the center, which appears like “donut-shaped” holes. Such holes should often be selected for data collection, because they provide a gradient of ice thickness in one micrograph and therefore improve the chances to collect protein particles in ice just thick enough to embed the protein particle, but thin enough to provide a strong signal-to-noise ratio. As common in cryoEM, we also observed the appearance of black spots on our grids, which are not ice contaminations, but protein aggregates. To reduce protein aggregation on the grid, centrifugation of the protein sample prior plunge freezing was helpful, however some aggregates will always remain. The micrographs obtained from these grids were very promising and, after processing, we managed to obtain nice and promising looking 2D classes (Figure 18. B). Those classes show that the particles are relatively homogenous, equally distributed and adopt various orientations in ice. Most importantly, these 2D classes (averages of hundreds to thousands of protein particles) show typical protein features such as alpha helices (Figure 18. B).

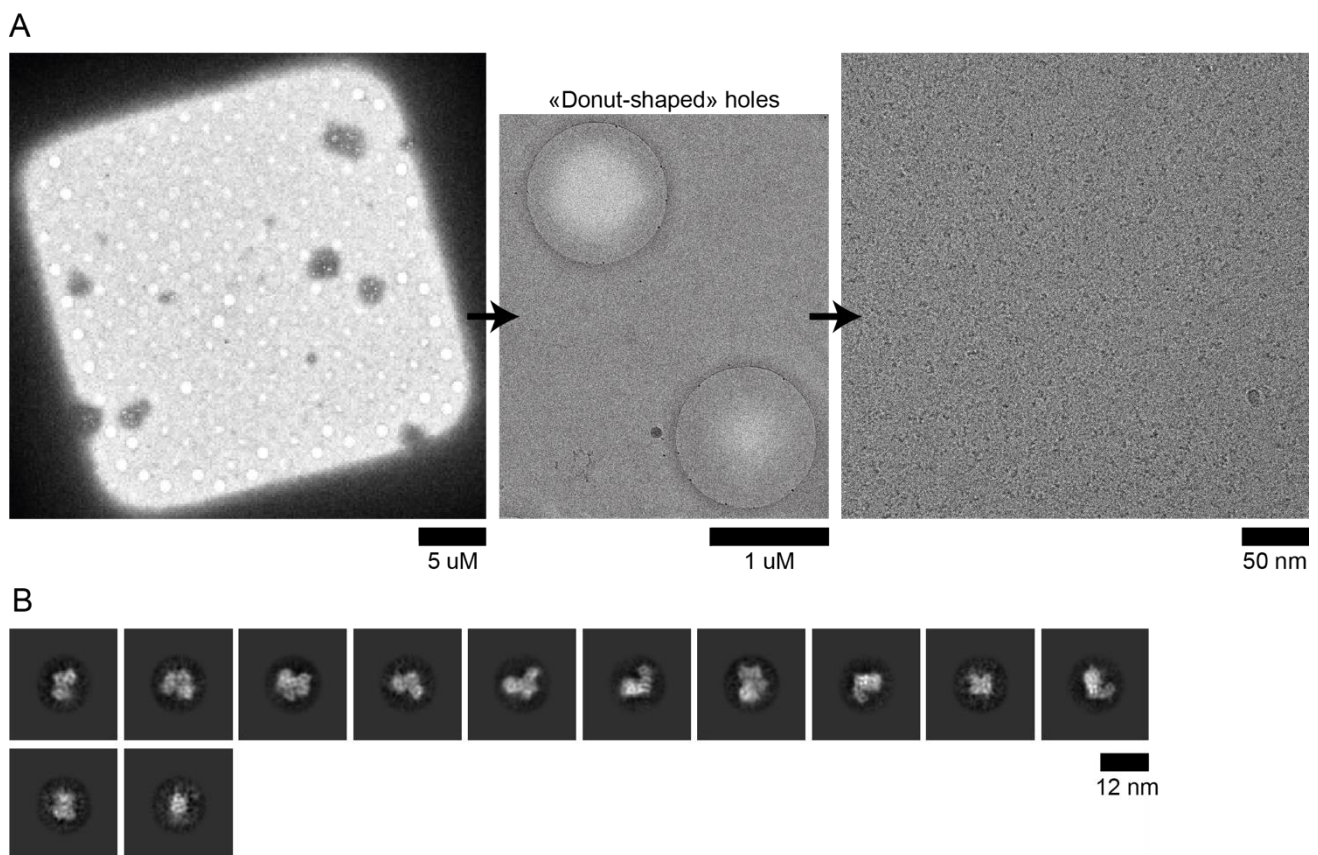


Figure 18. CryoEM grids using a non-tagged CCCP21

A, On the left is an image of a square on an EM grid. These squares are perforated. Protein particles are frozen in a thin ice layer that forms inside these holes during the rapid plunge-freezing step described before. Different

ice thicknesses can be observed, and “donut-shaped” holes with thin ice in the center and thick ice on the side are best for CCCP21 imaging (image in the middle). A representative cryo-electron micrograph with CCCP21 embedded particles in ice showing equal distribution. **B**, 500 movies were collected on this grid and subsequently processed. 2D classes of picked and averaged particles are shown on the right. A 12 nm circular mask had been applied. Various particle orientations can be observed, with the 2D classes showing clear protein features

In a next step, more micrographs were collected with the aim to reconstruct a high-resolution 3D model of the CCCP21 (Figure 19.) complex. We collected five datasets on our 200 keV Talos Arctica which resulted in a total sum of 5142 movies. After motion correction and CTF estimation, manual inspection of all micrographs was performed, bad micrographs were removed, which led to a final data set of 4530 micrographs. Next, I started picking particles using the automated particle picking pipeline Topaz, a machine learning tool implemented into cryoSPARC that picks particles with high accuracy, once correctly trained¹⁶⁵. Topaz also performs well on detecting and picking CCCP21 particles, especially when compared to conventional particle picking tools that rely on classical methods such as using Laplacian of Gaussian (LoG) spatial filter. 1.113.646 particles were extracted and subjected to a first round of 2D classification. The best resulting particles were used to train the automated particles picking tool Topaz. After each iteration, the program will identify and pick particles with higher accuracy on each micrograph, leading to improved 2D classifications. After several iterative rounds of 2D classification, a total of 401.081 selected particles were used to reconstruct a 3D model. Four *ab-initio* models were generated from the cleaned particle set. As commonly observed in cryoEM these initial models differ substantially between each other, a phenomenon that can be explained by particle damage caused by the ionization, during the freezing process or simply due to structural heterogeneity of the complex (multiple structural conformations). Moreover, the data show that our data set suffers from a dominant particle orientation, a problem often encountered in cryoEM. As a consequence, rare particle orientations are limiting the information along certain axes (low signal-to-noise ratio) and prevent a high-resolution reconstruction. We excluded all particles that seemed broken and started rounds of refinement resulting in a map resolution of around 10 Å. After several rounds of refinement, and careful selection of particles, I could improve the map to a final nominal resolution of 7.4 Å. This map is generated with a final data set of 154.893 selected particles.

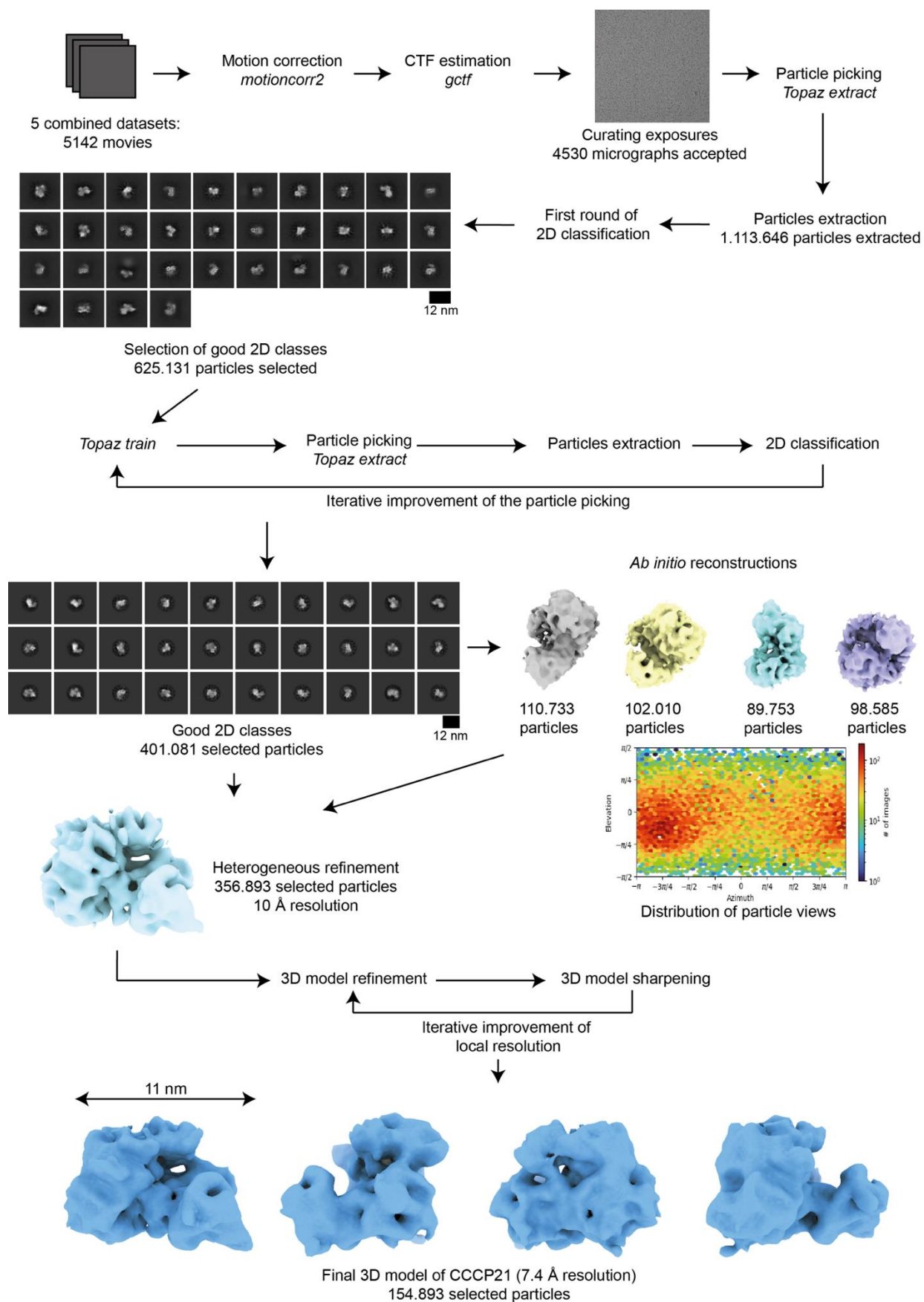


Figure 19. Processing pipeline of CCCP21

See methods. CCCP21 has a diameter of 110 Å (11 nm) and a circular mask with a radius of 120 Å (12 nm) had been applied (in order to reduce noise and promote particle alignment).

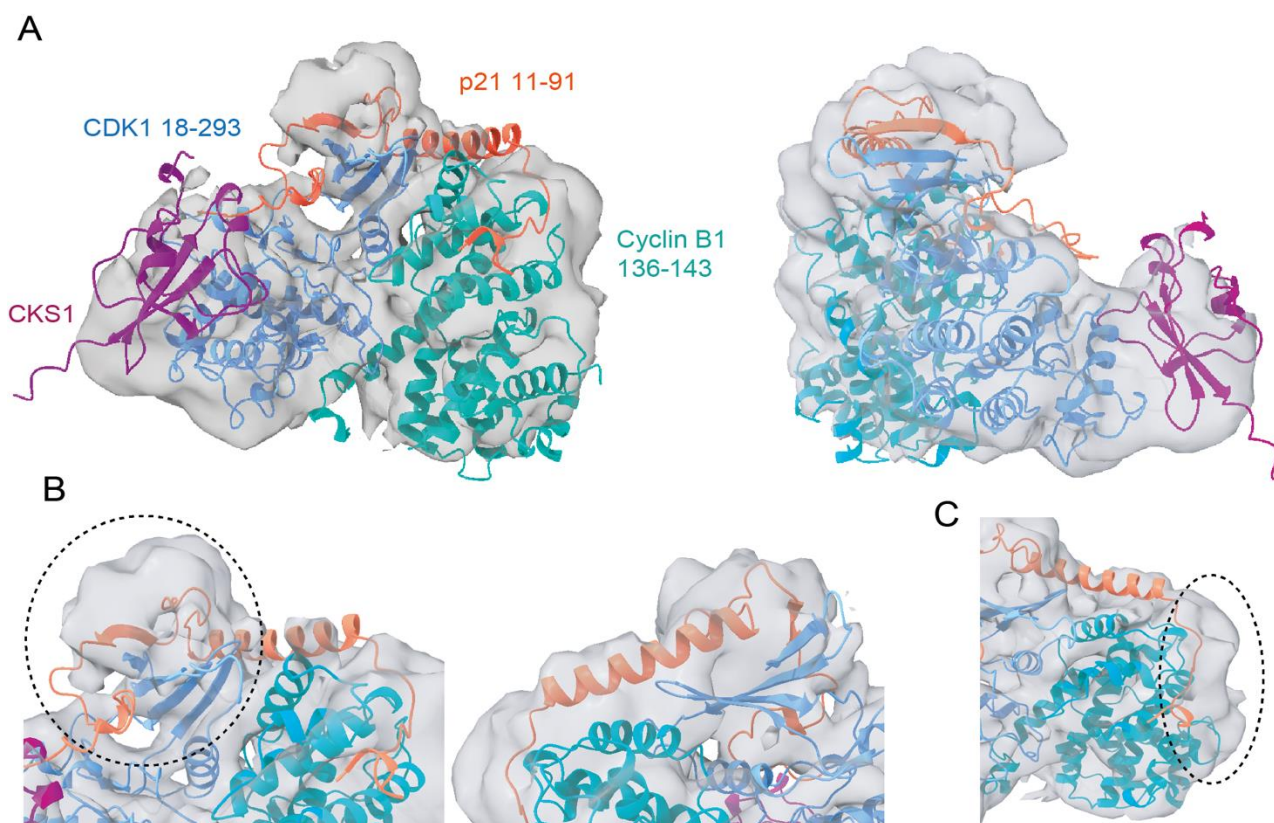


Figure 20. Coordinates of the AF2 prediction of the CCCP21 complex fitted in the EM density.

Unstructured regions (of all proteins) that are predicted not to be involved in complex formation are removed for clarity. **A**, Overall structure of the CCCP21 complex in different orientations. The predicted complex fits well into our EM density map. Proteins are color-coded as in Figure 16. The density map is shown in grey. **B**, Close-up zoom on the extra-density on top of the CDK1 N-lobe, fitting the predicted K domain of p21 and the adjacent β -strand. Additional density is also visible following the CDK1/p21 β -sheet and preceding the 3_{10} helix domain. It is tempting to speculate that this interaction causes the inhibition of CDK1 by p21. The image below shows the fit of the LH domain. **C**, Zoom on cyclin B1, with additional density being visible on the right. This extra density might represent binding of the RxL domain and the D1 segment of p21 to cyclin B1.

The obtained map is of sufficient quality to allow correct fitting of the coordinates of the AF2 prediction of CCCP21 in the generated EM density (Figure 20. A). Additional density is visible at the site where the p21 KID domain is predicted to bind to the CCC complex, validating the obtained 3D volume. The region where the main extra density can be observed near the CDK1 N-lobe, more precisely on top of the protein in the displayed orientation (Figure 20.B). Our AF2 model predicts that residues of p21 form a β -strand after the K domain upon binding to CDK1, thereby extending the β -sheet of CDK1. Moreover, the K domain is predicted to form a hairpin-like loop structure that fits well into the observed extra density. The bridging α -helix that connects CDK1 and cyclin B1 is not very well defined in our 3D volume, nevertheless clear additional density is visible. This density, however, is not well defined due to the limitations in resolution, likely caused by the aforementioned preferential orientation of protein particles. The predicted model likely shows the correct binding of this

helix to CDK1 and cyclin, which is expected to bind similarly to cyclin B1/CDK1 as p27 binds to CDK2 from which the crystal structure has been solved in the past. Structural information on the binding mode of the D2 segment and the 3₁₀-helix domain of p21 is not available and therefore an accurate prediction is not supported by the obtained EM density yet. Additional density from the LH region of p21 can be observed nearby cyclin B1, indicating that the D1 segment bearing the RxL binding site is binding to cyclin B1 opposite to the cyclin B1/CDK1 interface. Our structural and functional data suggest that p21 is bound to the CCC complex and additional density can be observed at regions where the KID domain is expected to bind. Excitingly, we also observe additional densities around cyclin B1 and CDK1 that cannot be explained by the existing p27/CDK2 crystal structure, providing further insights into the p21-mediated cyclin B1/CDK1 regulation.

Our first hypothesis was that the limitation of the resolution was due to a limited amount of finally selected particles and to overcome this problem, we needed to collect more movies. At the time, our Talos Arctica was not equipped with the latest generation detector and consequently we were not able to collect a large number of movies in few days. To collect a larger dataset, we decided to analyze our grids on a 300 keV Titan Krios (DCI Lausanne) equipped with a better and faster detector. We collected 16.598 micrographs and processed them the same way as we did for the previous datasets. Most micrographs were collected close to focus and combined to the accelerated electrons; it is resulting in particles that are less visible on a micrograph for the human eye (Figure 21). Due to poor picking performance by Topaz, I decided to pick particles manually and use those particles as a reference to train Topaz. This procedure enabled me to automatically pick and extract 7.106.575 particles from this data set. Subsequent 2D classification resulted in classes that did not appear as sharp as from previous data sets and particles seemed to be damaged. While several rounds of 2D classification improved the quality of the classes (now using around 612.003 selected particles), the classes remained mostly blurry, and the particles appeared to be damaged, potentially due to a collection in too thin ice. As a consequence, we were not able to generate a map at high resolution.

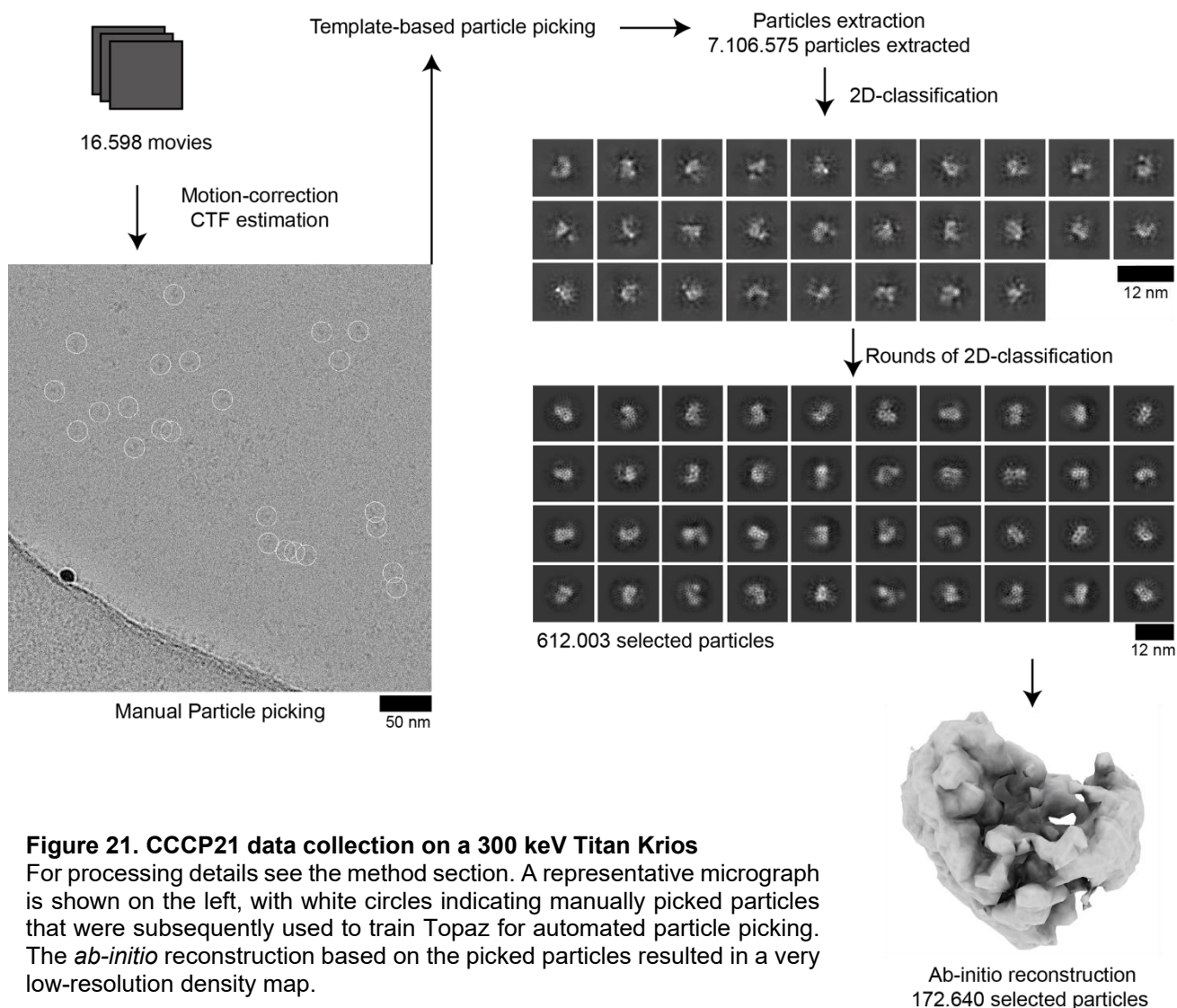


Figure 21. CCCP21 data collection on a 300 keV Titan Krios

For processing details see the method section. A representative micrograph is shown on the left, with white circles indicating manually picked particles that were subsequently used to train Topaz for automated particle picking. The *ab-initio* reconstruction based on the picked particles resulted in a very low-resolution density map.

After this attempt, we wondered if the particles were too damaged because we needed to collect micrographs in thicker ice on the Titan Krios. To assess this hypothesis, we sent a set of plunged EM grids with different ice thickness to be imaged by the ThermoFisher Center for Electron Microscopy in Eindhoven. To our surprise, even with checking the quality of the grid on our microscope before sending them, very few of our grids gave good micrographs and, most of the time, the ice was severely damaged during data collection (Figure 22.). The thin ice on which we were collecting on the Talos Arctica was too thin for this data collection and the ice was too damaged to be analyzed. The very few micrographs that we could analyze gave us results similar to our large Titan Krios dataset and when we analyzed the local motion on each micrograph, we noticed that the sample was getting severely damaged during the data acquisition.

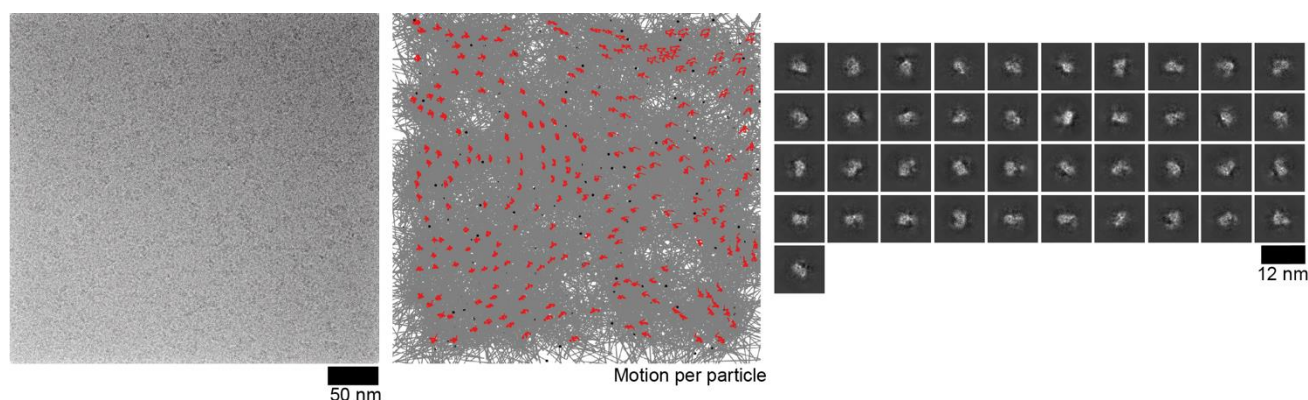


Figure 22. Small scale data collection of CCCP21 grids with different ice thicknesses

On the left is a representative cryo-electron micrograph of one of the six datasets collected on different grids previously screened on our 200 keV Talos Arctica. Particles are visible and could be picked but analysis of the local motion trajectory on these micrographs revealed strong movements of particles during data collection. Black dots represent coordinates from picked particles on the grid and grey lines their trajectories. On the right are the resulting 2D classes of these data sets.

Collecting many large data sets helped us to realize, that quality of the vitrified EM grids was still highly variable when applying the CCCP21 sample, and parameters for reliable grid preparation still needed to be identified. We therefore decided to address the problem of variable grid quality next. To do so, we modified the most common parameters during sample and grid preparation and analyzed their influence on grid quality systematically. I will discuss some of these efforts in the subsequent paragraphs.

3.2.3. Improving grid quality

Grid preparation and optimization is one of the most difficult tasks when performing cryoEM studies, but is essential, because poor grid quality will limit the result of the final outcome, a high-resolution density map. To optimize grid preparation, I systematically tested different grid types, detergents that were added to the protein buffer, and used different concentrations and type of salts in the buffer. The aim was to obtain grids with large areas of reproducible ice (thickness), suitable for data collection. Besides, I also used different protein concentrations during the grid preparation, a parameter that has been observed in many independent studies to influence ice thickness. Generally, more consistent ice formation can be observed on a grid, when the applied protein concentration is higher. However, a protein concentration that is too high, can result in several layers of protein particles resulting in impossible data processing. Further, protein particles that are too close to each other introduce “noise”, thereby reducing accurate particles alignments. Modifying the protein concentration during the vitrification process can often not be linearly translated into particle

concentration on the final EM grid. Even slight changes in the amount of protein used for EM grid preparation, might have drastic effect of the number of particles observed in holes, when screening the grids in the electron microscope. In a tedious and time-consuming effort, I noticed that using a CCCP21 protein solution concentration in the range of 0,1 to 0,2 mg/ml during grid preparation, often resulted in EM grids that showed a useful concentration and even distribution of particles that allowed efficient picking and consistent ice (with the right thickness). Having been able to optimize the protein concentration, I next tried to improve the grid quality through different chemicals and grid materials screening. In the next sections, I will describe my efforts adding different types of detergents to my protein solution before focusing on buffer conditions and EM grid types (size of the holes, addition of support layers but also the metal composition of the grids).

3.2.3.1. Detergents

Originally, detergents were used to extract, purify and stabilize membrane proteins because of their amphiphilic character and capacity to form a bilayer environment (micelles). The concentration at which a detergent starts to form micelles is called the Critical Micelle Concentration (CMC). Detergents have been shown to influence particle distribution in ice. By using specific detergents such as CHAPSO, particles adsorption to the air-water interface can often be reduced (references), probably by modifying the surface tension of the drop. As a result, particles distribution and orientations are randomized (no preferred interaction interface within the ice layer). To examine if detergents can also help optimizing CCCP21 grids, I tried a different range of detergents, including CHAPS, CHAPSO, LMNG, Tween20 and Triton, at different CMC. For all of them, the plunged grids looked nice with evenly distributed ice in most holes. However, except for CHAPSO and Tween 20 (Figure 23) most of the grids were empty. The detergent was added to the solution just before plunging the grids, to avoid any negative effects of the detergent such as the formation of aggregates. With 0,01% Tween 20, particles are visible on the micrographs (Figure 23. A) showing a relatively strong contrast. But after particle picking and 2D class averaging, the 2D classes appear blurry and do not align well. Moreover, the classes imply that the particles might be damaged. The usage of CHAPSO was shown in literature to improve the final resolution of small particles, with a similar shape and architecture to CCCP21¹⁶⁶, and therefore we tried the use of CHAPSO at different concentrations. We could see in all the trials the formation of evenly distributed ice in the holes (Figure 23. B) but even at the lowest of the tested CMC,

0,1%, we observed that the particles are mostly disassembled, and we can only see “dots” everywhere on the micrographs and no clear intact CCCP21 particles. We still collected a small dataset to make sure that no good conditions are missed, but the resulting 2D classes confirmed that the particles were damaged by the addition of CHAPSO. Combined, these results show that CCCP21 is very sensitive to changes in solution composition and addition of detergent often resulted in completely destabilized particles. I therefore decided to next screen for the influence of salt concentration on the stability of the CCP21 complex and the vitrification process to further refine the right conditions for cryoEM imaging.

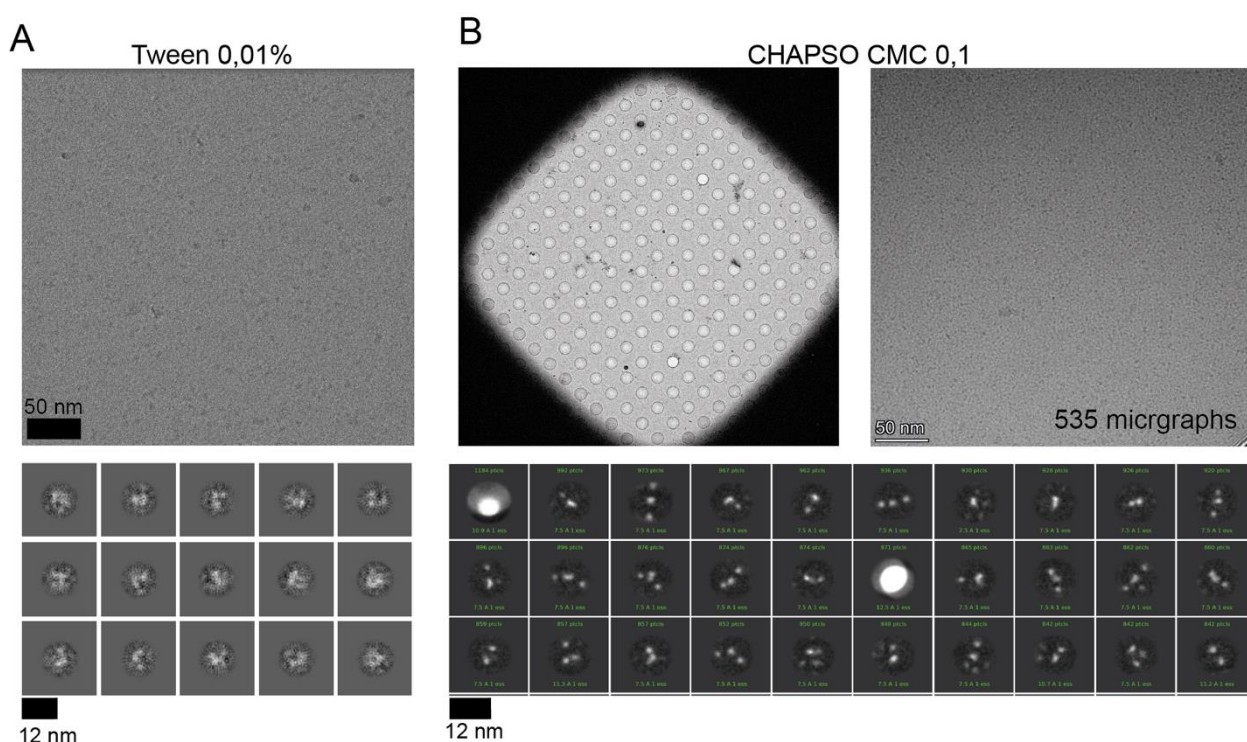


Figure 23. Trials to improve the quality of cryoEM grid preparations with detergents

For details of grid preparation, see methods. **A**, A representative cryo-electron micrograph of CCCP21 embedded into ice supplemented with 0,01% of Tween 20 is shown on the top. A small data set was collected, particles picked, extracted and averaged into 2D classes, shown on the bottom. **B**, On the upper left is an image representing a grid square where CCCP21 is embedded into ice supplemented with CHAPSO (0,1x CMC of 8 mM). A typical cryo-electron micrograph is shown on the upper right. 2D-classes are shown in the lower panel.

3.2.3.2. Salts

Changing the buffer conditions of the protein is an obvious experiment to do and another component that influences the behavior of a protein in solution is the type of salt used. Since the start of our experiments, we used NaCl, as it is the most commonly used salt but in many other cases, other salts are necessary for the stability of proteins and protein complexes. Publications that work on similar complexes in term of size and shape compared

to CCCP21¹⁶⁷, used KCl to purify the proteins and to prepare cryoEM grids. It is also known that the salt concentration and type can influence the formation of ice and particle distribution.

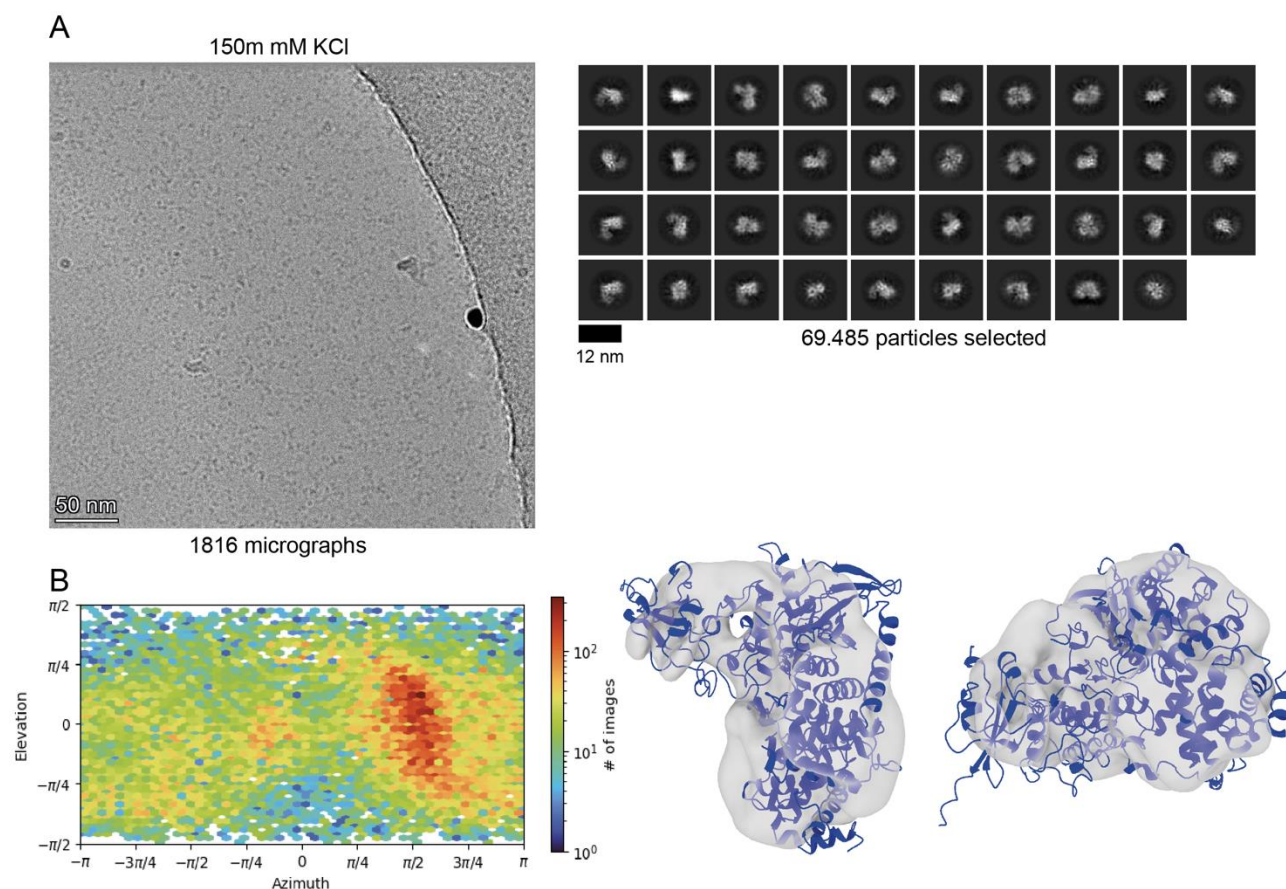


Figure 24. Trials to improve the quality of the cryo-EM grids with different salts

For details of grid preparation, see methods. **A**, Representative cryo-electron micrograph of CCCP21 embedded into ice. Before cryo-plunging, CCCP21 has been purified with a buffer containing 150 mM KCl instead of 150 mM NaCl. 1816 micrographs have been collected, particles were picked, extracted and classified into 2D classes shown on the right. **B**, The left panel shows a graph representing the particles orientation distribution used to reconstruct the *ab-initio* EM density map shown on the right. The color illustrates the number of particles found for each orientation from just a few (blue) to many (red). AlphaFold2 prediction of CCCP21 has been fitted into the *ab-initio* reconstruction on the right panel.

KCl has a higher ionic strength than NaCl, potentially directly influencing the protein surface interactions. We therefore decided to re-do the protein complex purification but exchanging NaCl with KCl. During the complex purification we did not notice any noticeable differences in terms of protein complex stability. We plunged the protein on an EM-grid and analyzed the micrographs. Many holes that show a nice gradient of ice thickness with thin, melted ice in the center of the hole, and a ring of ice with the right ice thickness including seemingly intact particles can be observed (Figure 24. A). We collected a small dataset on our 200 keV Talos Arctica and picked 69,485 particles which were used to perform an *ab-initio* 3D reconstruction (Figure 24. B). Strikingly, we observed a strong preferential

orientation of particles, resulting in a non-homogenous *ab-initio* map. In addition, parts of the CCCP21 seemed missing when fitting the AF2 prediction into the map. Specifically, the CKS1 protein seems to be missing or not well reconstructed due to the lack of particles. KCl grids resulted in micrographs with an even distribution of ice thickness and 2D classes that showed clear protein features but, again, we observe similar problems as with the NaCl grids, and therefore did not see a drastic improvement in the quality of our data. The collection of a larger data set on the upgraded microscope at the DCI-Geneva is planned next

3.2.3.3. EM grid-types

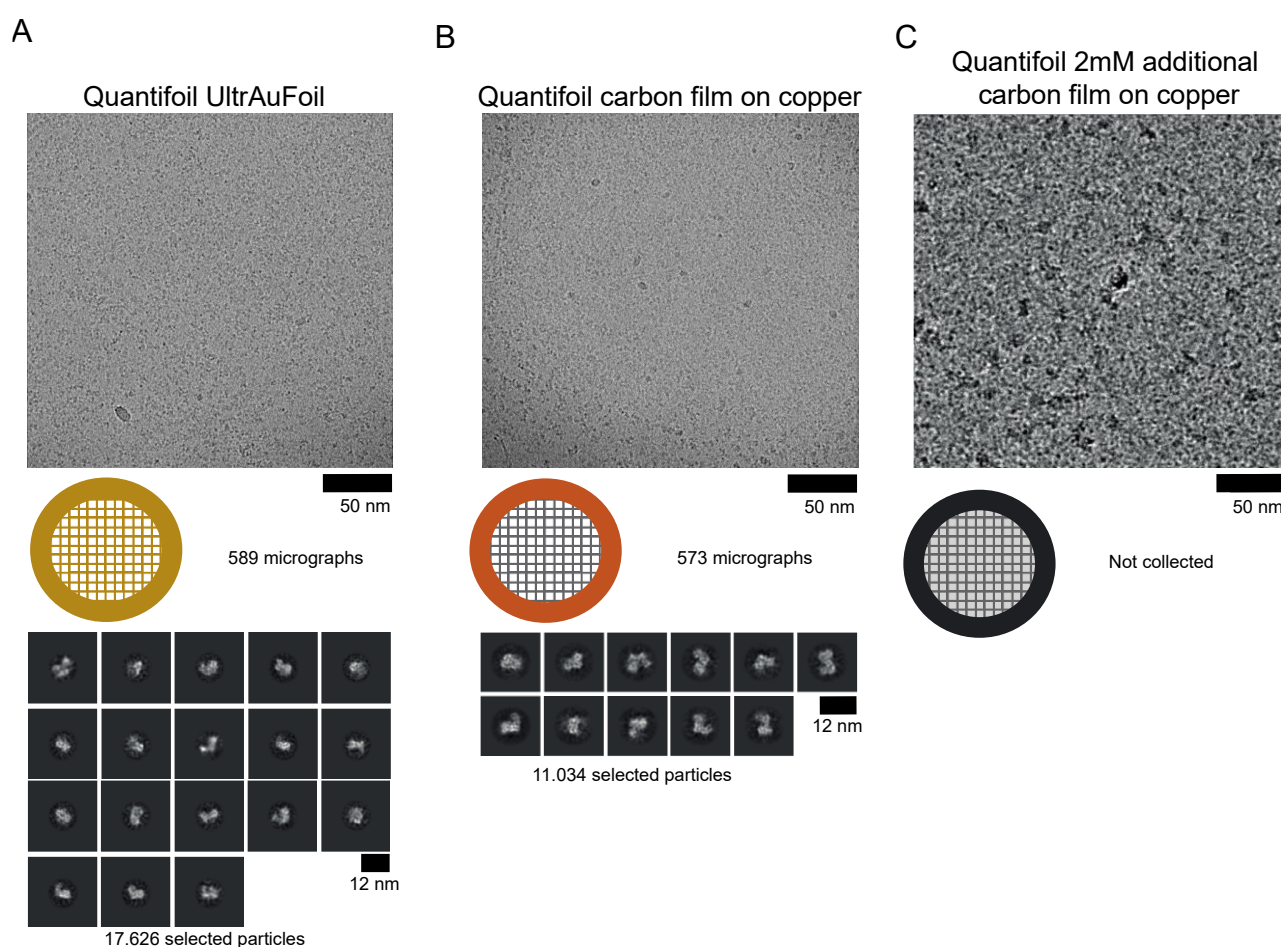


Figure 25. Improvement of cryoEM grids quality by screened different EM grid types

For grid preparation details, see the methods section. **A**, CCCP21 was cryo-plunged on Quantifoil UltrAuFoil grids and analyzed on our 200 kev Talos Arctica. On the top panel is a representative cryo electron micrograph of the collected 589 micrographs dataset. This dataset was furthermore analyzed, and particles were picked, extracted and classified into 2D classes visible on the lower panel. **B**, Same as for A on a Quantifoil copper grid supplemented with a carbon film. 573 micrographs were collected with corresponding 2D classes on the lower panel. **C**, CCCP21 was cryo plunged on a Quantifoil copper grid covered with an additional 2mm carbon support layer and a representative cryo electron micrograph analysis is displayed here.

Developments in cryoEM are not made only on the improvement of the microscopes but also on the EM grids themselves. For example, the grid material directly influences the behavior of the sample during vitrification. Over the last years, a lot of different grid types have been developed using different metal composition for their mesh but also on covering support layers. Most grids are covered by a flat film of carbon, roughly 10-12nm thick, with perforated holes in which the sample is measured. The spacing and the size of the holes differs between the grid types and those parameters also influence the quality of the embedded particles. For CCCP21, the grid type which yielded in our best data sets was the Quantifoil 300 gold mesh covered with a continuous layer of carbon and 1,2 μM diameter of holes, spaced by 1,3 μM (1,2/1,3). We therefore used mainly those grids to image CCCP21, but we also considered other grid types to see if we could change ice formation and particles distribution. First, we tried Quantifoil 300 gold mesh grids with a 50nm thick gold layer, also called UltrAuFoil grids¹⁶⁸. These grids were developed to reduce beam-induced movement by replacing the carbon foil by a flatter, better behaving gold foil, ideal for challenging small particles. Moreover, gold is a less conductive material that, unlike carbon, will not contract upon cryo-plunging. We tried 1,2 μM and 2 μM diameter holes and 1,3 μM and 2 μM spacing between the holes (respectively 1,2/1,3 and 2/2) combined with different protein concentrations and blotting conditions (Table 2). Imaged micrographs showed different protein distribution than on the previously screened EM-grids. As a test, 589 micrographs were collected (Figure 25. A). After few processing steps, it became clear that most particles are damaged regardless of the thickness of the ice. It must be noted that identifying ice thickness is more challenging on the UltrAuFoil grids as gold has more contrast than carbon, therefore the holes appear all empty to a human eye. As a second experiment, we tried Quantifoil 300 copper mesh grids covered with a carbon film with 0,6 μM , 1,2 μM and 2 μM diameter holes and 1 μM and 1,3 μM spacing between the holes (0,6/1, 1,2/1,3 and 2/1 respectively) combined with different protein concentrations and blotting conditions (Figure 25. B and Table 2). Micrographs with good particles distribution in ice were obtained and a little dataset was collected from which we could extract descent 2D classes. Nevertheless, 2D classes and particles orientation remain similar to the ones obtained with our big dataset on the Talos Arctica. Additionally, the 3D reconstruction of CCCP21 was not improved and more preferential particle orientation in ice was observed. We observed in the past that the usage of graphene oxide-covered grids resulted in high background noise and difficulty to obtain high-contrast particles, but as previously mentioned, different grids support layer were developed to attract particles, like ultrathin 2nm layers of carbon (UTC)¹⁶⁴. We tried this

support layer on Quantifoil 300 copper mesh 2/1 using a low protein concentration (Table 2) but obtained a higher background noise than for graphene oxide and we could not even distinguish the particles in ice. We decided to not collect a dataset on those grids (Figure 25. C). Additionally to the grids we just described, we also tried a diverse range of EM grids that did not give conclusive results (Table 2.) like, per example, C-flat grids designed to have a flatter surface than Quantifoil grids and promote ice formation in the holes but also Quantifoil HexAuFoil grids with hexameric-shaped squares and smaller holes designed to both facilitate ice formation and data collection procedures¹⁶⁹. For these last ones, we did observe very good ice formation but because of the small size of the holes and our large EM beam at the time, we could not image them on our 200kv Talos Arctica. Imaging these grids at high magnification on a 300 kev Titan Krios in DCI Lausanne to avoid the high-contrast gold overcame this issue, but it reduced dramatically the particles contrast. These experiments highlight the complexity and importance of EM grid choice when one decides to image protein particles with cryoEM and, in our case, showed that the correct type for CCCP21 remain the Quantifoil 300 mesh gold with a carbon layer 1,2/1,3.

EM-grid type	Protein concentration	Blotting time (seconds)	Additional move	Glow-discharge (seconds)	Temperature/% humidity
Quantifoil carbon film on gold 1.2/1.3	0,1 mg/ml 0,15 mg/ml 0,2 mg/ml	1 – 3,5s 1 - 2,5s 1,5s 2,5s	0 - 1mM 0 - 1mM 1mM 1mM	60s	14°C/95%
Quantifoil carbon film on gold 2/2	0,1 mg/ml 0,15 mg/ml 0,2 mg/ml	1,5s	1mM	60s	14°C/95%
Quantifoil carbon film on copper 1.2/1.3	0,1 mg/ml	1 – 2s	1 – 1,5mM	60s	14°C/95%
Quantifoil carbon film on copper 0,6/1	0,1 mg/ml	1 – 1,5s	1,5mM	60s	14°C/95%
Quantifoil 2mM additional carbon film on copper 2/1	0,05 mg/ml	1,5 – 2s	1 - 1,5mM	20s	14°C/95%
C-flat carbon film on copper 1,2/1,3	0,1 mg/ml	1– 2s	1 – 1,5mM	20s	14°C/95%
C-flat carbon film on copper 2/2	0,1 mg/ml	1– 2s	1 – 1,5mM	20s	14°C/95%
Quantifoil UltrAuFoil 1,2/1,3	0,1 mg/ml 0,2 mg/ml	1– 2s	1 – 1,5mM	60s	14°C/95%
Quantifoil HexAuFoil 0,3/0,6	0,1 mg/ml	1,5 – 2s	1mM	60s	14°C/95%

Table 2. Blotting and plunging conditions of different screened EM grid types

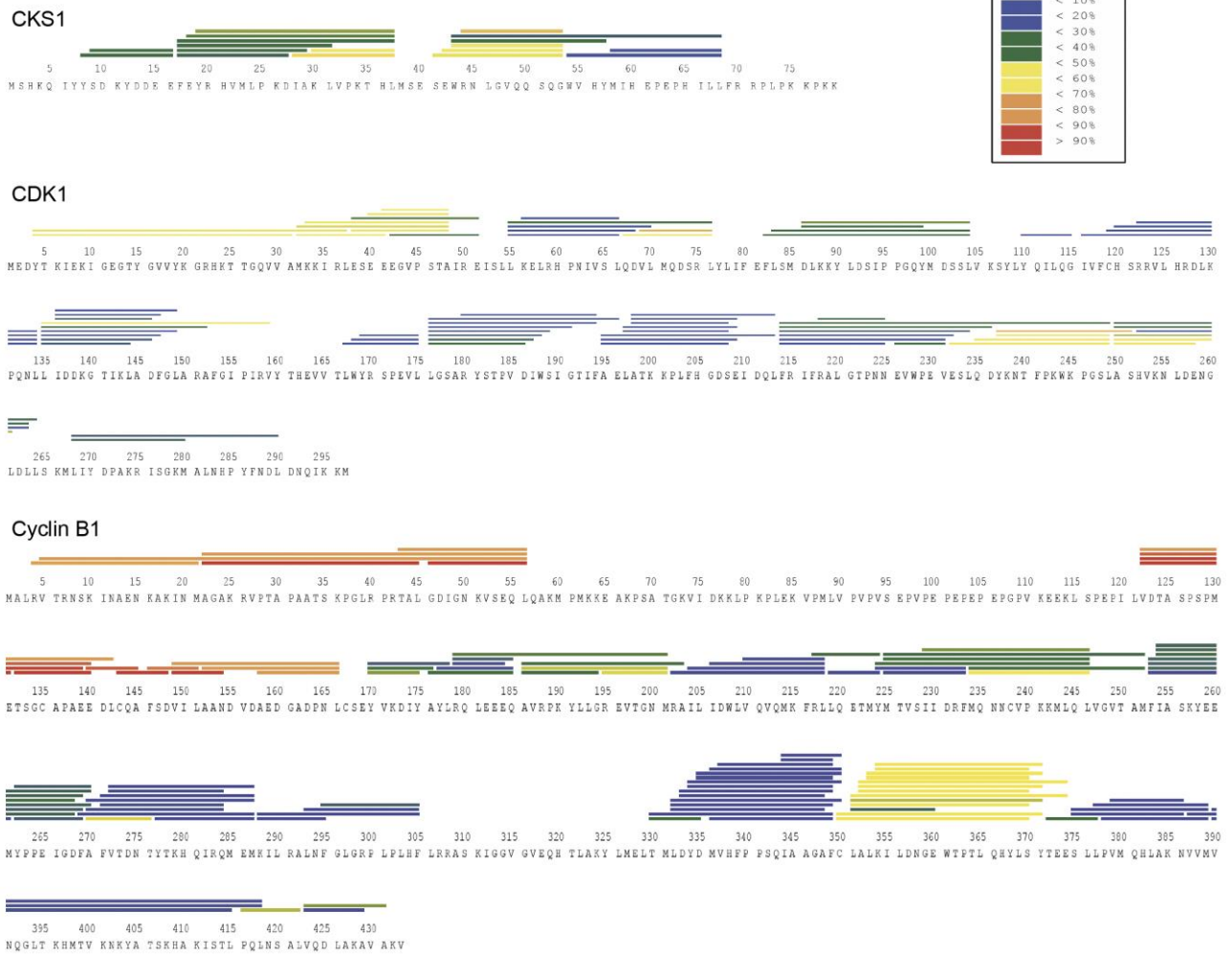
Several blotting and plunging conditions on different grid types were screened in order to improve the quality of the collected datasets. Details on the grid types are described in the methods section.

4. HDX-MS studies

CCCP21 is a challenging sample, and to determine its three-dimensional structure by cryoEM to a resolution that allows atomic modeling will require more time and additional collection of datasets. To overcome this lack of high-resolution information on the interaction of p21 with the CCC complex and to understand its structural organization, we opted for a different approach. Mass-spectrometry (MS) is a technique that can be used to analyze biomolecules, and which is divided in many sub-categories of specific techniques detecting the structural organization of protein and protein complexes in their native state. While traditional structural biology techniques such as X-ray crystallography, cryoEM, Nuclear Magnetic Resonance spectroscopy (NMR) or Small Angle X-ray Scattering (SAXS) allow the determination of high-resolution structures of target proteins, these techniques remain still challenging and are not well suited to analyze dynamic proteins like IDPs. MS techniques, however, can be applied to *in vitro* and *in cellulo* protein samples and provide information on protein dynamics, such as conformational changes in solution and are therefore valuable complementary techniques to structural biology¹⁷⁰. Structural MS techniques can be divided into two categories, the peptide-based ones, in which the protein of interest will be digested into peptides before being analyzed and the protein-based ones, in which the protein will remain intact. Hydrogen/deuterium exchange MS (HDX-MS) is a peptide-based technique and is used to study the folding of protein as well as their conformational changes in different states, for example upon binding to a ligand. HDX-MS is a method that is widely used in combination to cryo-EM studies¹⁷¹. On a protein backbone, amide hydrogens exchange constantly with free hydrogen that is present in the surrounding solution in a measurable rate. HDX-MS measures the replacement of hydrogens with a heavier isotope of hydrogen, namely deuterium (D₂). This exchange rate depends on the overall structure of the protein, because some regions of the protein are solvent exposed, whereas others form the hydrophobic core of the protein (complex). Therefore, the exchange rate is providing information on the structural organization of the protein. Upon binding to a ligand or a protein partner, the structure of the protein target changes and as a consequence also the deuterium exchange rate. By comparing D₂ incorporation before and after binding, we can get information on the binding sites and protein rearrangements. During my thesis, I used HDX-MS to study the conformational changes on CCC upon binding to p21.

4.1. HDX-MS results upon p21 binding

A



B

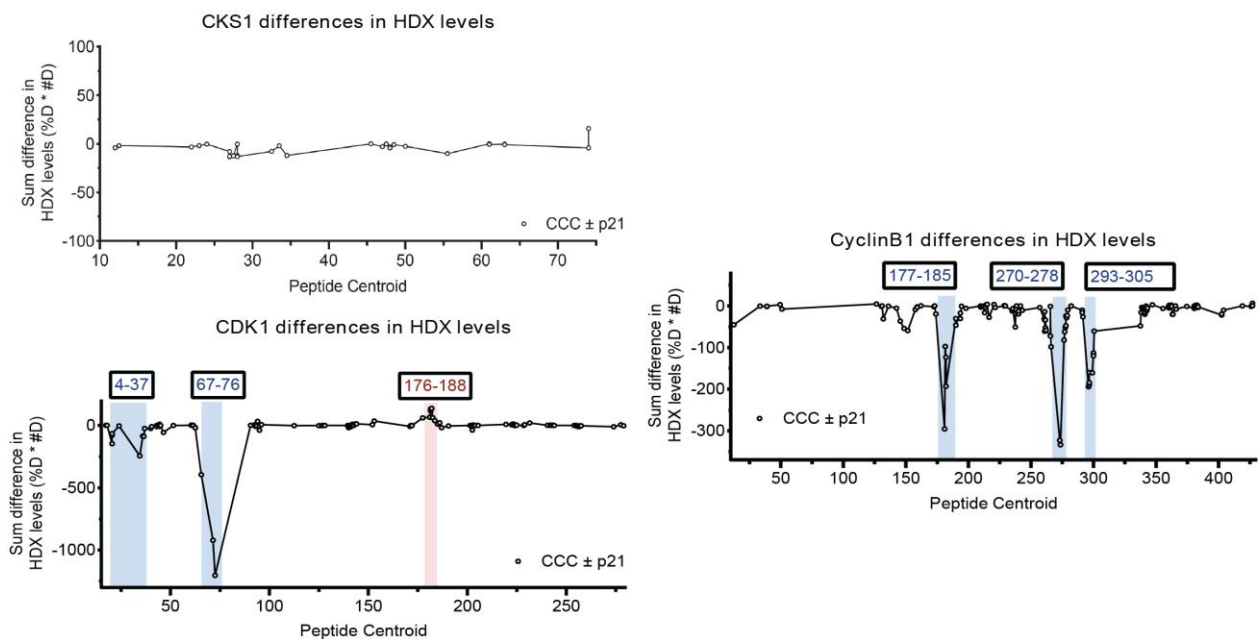


Figure 26. HDX-MS experiments on cyclin B1, CDK1 and CKS1 using CCC and CCCP21 complexes

For details, see the methods section. **A**, Schematic representation of the peptide map of each of the proteins. Each line represents a high-quality peptide that was analyzed to measure its H/D exchange rate. Peptides are colored according to their dynamics following the legend, with highly dynamic peptides shown in red and protected peptides colored in blue. **B**, Graphs representing the differences in deuteration levels for each of the proteins. Comparison of deuterium incorporation for each studied peptide is represented. The sum of differences for all timepoints is calculated, both as percentage deuteration (%D) and number of deuterons (#D). The product of the sum in differences in %D and #D is then multiplied to increase signal-to-noise ratio. Regions highlighted in blue indicate protection from exchange by the ligand, whereas regions highlighted in red show increases in H/D exchange rate. Protection indicates either direct binding site or stabilization of overall structure, whereas increase indicates either allosteric conformational change or disruption of intra-molecular contact.

The aim of these experiments was to identify where p21 is binding to the CCC complex and what changes in conformation of the three proteins binding to p21 induces. For this, we measured the D₂ exchange rate of CCC and CCCP21 and compared them. First, as stipulated before, HDX-MS is a peptide-based method that requires digestion of proteins into peptides before being analyzed. Peptides are analyzed based on their dynamics where highly dynamic peptides display limited amount of information. This digestion is protein-dependent and depends on the stability of the protein in solution. For the CCC complex, the mass-spectrometry analysis showed a good peptide-coverage of all three proteins (Figure 26. A). For CKS1, the obtained peptides covered almost the entire protein sequence and were not so dynamic. The same observation can be done for CDK1 with the exception of the residues 155 to 166, which correspond to the T-loop and were unfortunately not well covered. However, the regions surrounding the T-loop showed good coverage and provide information referring to the T-loop structural dynamics. For cyclin B1, the peptide coverage is lower than for the other two CCC proteins. According to the AF2 prediction of CCCP21 (and other secondary structure prediction programs), the cyclin B1 N-terminus is unstructured, including residues 1 to 136. We can see that the peptide coverage in that region is very low with highly flexible peptides. The determined structures of cyclin B1 using cryoEM or X-ray crystallography have either removed the first 165 residues or can't resolve them, indicating that these residues are indeed unstructured and flexible. We observe a good peptide coverage for residues 165 to 433, the structured part of cyclin B1. Interestingly, there is a portion in the C-CBD of cyclin B1 lacking peptide coverage (305-330) where the PBP is located. Therefore, we could not gather information regarding this region with HDX-MS. The D₂ exchange rate was measured for CKS1, CDK1 and cyclin B1 in CCC and CCCP21 and compared to each other (Figure 26. B). Interestingly, we cannot observe any differences in CKS1 upon binding to p21 which implicates that CKS1 is not directly involved into p21-mediated regulation of CDK1 and cyclin B1. For both, CDK1 and cyclin B1, we can detect reproducible changes in the exchange rate. In the case of CDK1, the regions 4-37 and 67-76

showed a lower exchange rate (in blue) upon binding to p21, indicating a direct interaction with p21 or a stabilization of the overall structure. The region 176-188 showed a higher exchange rate (in red) upon binding to p21, which indicates either conformational changes or a disruption of intra-molecular contacts. In the case of cyclin B1, three regions showed a lower exchange rate upon binding to p21, namely aa 177-185, aa 270-278 and aa 293-305. While there is no existing structure of CCCP21, AF2 can predict the CCCP21 complex with high confidence. Therefore we decided to use this predicted model, in combination with the existing structure of p27 in complex with diverse cyclin/CDK^{130,59} complexes, and p21 bound to cyclin D1/CDK4¹³¹, to better analyze our HDX-MS results. Comparing the AF2 prediction of CCCP21 with the observed differences in D₂ exchange rates, indicates that changes are mainly located in the regions where the KID domain of p21 is predicted to bind on cyclin B1 and CDK1 and in the contact zones between cyclin B1 and CDK1 (Figure 27. A).

A cyclin B1 region close to the MRAIL (Ala270-Thr278) shows a reduced exchange rate (Figure 27. B), but not the MRAIL patch itself. In particular, neither Gln211 nor Lys215 residues, predicted to bind p21, showed a difference in D₂ incorporation upon p21's binding questioning their implication in complex formation. A multiple sequence alignment of these residues from cyclin A2, cyclin D1 and cyclin B1, shows that the residues Thr273, Asp274, Asn275 and Thr276 are relatively well conserved. The Thr276 points towards the p21 RxL motif and its homologue in cyclin D1, Ser131, has been shown to form a hydrogen bond with p27. Therefore, we postulate that this residue is also mediating binding of the p21 D1 fragment to cyclin B1. This conserved region in cyclins is involved in substrate-binding and residues forming this region can be phosphorylated. P21-binding to this region might prevent phosphorylation of these residues and occlude other substrates to bind. These results indicate that these residues located at the interface between the N- and C-CBD might be important for docking of the D1 segment of p21's KID, and therefore the formation of a stable contact between the first p21 RxL motif and the cyclin B1 MRAIL patch as well as other associated residues. The two other regions of cyclin B1 that show a reduced exchange rate (177-185 and 293-305) upon p21-binding are not located in close proximity to the AF2-predicted binding site to p21 (Figure 27. C). The first region includes the residues Tyr177 to Gln185 that are located in cyclin B1's N-terminal helix. It has been shown that the N-terminal helices in B and A-type cyclins are necessary for binding to their CDK partner and to provide substrate specificity. Moreover, mutations of the conserved residues Tyr177 and Glu182 (present in our identified region) abrogate binding to CDK1¹⁷².

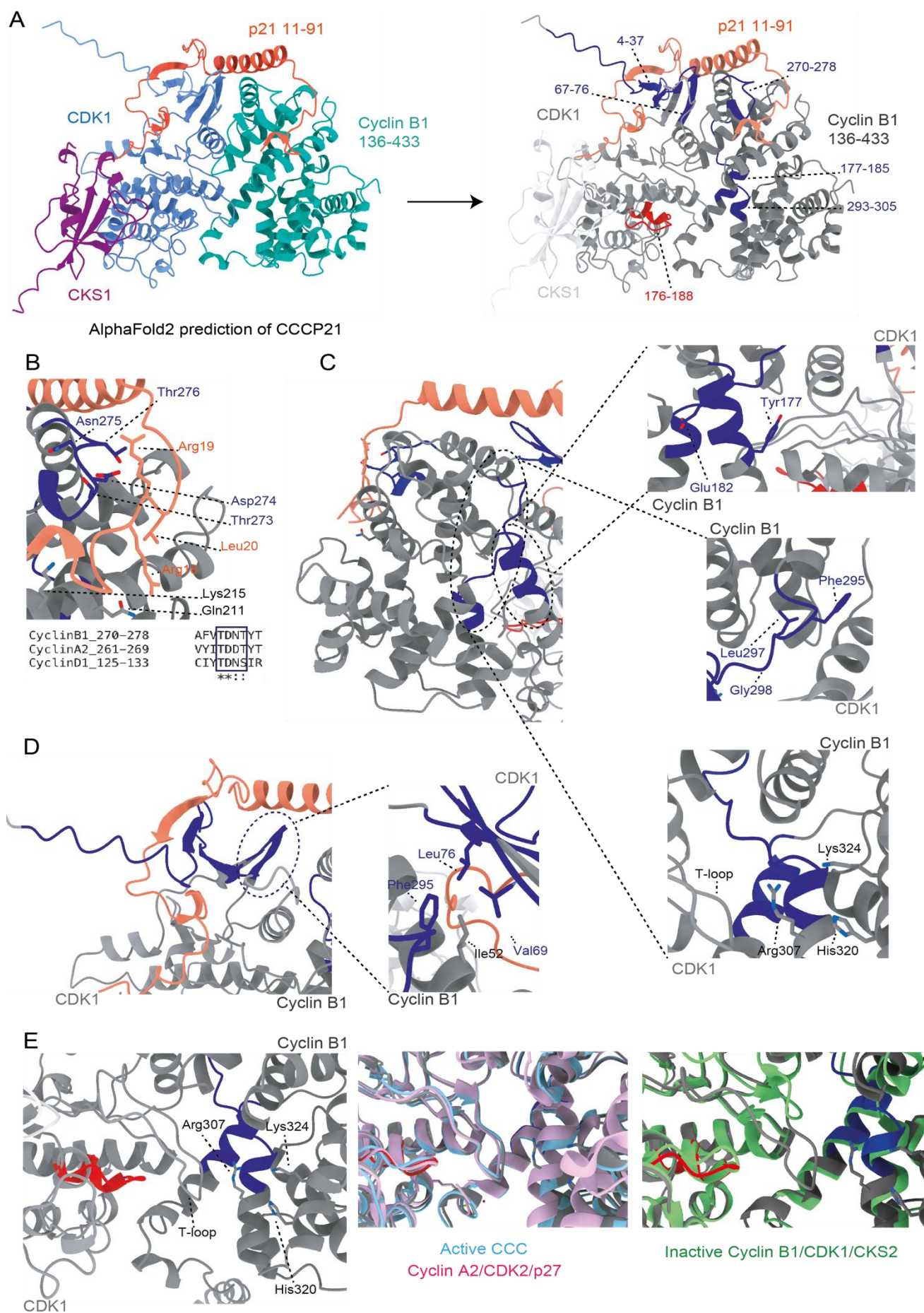


Figure 27. HDX-MS results mapped onto the AF2 predicted CCCP21 structure

A, The left panel shows the AF2 prediction of the CCCP21 complex structure, with the first 135 residues of cyclin B1 hidden, as well as the residues 1-10 and 91-164 of p21 not shown (all unstructured). Cyclin B1, CDK1 and CKS1 are shown in blue green, blue and purple, respectively. To visualize the changes observed by HDX-MS upon p21-binding (right panel), the CCC complex subunits are illustrated in dark grey (cyclin B1), grey (CDK1) and light grey (CKS1). Following the same color-code as previously described for the HDX-MS data analysis, regions highlighted in blue indicate protection from exchange by the ligand p21, whereas regions highlighted in red show an increase in the H/D exchange rate. Protection indicates either a direct binding event or stabilization of the overall structure, whereas an increase indicates either allosteric conformational changes or disruption of intra-molecular contacts. Because p21 was not analyzed with HDX-MS, the protein remains colored in orange. **B**, Zoom on the highlighted protected cyclin B1 residues 270-278 and their predicted interactions and proximity to p21. Sequence alignments of p21-interacting residues are shown below between cyclin B1, cyclin A2 and cyclin D1. MSA were generated using Clustal Omega and conserved residues are framed (T (Thr273), D (Asp274), N (Asn275) and T (Thr276) in cyclin B1). Stars (*) mark invariant residues in all sequences. Colons (:) label highly similar residues. Dots (.) mark partially conserved residues. Conserved residues are shown as sticks in the AF2 prediction, as well as the p21 RxL motif (Arg18, Arg19 and Lys20). The cyclin B1 residues Gln211 and Lys215, not resolved in the measured protected region but important of RxL-containing protein substrate docking, are also highlighted as sticks. **C**, Zoom on the two other highlighted protected cyclin B1 residues 177-185 and 293-305. The AF2 prediction shows no proximity/interactions with p21. Residues 177-185 in the cyclin B1 N-terminal helix are important for binding to CDK1, whereas residues 293-305 are in an important hydrophobic region involved in substrate recognition and phosphorylation by CDK1. Residues 293-305 are also in close proximity to the PBP key residues Arg307, His320 and Lys324. The CDK1 T-loop is labelled. **D**, Close-up view of the highlighted protected CDK1 residues 4-37 and 67-76 both located in the N-terminal lobe of the protein. The β -sheets forming the lobe rearrange upon binding to p21 and allow the 3_{10} helix to enter the catalytic cleft and inhibit the complex. **E**, Zoom on the highlighted CDK1 residues 176-188 with a higher exchange rate. This region is close to the CDK1 T-loop, cyclin B1 PBP as well as protected regions in cyclin B1. These residues in the AF2 predicted structures do not adopt a different conformation than in active CCC⁶⁴ (PDB:7NJ0) and the active cyclin A2/CDK2/p27 structure¹⁴⁴ (PDB:6ATH). When compared with inactive cyclin B1/CDK1/CKS2¹⁷³ (PDB:4YC3), these residues are slightly rearranged but the main difference is the retraction of the CDK1 T-loop from cyclin B1.

By combining this information with our HDX-MS, we propose that the N-terminal helix of cyclin B1, in addition to being a key region for binding to CDK1, also undergoes specific rearrangement upon binding to p21, thereby mediating p21 binding to the CCC complex. The second region spans the residues Leu293 to Phe305 and has been described in literature to make extensive hydrophobic contacts with CDK2⁵⁹. Moreover, it shows very close proximity to the RRASK (307-311) motif conserved in all B-type cyclins, but different in A and E-types cyclins. This motif was shown to be important for Cdc25 substrate recognition by cyclin B2 in *X. Laevis*¹⁷⁴ as its mutation abolished CDK1 activation and kinase activity on its substrates. The residues Phe295, Leu297 and Gly298 were shown to directly interact with CDK2 with a B-type cyclin specificity for Gly298. Upon binding to p21, this region is stronger protected from D₂ exchange suggesting that binding of p21 reinforces those hydrophobic contacts between cyclin B1 and CDK1. Another important feature of this region of cyclin B1 is its close proximity to the cyclin B1 PBP and the CDK1 T-loop, although we did not observe any changes in the exchange rates at T-loop when p21 is bound and therefore it does not seem to impact the conformation of the loop significantly.

Two regions in CDK1's N-lobe showed a reduced D_2 exchange rate (4-37 and 67-76) and both are in CDK1 β -sheet showed to interact with p21 (Figure 27. D). It is known that upon binding to the D2 segment of p21's KID, the first two beta strands of the CDK1 N-lobe undergo structural rearrangements and the same effect have been observed for p27 binding to other CDKs. AF2 predicts the first β -strand in CDK1 (4-13) to be unstructured and the second β -strand (16-23) to directly bind to p21 but we propose, that similarly to CDK4 upon binding to p27, both β -strands become structured upon binding to p21. As a consequence, these sequences will be pulled up and therefore participate in the rearrangement of the active site and the T-loop, as well as the inhibition by the 3_{10} helix domain as described earlier. Importantly, residues Lys20 and Val18, predicted to interact with p21, are stronger protected upon p21 binding, validating that our predictions on p21's KID binding mode to CDK1 are likely to be correct. We therefore conclude that p21 interacts with CDK1 by making direct contacts with the N-terminal β -strands through its KID domain. The second region, from residues Gln67 to Leu76, provides key residues responsible for binding of cyclin B1. Two regions that show lower D_2 incorporation and are therefore more protected, face each other in cyclin B1 and CDK1. It is known that cyclin B1 Phe295 makes hydrophobic contacts with CDK2 Ile52, Val69 and Leu76, residues well conserved in CDK1 (Figure 27. D), confirming our prediction that these interactions are enhanced upon p21-binding. Moreover, other known interactions between cyclin B1 and CDK1 do not show any changes, prompting us to propose that p21 is driving these structural rearrangements, which could provide an explanation of the specific gradient inhibition on the CDK1 kinase activity. A third region in the C-lobe of CDK1 shows a higher D_2 exchange rate from Leu176 to Trp188. These residues are closely located to the T-loop (Figure 27. E). P21 is known to pause cells during the G2/M checkpoint by inhibiting the phosphorylation of the CDK1 T-loop by the CAK⁸⁷. However, a structural rationale that explains this effect has not been described yet. In our experimental setup, the CCC is expressed as an active complex with a phosphorylated T-loop Thr161 residue. Further, p21 binds in stoichiometric amounts to the CCC complex, which is not the case when the T-loop is not in an active conformation. We propose that the loop is displaced upon binding to p21 becoming exposed to the solvent. Interestingly, AF2 is predicting the T-loop in its active form, and when compared to experimentally determined structures of the CCC⁶⁴ and the cyclin A2/CDK2/p27¹⁴⁴ complexes, we can observe that they are superposing well (Figure 27. E). On the other hand, when superposing the AF2 prediction with the crystal structure of cyclin B1/CDK1/CKS2¹⁷³, we can see that the highlighted region of CDK1 by HDX-MS does not alter its conformation, but the non-activated T-loop is positioned closer to the C-lobe. It can

be hypothesized that p21 is pushing away the T-loop, while binding to the activate CCC complex. As a consequence, residues 176-188 are exposed to the solvent, promoting a higher D₂ incorporation. Another important point to highlight is the proximity to the cyclin B1 PBP and regions in cyclin B1 showing a lower D₂ exchange rate, and therefore the potential participation of the PBP in p21-mediated regulation of CDK1. To further shed light on this mechanism, we also performed HDX-MS analyses on the PBP mutated CCC(EFE)P21 and CCC(TEM)P21 complexes, described in the following section.

4.2. HDX-MS results when the PBP is mutated

We know that p21 binds to CCC even if the pocket is mutated (Figure 12.C), demonstrated by using the TEM and EFE mutations (that notably can't bind to separate anymore). We therefore wanted to study if there was a difference in the D₂ incorporation levels when the PBP was mutated. Changes in exchange rates would hint towards a possible involvement of the PBP into the regulation of CDK1 by p21. The results that we observe in this experiment are comparing the differences between CCCP21 and CCCP21 with a mutated PBP. If we first look at the overall AF2 structure prediction of CCCP21 with the mutated PBP, the regions around the KID domain of p21 do not seem to change when the pocket is mutated, indicating that the known interactions between p21 and the CCC complex are formed the same way for both, WT and PBP-mutated CCCP21 (Figure 28. A). Interestingly, all the changes in the exchange rate are clustered in the region surrounding the PBP and the CDK1 T-loop, showing us that mutating the key residues of the PBP did cause changes in the structural organization of the CCC.

Focusing on the region that contains all the observed different D₂ exchange rates illustrates that a part of the CDK1 T-loop and activation segment (residues 150-159) shows a lower D₂ incorporation level and therefore is less exposed to the solvent (Figure 28. B). Moreover, the residues 293-305 of cyclin B1, which are facing the CDK1 T-loop, are also less exposed to the solvent. We therefore postulate that mutation of the cyclin B1 PBP abolish p21's ability to displace the T-loop and consequently the T-loop adopts an active complex conformation. Indeed, in CDK2, the region 152-163 (153-164 in CDK1) undergoes a structural rearrangement upon binding to its activating cyclin A2. The CDK2 L12 helix (148-152 in CDK1) converts into a loop upon binding to cyclin A2 and exposes the highly conserved Thr160 (Thr161 in CDK1) to allow its phosphorylation. A final rearrangement of the T-loop is needed for the full activation of the complex. These same residues in CDK1 are

highlighted by our HDX-MS results and by superimposing the structures of partially active CDK2¹⁷⁵ with our AF2 prediction, a displacement of the T-loop is visible (Figure 28. B). We can therefore propose that p21 is inhibiting active CDK1 by displacing its phosphorylated T-loop, however mutations in the PBP of cyclin B1 abolishes p21 binding leaving the T-loop in its active conformation. It is interesting to note that the CDK1 region 176-188 shows a higher D₂ exchange rate compared to the WT CCC complex, which was already enhanced upon p21 binding, implying that the mutation of the PBP exposes this region even more to the surrounding solvent.

The N-terminal helix of cyclin B1 shows a higher D₂ incorporation for the residues 170-177 compared to the N-terminal helix of cyclin B1 upon p21 binding when the cyclin B1 PBP is not mutated (figure 28. C). It has been described before that this helix is necessary for the binding interphase between cyclin B1 and CDK1. Here, we do not observe complex disassembling when cyclin B1 PBP is mutated but rather a weakened interaction between cyclin B1 N-terminal helix and CDK1. It is nevertheless difficult to correlate this observation to p21's binding as the HDX-measured residues are different than in the previous section. Indeed, we cannot exclude that weakened interaction is due to only cyclin B1 PBP mutation and changes in protein residues composition. In summary, our results indicate that mutating the PBP did not affect the binding of p21 KID to cyclin B1 or CDK1 but rather affected the region surrounding the T-loop. Because we do not have structural information if p21 is binding to this part of the complex, we cannot conclude that this observation is caused by a PBP-mediated effect on p21 binding or a pocket-mediated effect on the conformation on the T-loop, both possibilities implying exciting regulation mechanisms. The HDX-MS results are consistent between the TEM and EFE mutants for CDK1 but not cyclin B1, where changes in D₂ incorporation are more visible with the EFE mutant rather than the TEM. Such effects and results could also be attributed to a reduced stability of those mutated compared to the WT complex and therefore more challenging to analyze. To further analyze the role of the cyclin B1 PBP, we wanted to analyze the changes in the structures of CCC when mutated with HDX-MS in the absence of p21.

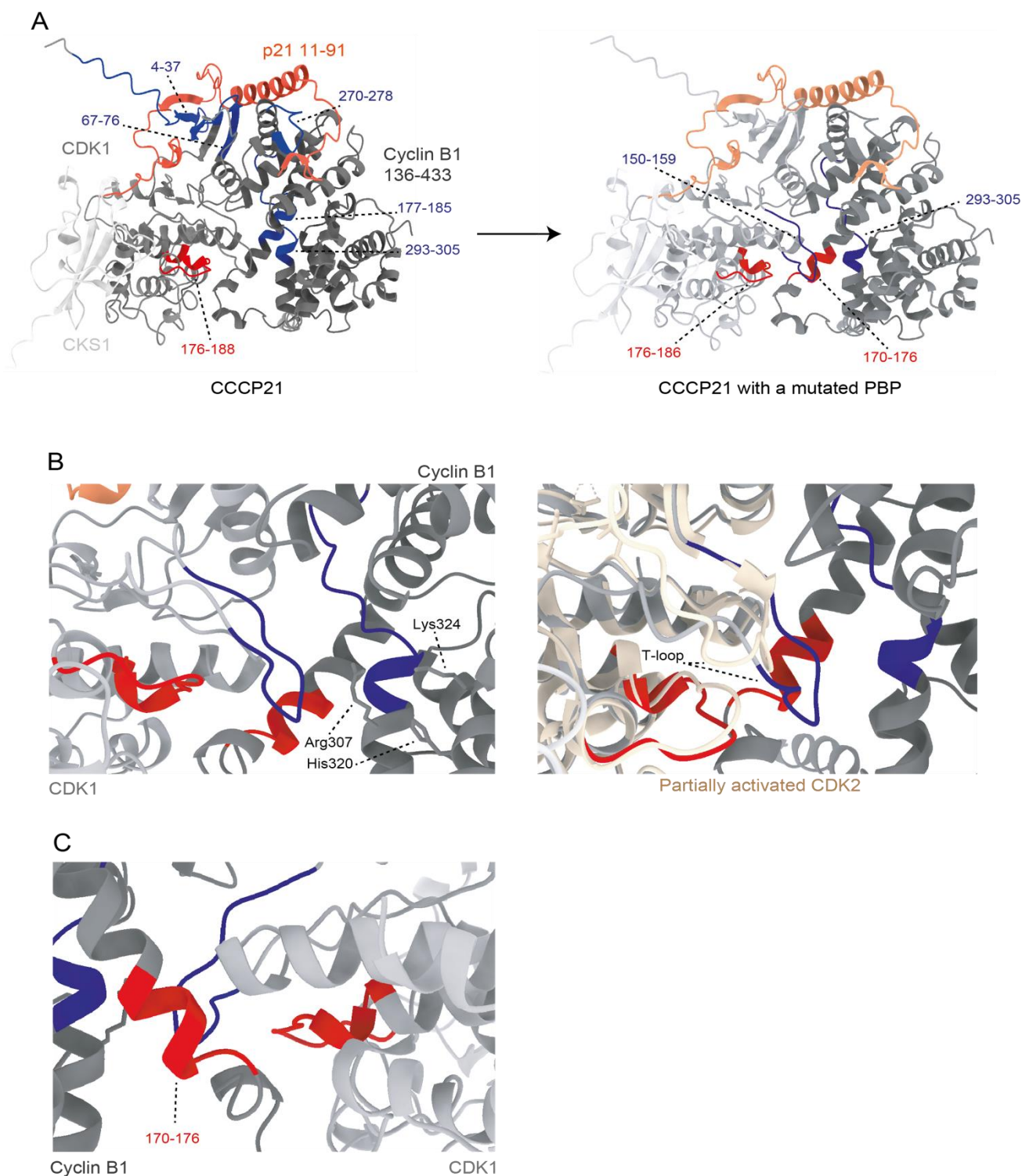


Figure 28. HDX-MS results mapped onto the AF2 predicted CCCP21 structure in the case of a mutated cyclin B1 PBP

A, On the left are the results previously described of the HDX-MS-measured differences in exchange rate upon binding of p21 to CCC displayed on the AF2 prediction of CCCP21. On the right is the regions showing a difference in exchange rate compared to CCCP21 when the PBP is mutated. The same color code is used as before to describe the results and the proteins. **B**, Close-up zoom on the four regions in cyclin B1 and CDK1 measured with HDX-MS showing differences when the PBP is mutated. CDK1 T-loop becomes more protected and its proximity to cyclin B1 PBP key residues suggests that the mutations directly played a role on the observed structural rearrangements. If we compare our AF2 prediction with highlighted regions with differences

in exchange rates with the structure of a partially activated CDK2¹⁷⁵ (PDB:1FQ1), CDK2 T-loop is placed in the back compared to active CDK1 T-loop. **C**, Close-up zoom on the back of the complex to highlight the higher exchange rate measured on cyclin B1 N-terminal helix compared to when the PBP is not mutated.

4.3. CCC EFE and TEM

Mutating cyclin B1 PBP into a TEM or a EFE mutant is causing an inversion of the positive into negative surface charges in this region. We previously observed that D₂ incorporation was increased when cyclin B1 PBP is mutated in EFE than TEM. The TEM sequence is found in cyclin A2 in contrary to EFE which is not present in any cyclin but is simply an inversion of cyclin B1 PBP positively charged residues into negatively charged ones. Our results showed that the main differences in the structural organization of the CCCP21 complex when the PBP was mutated were localized around the PBP and the surrounding positively charged patch facing CDK1 T-loop. These observations question the implication of a PBP mutation in a direct CCC structural rearrangement in the absence of p21. To assess this question, we performed HDX-MS experiments on CCC TEM and EFE and compared D₂ incorporation rate to CCC. First, a significant difference between TEM and EFE exchange rate is observed, guiding us to analyze them separately. In the case of CKS1, neither EFE nor TEM PBP mutants exhibited a significant difference in D₂ exchange rate indicating that its structural arrangement is not impacted by a mutation in cyclin B1 PBP.

Mutating cyclin B1 PBP into a EFE mutant caused strong structural rearrangements into the complex (Figure 29. A). First, in cyclin B1, a region spanning from Tyr170 to Tyr175 showed an increased D₂ exchange rate whereas a second region from Lys288 to Phe305 showed a reduced D₂ exchange rate and therefore a stabilization of the structure. For CDK1, only increased D₂ exchange rate was measured into three regions: Tyr4 to Ala31, Gly177 to Val185 and Phe214 to Glu232. Strikingly, the first measured region witnesses the presence of structural instability in the N-lobe of CDK1 and therefore in a far distance from cyclin B1 PBP. Moreover, this region englobes the glycine-rich sequence 11-17 where the key residues Thr14 and Tyr15 are located, residues that, when phosphorylated during G2/M transition, inhibit the activity of CDK1^{176,177}. This region of CDK1 has been shown to be protected upon p21-binding to the complex and we did not notice any difference in D₂ exchange rate when with a mutated PBP upon binding to p21 compared to WT CCCP21 complex. Therefore, the observed instability in mutated CCC does not impact the binding of p21 to CCC in that specific region.

If we now focus on CCC TEM (Figure 29. B), we can see that the region 288-305 in cyclin B1 showing protection to solvent in EFE is not measured anymore. The same effect is observed for the region 214-232 in CDK1 that showed a higher D₂ exchange rate in EFE. This shows that the structural effect of the TEM mutation is milder compared to EFE. Interestingly, except for CDK1 N-lobe region 4-31, the two others common to EFE and TEM CCC (cyclin B1 177-185 and CDK1 170-175) are also found upon binding to p21. We can conclude therefore that the reduced D₂ exchange and allosteric conformational changes observed here are a direct consequence of disturbances in the PBP and its surrounding region. When CCC is not mutated and upon binding to p21, cyclin B1 region 177-185, located in the N-terminal helix, gets more protected highlighting the increasing importance of the reorganization of this region for p21-mediated regulation.

All the annotations performed for this experiment are done on a cryoEM-resolved structure of active CCC⁶⁴ and, without any protein dynamics information, we can only imagine the movements induced by a mutation in cyclin B1 PBP. To attempt visualizing such conformational changes, an AF2 prediction of cyclin B1 with a EFE mutation in complex with CDK1 (CC EFE) was generated and compared to our HDX-MS results (Figure 29. C). AF2 did not predict any structural rearrangement highlighting the technical limitations of AlphaFold, capable to generate a protein structure with a high certainty to a certain extent, but incapable yet to predict protein structure dynamics. If we compare now the electrostatic surface charges of the structure of CCC compared to the AF2 prediction of CC EFE, we can visualize the drastic changes imposed by the EFE mutation and notably potentially on the position of the facing CDK1 T-loop. We can postulate from those results that changing the surface charges on cyclin B1 PBP results in conformational rearrangements in CCC complex and influences the activity of CDK1 and its regulation by p21 but potentially by other binding partners as well. To further investigate the importance of CCC conformational remodeling upon p21 binding and the role of cyclin B1 PBP, we needed to investigate the kinase activity of CDK1 in our different protein complexes *in vitro*.

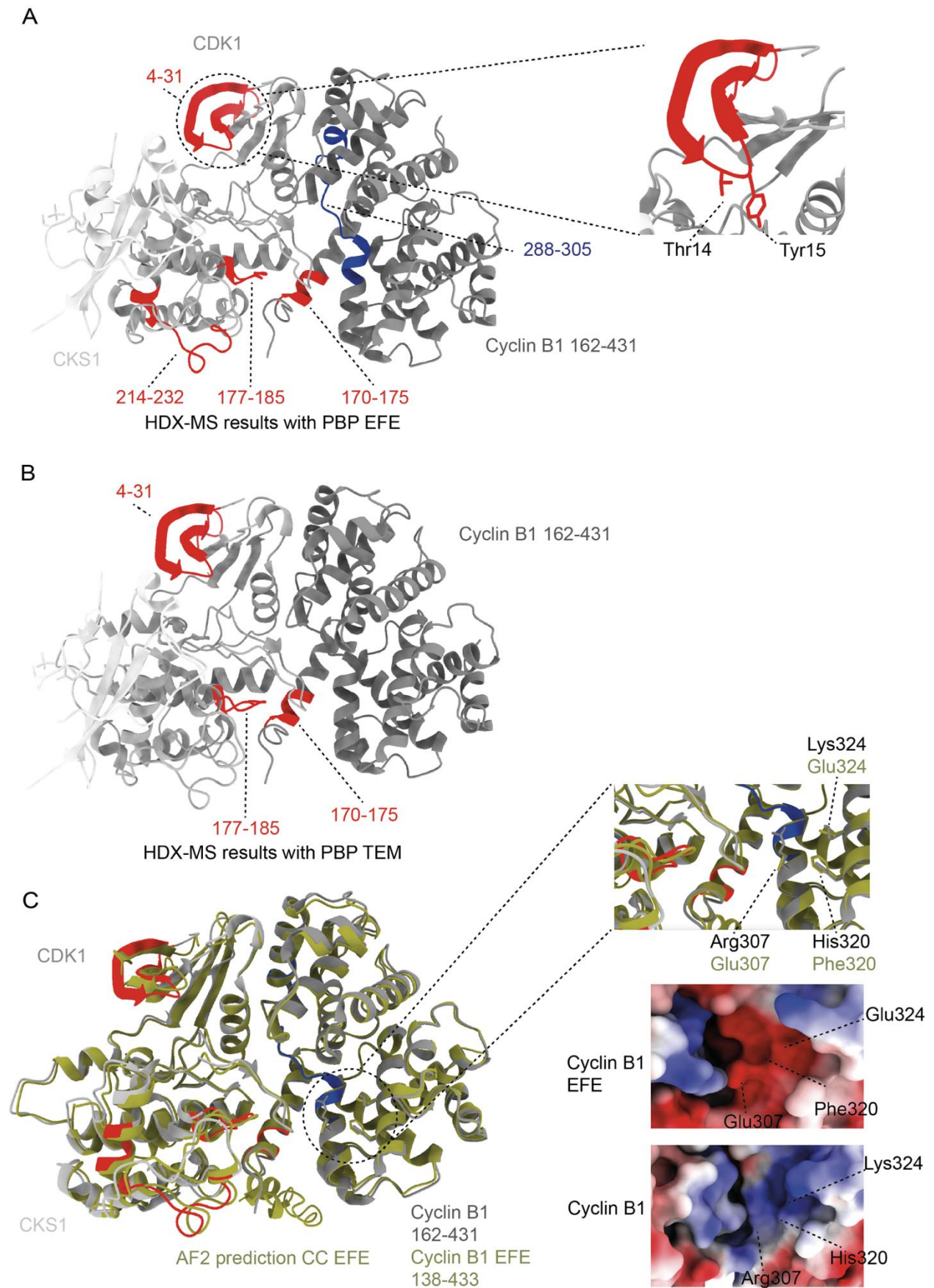


Figure 29. HDX-MS results when cyclin B1 PBP is mutated in CCC complex

A, Here are represented the differences in D₂ exchange rate measured with HDX-MS when cyclin B1 PBP is mutated into EFE. The cryoEM resolved structure of WT CCC (PDB:7NJ0⁶⁴) was used to visualize the measured regions and, same as in Figure 27, regions highlighted in blue indicate protection from exchange by the ligand, whereas regions highlighted in red show increases in H/D exchange rate. cyclin B1 is colored in dark grey, CDK1 in grey and CKS1 in light grey. On the right is a zoom on the CDK1 region 4-31 showing a higher D₂ exchange rate with its residues Thr14 and Tyr15, part of the glycine-rich loop, annotated. **B,** Here are represented the differences in D₂ exchange rate measured with HDX-MS when cyclin B1 PBP is mutated into TEM. The annotation system is the same as in A. **C,** The structure of CCC (PDB:7NJ0⁶⁴) with the annotated differences in D₂ exchange rate for the EFE mutant (in shades of grey) is compared here to the AF2 predicted structure of cyclin B1/CDK1 EFE (in olive green). On the right can be seen a zoom on the PBP region with its annotated residues for the WT (Arg307, His320 and Lys324) and EFE (Glu307, Phe320 and Glu324). Right below are displayed the electrostatic surface charges of those two proteins PBP region with the same residues annotated.

5. Activity of CCCP21 WT, TEM and EFE

To better understand the role of p21 regulating active CCC, we next analyzed the activity of the complex. As described earlier in this thesis, it is well-known that p21 inhibits the kinase activity of CDK1. However, the molecular mechanisms that underly this regulation might differ depending on if the CDK1 T-loop is phosphorylated or not. Therefore, regulation of CCC activity by p21 either happens during G2/M transition or during metaphase/anaphase transition. We conducted a series of kinase assays *in vitro* using wild-type CCC, as well as the TEM and EFE mutations of the CCC complex, with and without p21 bound, to better understand the possible role of the cyclin B1 PBP during this regulation. Using this kinase assay, we can visualize the phosphorylation of cyclin B1 by CDK1. Cyclin B1 binds CDK1 to form an active complex, but CDK1 will also phosphorylate cyclin B1 on diverse residues during mitosis allowing its progression to a fully active form. First, we assessed the possibility that the TEM and EFE mutants have a different activity than the WT CCC. To do so, we performed a kinase activity assay on the complex by monitoring the incorporation of radioactively labeled phosphate over time. After 20 minutes of reaction, all three complexes showed similar phosphorylation of cyclin B1 (Figure 30. A). We also performed the same experiment with commercial cyclin B1/CDK1 active complexes without CKS1 and noticed a slight phosphorylation of cyclin B1 after 20 minutes, but much weaker compared to when CKS1 is present (Figure 30. B).

We further analyzed the kinase activity of CCC WT, TEM and EFE upon binding to p21. First, we assessed if the presence of p21 would reduce the phosphorylation of Histone H1, a known substrate of CDK1 as described in the literature¹⁷⁸. To our surprise, we did not see

any major difference in H1 phosphorylation when p21 is present in the complex, with the exception for the EFE mutant after 3 minutes of reaction (Figure 30.C). However, this effect was no longer visible after 10 minutes and therefore we cannot correlate these results with a p21-mediated inhibition. We were surprised by this lack of inhibition on H1 phosphorylation by CDK1 but, however noticed a strong difference in cyclin B1 phosphorylation. In both, WT and mutated CCC, p21-binding inhibits the phosphorylation of cyclin B1, however in the absence of p21 it is phosphorylated (after 3 and 10 minutes of reaction). As mentioned before, the CCC purified from insect cells has a T-loop in an active conformation. Based on our finding with the HDX-MS results, we propose that CCC, when p21 is bound, is partially active, allowing the phosphorylation of targets like H1, but cannot phosphorylate its co-factor cyclin B1 or at least less efficiently compared to p21-free CCC. We also noticed a slight phosphorylation of p21 that we decided to investigate further.

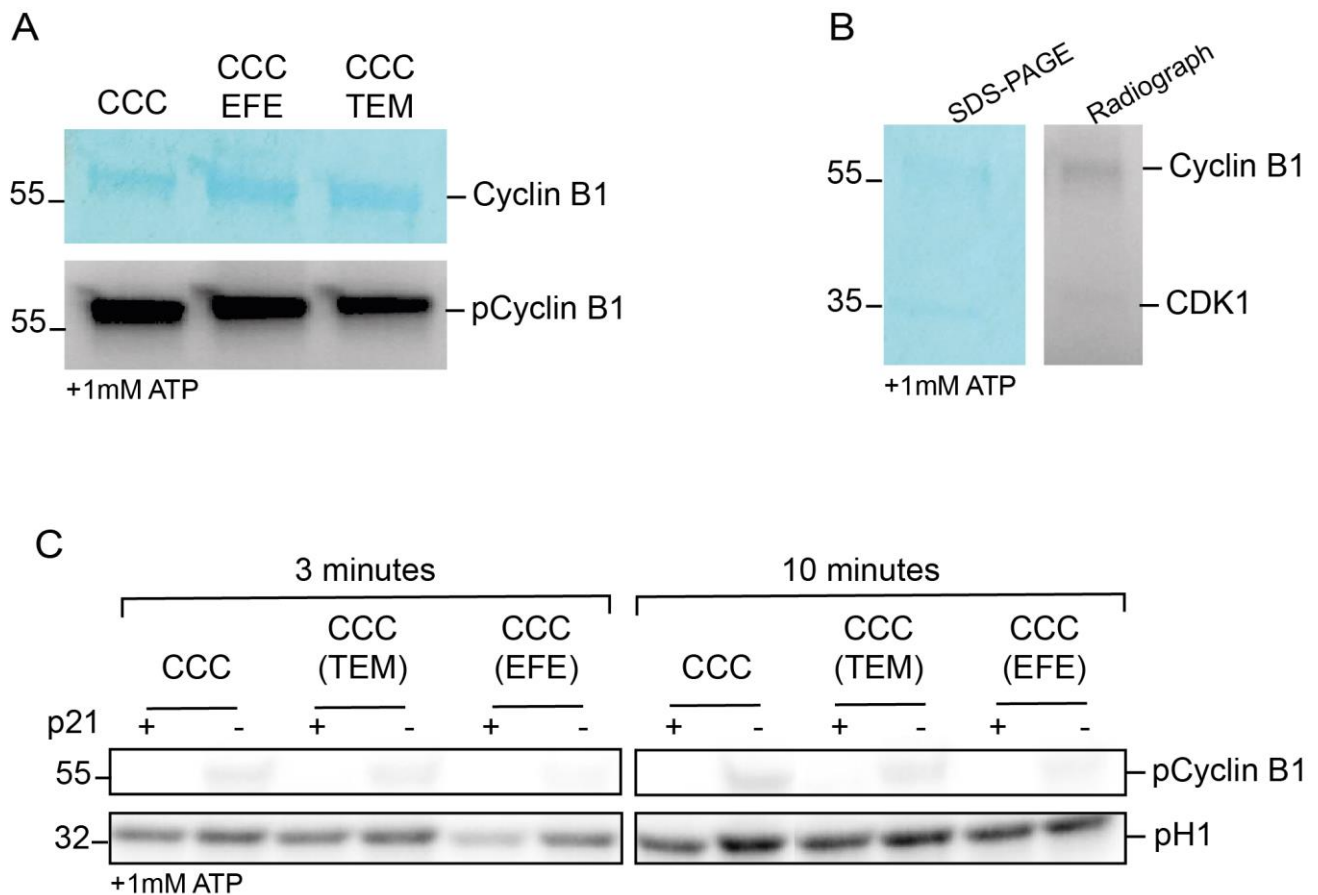


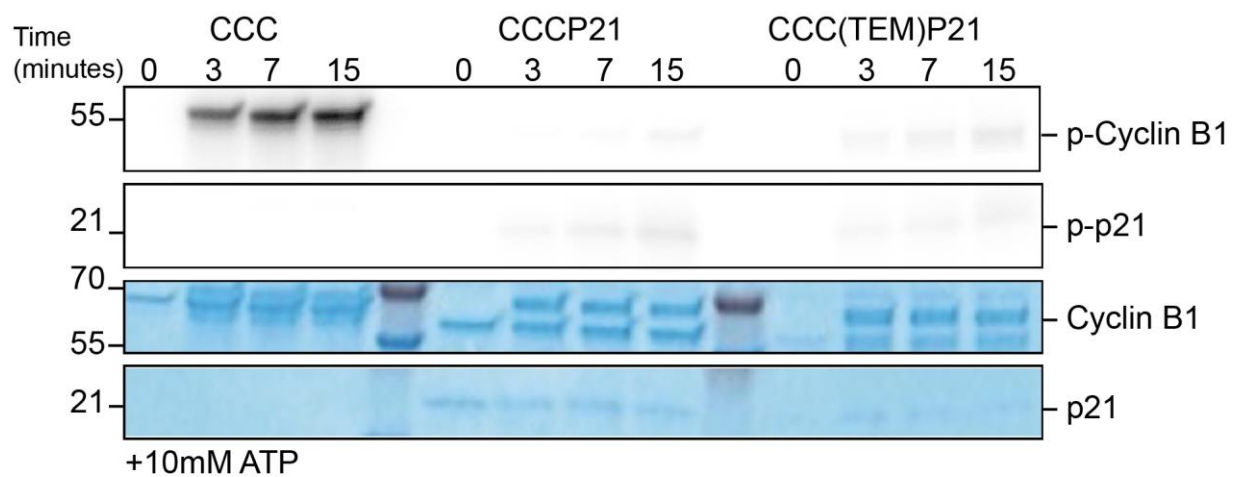
Figure 30. Activity of CCC WT and TEM and EFE mutants in p21 presence or absence

A, *In vitro* radioactive kinase assay of purified human CCC WT, EFE and TEM. The reaction took place for 20 minutes at 30°C with 1mM ATP before being loaded on an SDS-PAGE gel (upper lane). The incorporation of radioactively labelled phosphates on cyclin B1 was measured after one night in contact with a radiograph (lower lane). **B**, Same as in A but the protein used was commercially purified cyclin B1/CDK1 (ThermoFisher Scientific). **C**, *In vitro* radioactive kinase assay of purified human CCC WT, EFE and TEM and CCCP21 WT, EFE and TEM

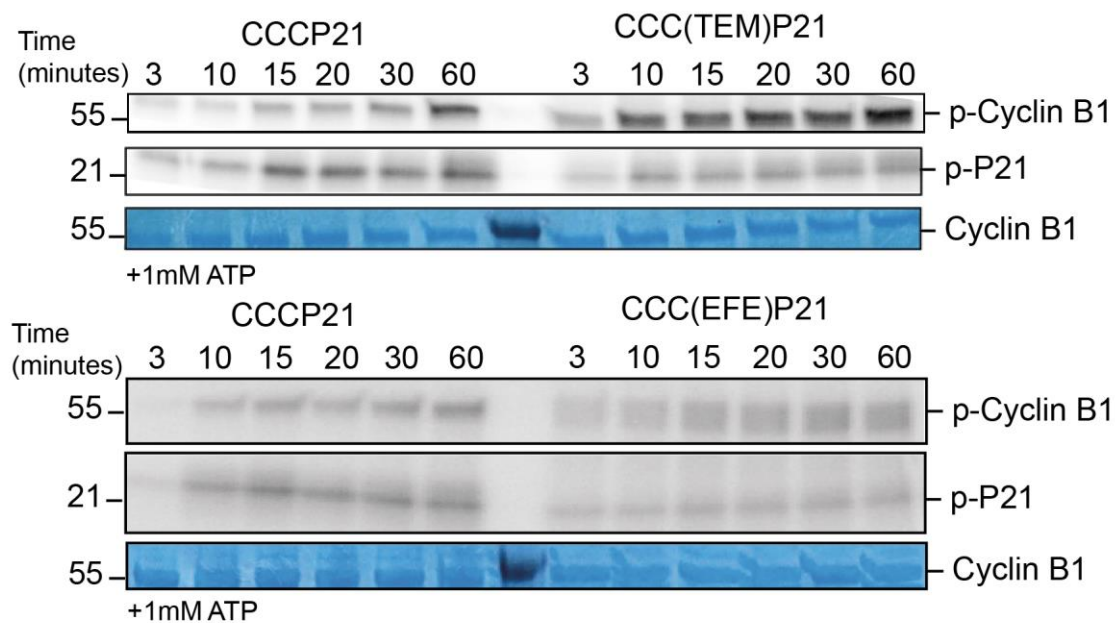
in contact with commercially purified Histone H1. The reaction was stopped after 3 and 10 minutes at 30°C with 1mM ATP and loaded on an SDS-PAGE gel before being transferred to a radiograph for one night. The incorporation of radioactively labelled phosphates on cyclin B1 and Histone H1 is observed here.

We performed a new kinase activity assay on CCCP21 but this time excluding H1 to focus only on the phosphorylation of cyclin B1 and p21 (Figure 31. A). We increased the concentration of ATP to provoke a strong reaction and analyzed the phosphorylation state of each of the proteins in the CCC, CCCP21 and CCC(TEM)P21 complexes at different time points. The appearance of a double band for cyclin B1 can be observed, an artefact caused by sample boiling prior to loading the sample on an SDS-PAGE gel. After 3 minutes, cyclin B1 is strongly phosphorylated in CCC and seems to reach a plateau after 7 minutes of reaction. Upon binding to p21, this phosphorylation is almost absent. Slight phosphorylation of cyclin B1 can be observed after 7 minutes. When cyclin B1 PBP is mutated in TEM, the phosphorylation of cyclin B1 is still extremely low compared to CCC but faster than in WT CCCP21. A progressive p21 phosphorylation can be detected in both CCCP21 and CCC(TEM)P21, roughly three minutes after the start of the reaction. To further investigate the effect of the TEM mutation on cyclin B1 phosphorylation upon binding to p21, we performed a kinase activity assay with longer time points comparing CCCP21 and CCC(TEM)P21 (Figure 31. B). A slow but progressive phosphorylation of cyclin B1 is observed in CCCP21, which is much faster when the PBP is mutated. Here, a strong phosphorylation can already be observed after 10 minutes of reaction. We do not detect strong differences in p21 phosphorylation in both conditions, indicating that this effect is independent of the integrity of the cyclin B1 PBP. The same experiment was performed with CCC(TEM)P21, but the results were not as clear as for the TEM mutant, possibly due to a reduced stability of this complex during the kinase reaction at 30°C. Still, we can observe that after 3 minutes, cyclin B1 is less phosphorylated in the WT than in the EFE mutant. It has been demonstrated that p21 can bind to more than one cyclin B1/CDK1 complex during G2/M transition. However, we did not observe a change in complex stoichiometries that would confirm this observation. To test if stoichiometries can change upon cyclin B1 and p21 phosphorylation, we performed analytical size exclusion chromatography runs before and after a 15-minute kinase reaction (figure 31. C). While we did observe a slight shift of the elution volume, we conclude that the different elution behavior is most likely caused by the phosphorylation of the complex but not higher molecular weight complexes.

A



B



C

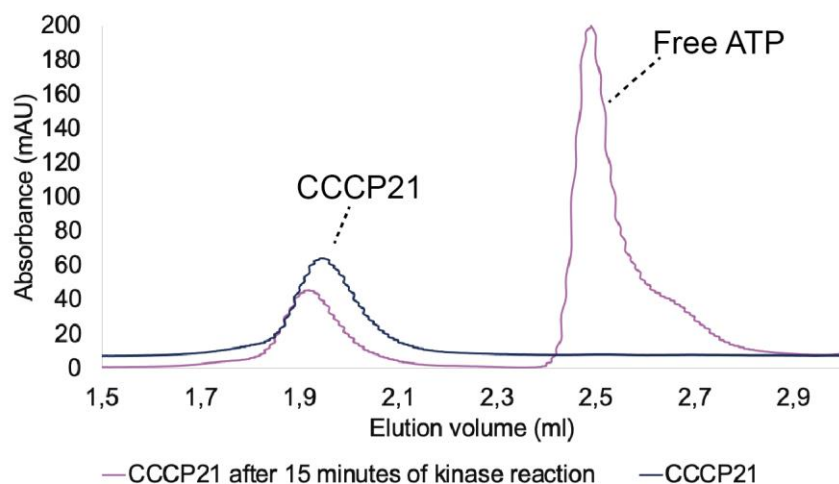


Figure 31. Activity of CCC autophosphorylation upon p21's binding with and without a mutated PBP

A, *In vitro* radioactive kinase assay of purified human CCC, CCCP21 and CCC(TEM)P21. The reaction was stopped after 0, 3, 7 and 15 minutes at 30°C with 10mM ATP and loaded on an SDS-PAGE gel (lower lanes) before being transferred to a radiograph for one hour. The incorporation of radioactively labelled phosphates on cyclin B1 and p21 is visible on the upper lanes. **B**, *In vitro* kinase assay of purified human CCCP21, CCC(TEM)P21 and CCC(EFE)P21. The reaction was stopped after 3, 10, 15, 20, 30 and 60 minutes at 30°C with 10mM ATP and loaded on an SDS-PAGE gel (lower lane for each mutant) before being transferred to a radiograph for one hour. The incorporation of radioactively labelled phosphates on cyclin B1 and p21 is visible for each mutant on the upper lanes. **C**, Result of analytical size-exclusion chromatography assays on CCCP21 before and after a 15-minutes kinase reaction with 1mM ATP at 30°C.

To better understand the progressive phosphorylation of WT cyclin B1, we sent the protein after 15 minutes of kinase reaction to the EMBL mass-spectrometry facility to identify phosphorylated residues. Our preliminary analyses show that 9 sites have been identified to be phosphorylated during this first try with high confidence. Most of them are in the unstructured N-terminal region of cyclin B1 (Thr6, Thr34, Thr43, Ser69 or Thr71 and Ser95) and 3 others in the C-CBD (Ser310, Thr395 and Thr399) (Figure 32. A). Ser310 is part of the RRASK motif described before as essential for CDK1 activation and CDK's capacity to phosphorylate its substrates. This motif was shown to be close to an HDX-MS-demonstrated solvent-protected area upon p21 binding and could play a role for p21 phosphorylation by CDK1. This region is also where the PBP is found as the first Arg of the RRASK motif is one of the key residues of this pocket (Arg307). Moreover, in the cyclin B1/CDK2 crystal structure, the Ser310 side chain is shown to hydrogen bond with His320⁵⁹, another PBP key residue. As showed before, without CKS1, these phosphorylations do not appear on the radiograph and are likely stimulated by CKS1 presence, which is known to enhance CDK1 multisite phosphorylation and in particular phosphorylations of non-consensus CDK1 sites. Analyzing the MS-identified phospho-sites, reveals that none of them are bearing the consensus CDK1 S/T-P-X-K/R site. Furthermore, it has been shown recently that, *in vitro*, cyclin A2, by binding CDK1, promotes the phosphorylation of non-proline-driven S/T-x-x-K consensus sites on its substrates, where the first x is any residue except a proline¹⁷⁹. This is also the case for cyclin B1. If CKS1 is present, phosphorylation of sites where the +1 residue is not a proline are enhanced, with no differences in the RxL motif in the substrates. We propose that CDK1 phosphorylates cyclin B1 on those non-proline driven consensus sites with the help of CKS1. This specific kinase activity is inhibited by p21 but if cyclin B1 PBP is mutated, this inhibition is less strong due to the release of the CDK1 T-loop in its active conformation. We also sent phosphorylated p21 to identify the phospho-sites of 21, and in a first experiment, we detected eight phosphorylated residues, and six of them with high confidence (Thr80, Ser130, Ser137, Thr148, Tyr151 and Ser153) (figure 32. B). Only one CDK1 consensus site is found, namely

Ser130 (S-P) known to be phosphorylated by cyclin B1/CDK1 and to stimulate p21 ubiquitination¹³⁸. The Thr80 residue is located just after the 3₁₀ helix and may have a role in the inhibition of the CDK1 ATP-binding pocket. All other residues are localized in a proposed second D1 region where the second Cy domain is found¹⁴⁶. This region has been shown to interact with cyclin A/CDK2 in a non-inhibitory mode independently to the KID, which we never observed for cyclin B1/CDK1. Tyrosine phosphorylation of Tyr151 and not Tyr77 was detected in our results, and it has been proposed in the literature that this first one was resulting into a weaker interaction of the second C-terminal D1 region, and therefore the Cy2 domain, with cyclin A2¹⁴⁶. Replications of this experiment are needed to validate our first observations and to propose data-driven ideas about the role of the p21 Cy2 domain on cyclin B1/CDK1 and their phosphorylated residues.

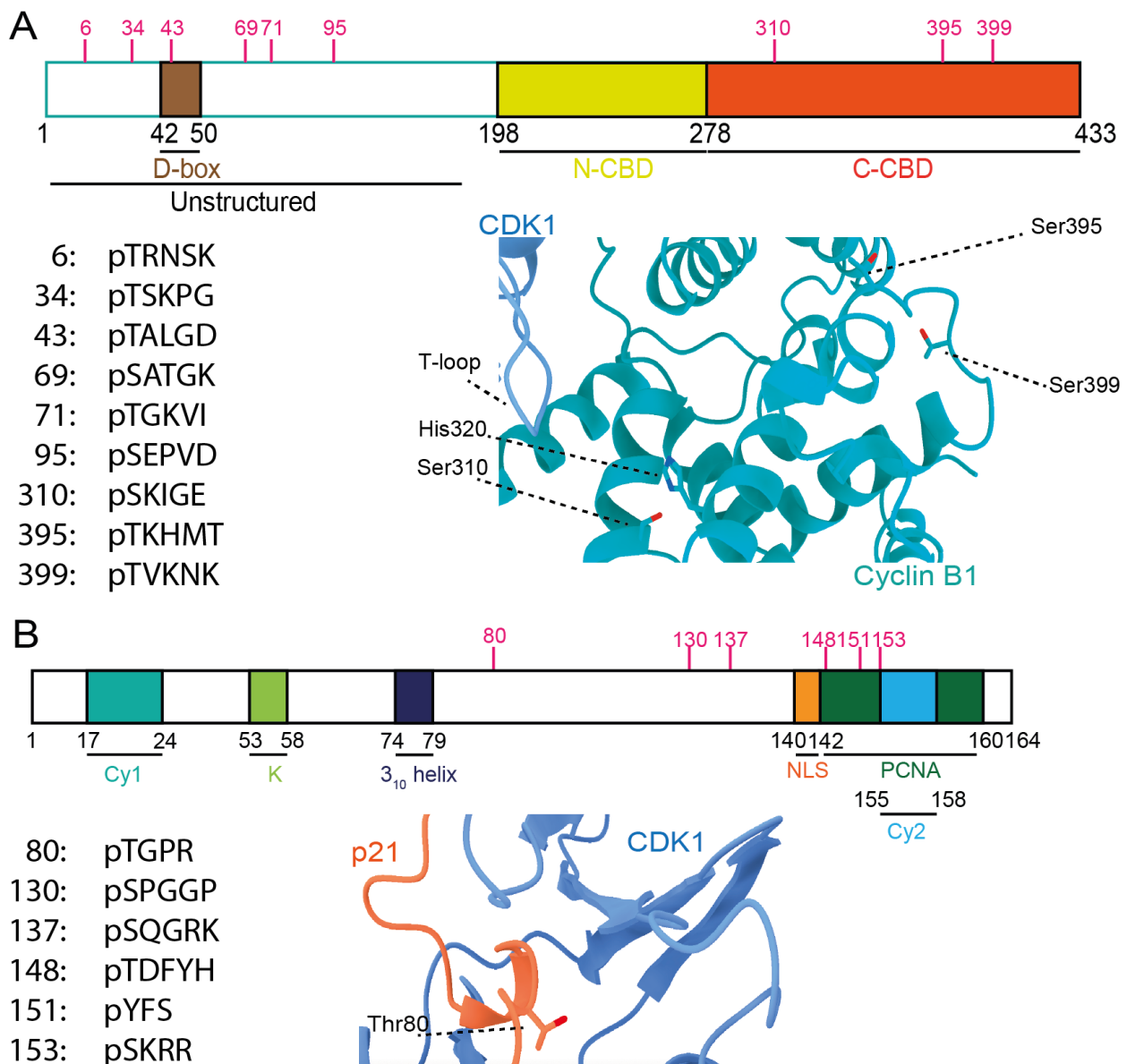


Figure 32. Identification of phosphorylated residues by mass-spectrometry on cyclin B1 and p21 after an *in vitro* kinase reaction.

A, First results on identified phosphorylated Serines and Threonines on cyclin B1 upon its phosphorylation by CDK1 *in vitro* in a CCCP21 complex. The schematic representation of the protein is showing the N-terminal unstructured part of the protein containing the D-box. cyclin B1 CBD starts from the residue 198 and the N- and C-CBD are annotated. Every identified phosphorylated residue is annotated in pink based on their position in the protein. Each of these residues and their following sequences are listed below. On the right part is the AF2 predicted CCCP21 structure with the identified phosphorylated residues on the structured part of cyclin B1 (CBD). **B,** First results on identified phosphorylated Serines, Threonines and Tyrosine on p21 upon its phosphorylation by CDK1 *in vitro* in a CCCP21 complex. Same as in A, the schematic representation of p21 shows the different domains of p21 described before in this thesis and the identified phosphorylated residues are annotated in pink based on their position in the protein. The phosphorylated residues and the following sequences are listed below. On the right in the AF2 predicted CCCP21 structure with the only identified phosphorylated residue (Thr80) predicted to bind to CCC and more specifically CDK1.

DISCUSSION

P21 was originally described as a universal inhibitor of cell cycle-regulatory cyclin/CDK complexes. In recent years, p21 role in the complex formation of certain cyclin/CDK¹⁴⁰ and the regulation of complex activity, as well as non-inhibitory roles has been studied. Little is known about p21 regulation of CDK1 and specially about its structural organization. During my thesis, I attempted to understand better how p21 binds and regulates CDK1 when bound to cyclin B1 and CKS1, and if the recently discovered cyclin B1 Phosphate-Binding Pocket and its three key residues play a role in these. We demonstrated that p21 is a very flexible protein making full-length analysis by conventional structure biology like cryo-EM difficult and to understand better the regulation of the CCC complex by p21, structural biology needs to be complemented with other biophysical techniques, such as mass-spectrometry or functional assays (kinase activity). I will now discuss the results and the resulting conclusions to design the future of this project and continue shedding light on the structural relationship between p21, cyclin B1, CDK1 and CKS1 during mitotic events.

1. Structural information about CCCP21

I first managed to purify and express soluble and stable CCCP21 complexes. The complex is found in a quaternary stable form and does not form higher order structures. Our experimental setup might however favor the formation of stable CCCP21 complexes because p21 proteins are applied in excess to purified CCC complexes, potentially forcing the formation of stable complexes. It would be interesting to test if the available amount of free protein plays an important role in complex formation, for example by reducing the amount of available p21. The formation of the complex does not seem to be impacted by the cyclin B1 pocket mutants – the previously described TEM and EFE mutations – unlike for separase where binding to the CCC complex is fully disrupted using a TEM or EFE mutant. Therefore, we conclude that p21 does not require the pocket to bind the CCC complex. Several domains have been described in p21 with diverse roles and binding partners, but our results prove that only the CDK-binding domain K has a strong binding affinity to the complex or to CDK1 in isolation, and not the remaining sequences of the protein. It is interesting to note that the two cyclin-binding domains Cy could not bind efficiently alone to the complex further substantiating the importance of the K domain and most likely, the preceding LH alpha helix, one of the few folded regions of the protein. To better understand complex formation and

organization on an atomic level, we tried to solve the structure of full-length p21 bound to cyclin B1, CDK1 and CKS1 using cryo-EM. Here, we could obtain a 7.4 Å resolution EM density map that allowed us to confidentially place an AF2-generated model into the density. By changing the EM-grid types, adding detergents and changing the buffer composition, we tried to improve the quality of the grids and consequently of the CCCP21 particles, with the aim to be able to reach a higher resolution, however so far, we were not able to obtain a resolution that allows the correct placement of side chains. It will be interesting to apply novel refinement methodologies such as BLUSH¹⁸⁰, now released with RELION 5.0, to test if they would improve structure determination at higher resolutions. In addition, electron filter such as the Selectris X filter should result in higher quality images, especially important to identify the correct particle alignments of small proteins that possess a much lower signal-to-noise ratio than large macromolecular complexes such as ribosomes. Other ideas that could help to improve the resolution are related to the proteins that we are using. Cyclin B1 full-length protein contains a large N-terminal unstructured region (1-136) that could perturb the distribution of the particles in ice and their alignment during data processing. Moreover, in the different structures available of cyclin B1, this region is deleted to promote protein stability. Removing this region and test if we can obtain better aligned particles can potentially represent another future research direction, but we cannot exclude that these residues are recognized by p21 and that their deletion may impact binding and regulation of the CCC complex. P21 is an unstructured and instable protein. While it is known that the KID domain mediates binding to cyclin B1 and CDK1 (and thereby also stabilizes the CC complex), the organization of the other parts of the protein are still unknown. The development of AlphaFold 2 during my thesis allowed me to predict the assembly of the CCCP21 complex. I used these structure predictions to generate hypothesis that were then tested experimentally. Importantly, AF2 predicts only the KID domain of p21 to bind to the CCC, while the rest of the protein is predicted to flop around the structure. So far, only the structures of the KID domain of p21 and p27 have been solved bound to cyclin/CDK complexes potentially introducing a prediction bias. In addition, post-translational modifications that can influence the binding mode of complex subunits are not considered. We also know that the N-terminal Cy1 domain, containing the RxL motif, binds to cyclin B1. From the D1 section emerges the LH α -helix that bridges cyclin B1 to the CDK1 N-lobe, on which the D2 section will sit rearranging the architecture of CDK1 β -strands allowing the 3₁₀ helix of p21 to enter the CDK1 catalytic cleft and inhibit its kinase activity. Our medium-resolution EM density map of CCCP21, did not allow us to confirm precisely those predictions and we therefore decided to

complement our structural studies with HDX-MS, to better understand which part of the CCC are targeted by p21 binding. Using HDX-MS, we observed that a region on cyclin B1 composed of mainly conserved residues amongst cell cycle cyclins, and close to the MRAIL hydrophobic patch, gets protected upon p21 binding. We thus propose that this region is the binding site of the RxL motif of p21, which is in line with the AF2 predictions. It has been postulated before that p21 binds first to cyclin B1 through its D1 segment and subsequently binds to CDK1. Mutating this potential RxL-recognizing patch in cyclin B1, especially Ala276, and potentially observe a reduced binding affinity of p21 and as a consequence a higher activity of the CCC complex could be an interesting experiment to perform. HDX-MS confirmed that p21 binds to the CDK1 N-lobe, as expected, and rearranges its structure to facilitate the inhibition of the CDK1 kinase activity. These results confirmed the AF2 prediction, and existing literature as well as determined structures of homologous proteins. HDX-MS also revealed rearrangements in cyclin B1 and CDK1 upon binding to p21 in regions where p21 was not predicted by AF2 to bind to the complex. AF2 predicted that p21 binds to the CCC complex up to the 3₁₀ helix, but interestingly it was shown that a p21 mutation found in human breast carcinoma changing Arg94 into Trp94 impairs the inhibition of CDK1 by p21¹¹². This result shows that contact sites between p21 and the CCC complex extend further than AF2 is predicting, and that these interactions play a key role in the regulation of CDK1 by p21. In our HDX-MS measurements we detected that the cyclin B1 N-terminal α -helix becomes more protected in presence of p21, however this effect is unlikely to be caused by directing protein-protein interactions. Instead, this α -helix is known to be important for binding to CDK1 and for the formation of CC complexes, which implies, that p21 is stabilizing this interaction between the two proteins, by acting like a sealing rope that glues the two proteins together – a mechanism important for p21-mediated regulation of CDK1 kinase activity. We speculate that at least in part this mechanism is governed by the cyclin B1 PBP and surrounding residues like the RRASK motif, because those regions get more protected upon the binding of p21. Similar HDX-MS experiments have been performed on active cyclin D1/CDK4 (phosphorylated Thr172) complexes. Here, CDK4 shows a higher protection of the N-lobe and specifically around the glycine rich phosphate binding sequence composed of residues 11-17 upon binding to p21¹²⁹. Similar results are also obtained in our experiments with CDK1. The rest of the CDK4 structure is not impacted, like CDK1, and the slow gradient injection of an inhibitor displaces p21 but not p27 which has more interaction sites with CDK4. It is plausible that a similar mechanism applies to CDK1, where the association of p21 remains relatively slow when Thr161 is phosphorylated and therefore p21 inhibits temporarily

CCC, before releasing it in metaphase. Moreover, it has been shown that p21 does not prevent dephosphorylation of CDK1 Thr14 and Tyr15 by Cdc25¹²⁸, but our HDX-MS results show protection of those unphosphorylated residues. In our experiments, CDK1 is already in its unphosphorylated, active form confirmed by phosphoMS and activity assays. Protection and rearrangement of those residues could be an important role for p21 to assure the inhibition and delay of the activation of CCC, until enough substrate is present. However, this does not explain why p21 can bind inactive and active CDK1 and if the role of p21 after entering M phase is to retain the T-loop in a semi-active conformation until its dislodgement in metaphase following its ubiquitination.

2. The important role of dynamic disorder in p21-mediated regulation and its regulation by post-translational modifications

Another important point to consider, is the unstructured and therefore flexible nature of p21 and how this impacts target binding. Research on protein dynamics is not an easy task and in the case of p21 or p27, the phosphorylation status of the protein mainly determines its binding mode to the target. Most of the previous studies on the protein dynamics focus on p27 and only to a lesser extent on p21. Existing literature and some of the results generated in this thesis indicate that those two proteins do not bind their target identically. The flexibility of p21 and p27 make them challenging targets to be analyzed purely by conventional structural biology techniques. For many years, crystallography was the dominant technique used to obtain high-resolution structure information. However, the formation of crystals is impaired by flexible protein regions and those have traditionally been removed. In the case of p27, the region of p27 protein co-crystallized with cyclin/CDK complexes was by design restricted to the KID domain. The development of cryoEM and its capacity to analyze intrinsically disordered regions of proteins, prompted us to structurally revisit this complex to deepen our understanding of p21-mediated regulation. Here, we aimed at structurally visualizing for the first time full-length p21 bound to the CCC complex. However, even after extensive screening attempts, we were not able to obtain a high-resolution structure that would allow us to precisely describe and visualize p21-binding to the CCC complex. It is known that heterogeneity in a sample impairs particle averaging and, as a consequence high-resolution reconstruction difficult. P21 and p27 are dynamic and disordered proteins¹⁸¹, that can adopt different conformational states. This flexibility allows IDPs such as p21 to interact

and regulate many partners in different ways and in different timings. It has been described that IDPs are normally less conserved than their binding partner and it is believed that their conserved folding-on-binding amino acid sequence co-evolves with their target¹⁸². Understanding the evolutionary relationship between two protein binding partners is necessary to investigate their specific role. We do not possess such tool yet, but other techniques exist to investigate the role of the binding dynamics of p21 to the CCC complex and how post-translational modifications, especially phosphorylations, control such interactions. P21 contains three consensus CDK1 S/T-P-X-K/R sites: Thr57, Ser98 and Ser130. It has already been experimentally demonstrated that CDK1 can phosphorylate Thr57 and Ser130^{128,138}. Ser130 phosphorylation by CDK1 reduces the stability of p21-binding to the CCC complex and promotes its degradation, suggesting that this is an event happening during later stages of mitosis. Our initial mass-spectrometry experiments indicate that this residue is phosphorylated in our complex after a kinase reaction, however our complex is not disassembled. One possible explanation is that the T57E mutation glues p21 to the CCC complex. Repeating this experiment using WT p21 would be necessary to exclude this possibility. Interestingly, p21 can be phosphorylated at two tyrosine sites, namely Tyr77 and Tyr151 by a Non-Receptor Tyrosine Kinase (NRTK). By Mass Spectrometry, we only observe the tyrosine at position 151 phosphorylated, but a phosphorylation at Tyr77 is an important step (together with Ser130) towards p21 degradation and the release of the CCC complex. Earlier in this thesis, I discuss the role of Tyr77 phosphorylation, less precise than its homologue in p27 Tyr88, for the ejection of the 3_{10} helix out of CDK1 ATP-binding pocket. Displacement of p27 out of the CDK4 ATP-binding pocket following its phosphorylation on Tyr88 is important¹⁸³. However, to generate a partially active cyclin D/CDK4 complex and allow the phosphorylation of its T-loop Thr174 by the CAK complex, as well as the phosphorylation of other residues of p27 that results in its degradation, the phosphorylation of two other tyrosine residues (Tyr74 and Tyr79) is necessary¹⁸⁴. These tyrosine residues do not exist in p21. Therefore, the mechanism for p21's displacement from the CCC complex remains enigmatic. Just below the 3_{10} -helix and close to the critical Arg94, Mass Spectrometry showed a phosphorylation on Thr80. We could postulate here that, in the case of p21, residues closer to the T-loop could be the key of p21-mediated regulation of a CDK, together with residues responsible for ejection of the 3_{10} helix by phosphorylation of Tyr77. An important information to note is that in our experimental setup, p21 is not co-expressed with CCC in insect cells, but separately in bacteria which results in an unphosphorylated form of p21. Therefore, the observed phosphorylations are most likely added by the CCC complex.

Indeed, we did not detect any phosphorylation on Tyr77, which is known to be phosphorylated by NRTKs and not by CDK1¹⁴⁶. Unlike p27, p21 contains a second Cy motif and this region was described as a potential second D1 fragment (D1c)¹⁴⁶, capable of binding to cyclins through its RxL motif. Experiments on cyclin A/CDK2 show that the phosphorylation of Tyr151 by NRTKs from the Src family is causing a reduced interaction between the D1c region and cyclin A, and a reduced phosphorylation of Ser130¹⁸⁵. This leads to a stabilization of the complex, until the phosphorylation of Tyr77 causes enhanced phosphorylation of Ser130, resulting in the degradation of p21. It seems therefore that our experiments are leading to the formation of such stabilized complexes, but we cannot explain how Tyr151 was phosphorylated because none of the NRTKs are present in our buffer. A non-specific phosphorylation due to high concentrations of the CCC complex in the buffer seems the most logical explanation. Tyr151 phosphorylation might also promote the formation of stoichiometric complexes, because if p21 was not phosphorylated on Tyr151, it could potentially bind another CCC complex through its second RxL motif. However, we strongly suggest here that the second motif is more likely to compete with the first RxL motif for the same cyclin B1 instead of forming higher-order complexes. Future studies should address this possibility by mutating Tyr151, and by co-expressing p21 with CCC in insect cells. Interestingly, amongst all phospho-sites identified in our MS studies, only Ser130 contains a S/TP CDK1 consensus site, whereas all other sites are non-consensus sites. This result likely highlights the role of CKS1 in phosphorylating non-consensus CDK sites. Unlike p27, p21 does not bind CKS proteins as shown by our HDX-MS results. However, we do not know if CKS1 stimulates CDK1-dependent p21 phosphorylation and thereby mediates the regulation of p21. Our MS results are preliminary and need to be repeated several times. Further, the role of CKS1 needs to be further dissected. Three p21 phosphorylations are found in the PCNA interacting domain and are known to play a role for this interaction, independent of cyclin/CDK complexes. In summary, investigating the phosphorylation status of p21 and correlating it to structural as well as functional studies like HDX-MS or cross-linking MS is the next step to better understand the enigmatic regulation of the kinase activity of CDK1 by p21. Another way to analyze the flexible nature of the IDP p21, is to perform molecular dynamics simulations combined with NMR spectroscopy experiments using the CCCP21 complex. NMR spectroscopy can provide information about conformational changes of proteins in time and in solution and such experiment has already been conducted highlighting the variety of p27 conformations bound to cyclin A2/CDK2 and was also correlated to molecular dynamics simulations¹⁸⁶. These studies provided dynamic insights into p27-binding to the complex,

considering its phosphorylation status and this gives us precious insights to improve our cryoEM reconstruction if the same mechanisms exist for p21. It would be necessary to combine these *in vitro* data to *in vivo* experiments correlating the availability of p21 and cellular localization of CCC to the release of its activity.

3. Cyclin B1 surface charges and hydrophobicity role on p21's binding and regulation

This work on p21 originated from the hypothesis that p21 regulation is controlled by its phosphorylation status that mediates binding to the cyclin B1 PBP. Cyclin B1 contains unique features in its CBD distinguishing it from other cyclins and we believe that these features lead to the specific regulation of CDK1 by p21 during mitosis. First, if we compare the RxL recognition site of cyclin B1 to cyclin A2, we can see that the surface charge of cyclin A2 is mainly negative, whereas the surface charge of cyclin B1 is mainly neutral, except for the region identified to be protected upon binding to p21 by our HDX-MS experiments (Figure 33.A). It has been hypothesized in a previous study that a change in charges, for example by replacing the negatively charged cyclin A2 Glu224 with a positively charged cyclin B1 Lys215, could impair substrate binding⁵⁹. We therefore propose that these differences in surface charge between cyclin B1 and cyclin A regulate p21-binding and allow for a specific regulation of the cyclin B1/CDK1/CKS1 complex compared to other cyclin/CDK complexes. P21 would be pushed closer to Asp274 of cyclin B1, part of the HDX-MS protected area (Figure 33. B). To assess this possibility, performing mutational analyses of this residue by converting it into a neutral or positively charged residue could be an interesting starting point to gain better insights into this interaction. This could also help to explain why altering the size of the LH segment does not have the same effect on all CCC complexes¹²⁵. In comparison, there is no drastic difference between the hydrophobicity of cyclin B1 and cyclin A2 in this region. A hydrophobic belt is known to direct B-type cyclins to centrosomes during mitosis¹⁸⁷ (Figure 33. A). Nevertheless, this region is hidden by p21 in our AF2 prediction (Figure 33. B). One could speculate that p21 inhibits CDK1 movements to centrosomes and capacity to phosphorylate its substrates until the right moment during mitosis. Other differences between cyclin B1 and cyclin A2 that I have already discussed in this thesis, are the cyclin PBP and its surrounding positively charged residues compared to the negatively charged region in cyclin A2 (Figure 33. C).

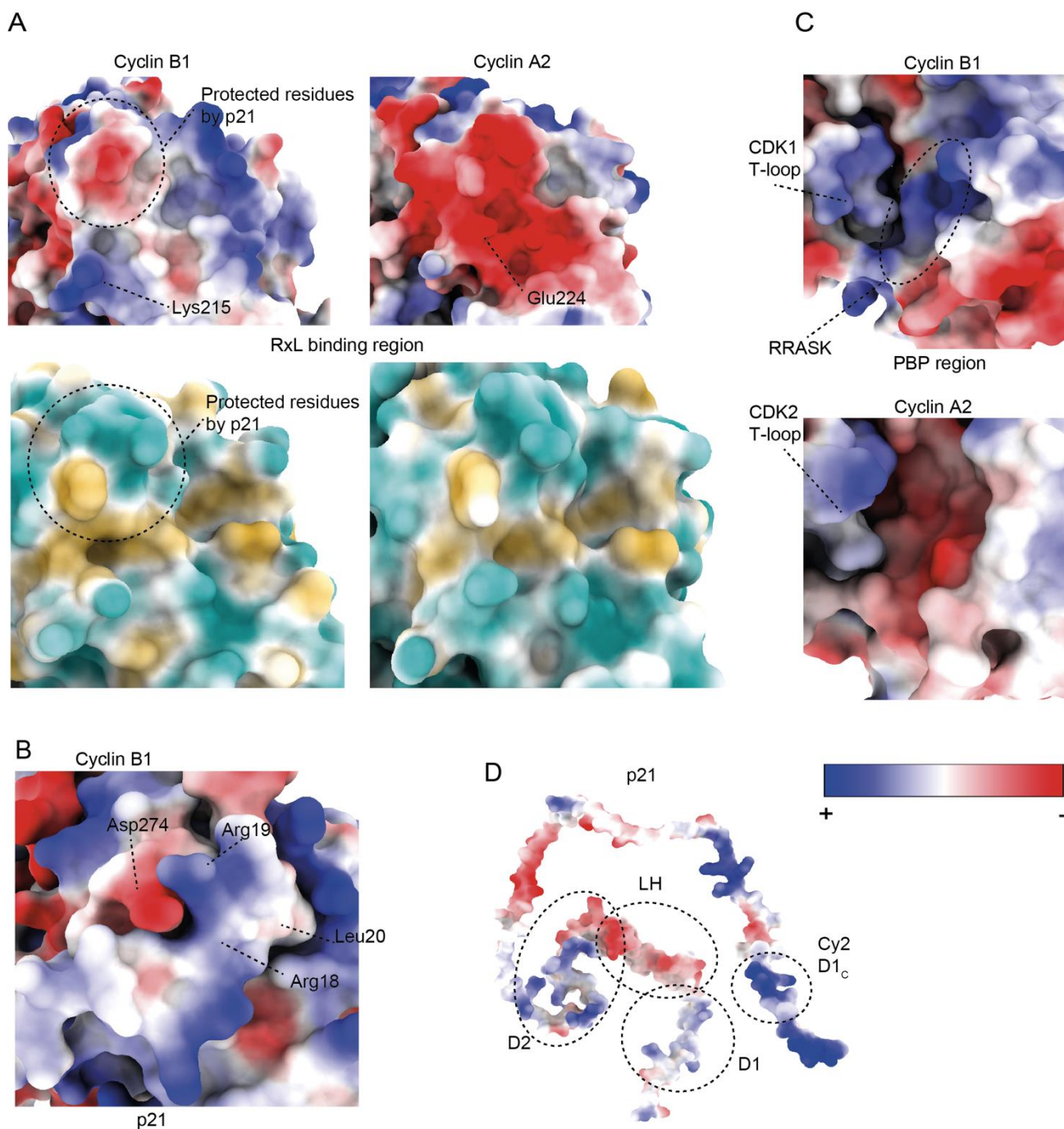


Figure 33. Implication of cyclin B1 electrostatic surface charges into the regulation of CCC by p21

A, In the upper panels are compared the electrostatic surface charges of the RxL binding sites of cyclin B1⁶⁴ (PDB:7NJ0) and cyclin A2 (PDB:1QMZ¹⁸⁸) with highlighted on cyclin B1 the protected region observed with HDX-MS upon binding to p21. The homologous residues Lys215 and Glu224 are also annotated showing the drastic differences of charges in these key residues of this region. In the lower panels are compared hydrophobicity of the same proteins showing no drastic differences between the two. **B**, Here are represented the electrostatic surface charges of p21 RxL (Arg18, Arg19 and Leu20) on cyclin B1 with a highlight on the facing negatively charged Asp274, residue found protect in our HDX-MS results. This structure is our AF2 CCCP21 prediction **C**, Here are compared the cyclin B1⁶⁴ (PDB:7NJ0) PBP and surrounding negatively charged patch with the equivalent region in cyclin A2 (PDB:1QMZ¹⁸⁸) with their respecting facing CDK1 and 2 T-loop. The RRASK region was shown to be protected upon p21 binding. **D**, The AF2 prediction of CCCP21 is used here but only the electrostatic surface charges of the full length p21 are shown. The different segments D1, LH, D2 and D1_c are annotated where the three first ones are predicted bound to the CCC complex but not the rest of the protein.

We generated a TEM and a EFE mutant to investigate the role of the positive charges of the PBP in p21-mediated regulation and observed differences in both, HDX-MS results and kinase activity assays, indicating that this region indeed plays a key role in the regulation of the CCC complex and is not only involved in the regulation of separase. Analyzing the surface charges of p21 (Figure 33. D) reveals a negatively charged area after the D2 segment. Our studies indicate that p21's contact sites continue after the D2 segment and therefore it is plausible that this region interacts with the cyclin B1 PBP and the positively charged patch, thereby bringing the C-terminal residues of p21 closer to the CDK1 active site, promoting their phosphorylation. This model is only valid if the D1_C segment does not bind another CCC complex through its RxL motif but instead competes with first RxL motif for cyclin B1 binding. To test this, we would need to detect p21's movements relative to the CCC with or without a positively charged patch, and also monitor the formation of complexes in solution by mass photometry or NMR.

4. A potential relationship between the CDK1 T-loop and the cyclin B1 PBP

The involvement of the CDK1 T-loop in the regulation of the activity of the CCC complex is another question that needs to be further addressed in future. Our HDX-MS results show that upon binding to p21 a region close to the T-loop and to the cyclin B1 PBP undergoes structural rearrangements. In human cells, when the expression of p21 is induced during the G2/M transition, the cells remain in G2 phase and their entrance into M phase is significantly delayed or totally blocked⁸⁷. This pause correlates with an inhibition of cyclin B1/CDK1 activity and more specifically with the phosphorylation of the well-conserved Thr160 on the CDK1 T-loop by the CAK complex. Interestingly, during this phase of the cell cycle, p21 was shown not to bind in a stoichiometric way to cyclin B1/CDK1 complexes¹²⁸. This raises the question how it is possible for p21 to inhibit efficiently cyclin B1/CDK1 complexes, without binding directly to all of them. A possible explanation is that the expression of p53 and consequently p21 is not the only factor for the initiation of the G2/M arrest but required for maintaining it. This allows a smooth transition through the M phase like a buffer protein¹³⁵ and a slow activation by the CAK complex. In our research setting, CCC is expressed as an active complex, and therefore in a conformation typical for late steps in mitosis. Once Thr161 is phosphorylated, one molecule of p21 binds one CCC complex, at least *in vitro*. Indeed, when we analyze the complex on SDS-PAGE gels and by size exclusion chromatography we do

not observe anything else than quaternary CCCP21 complexes (1:1:1:1 stoichiometry). Additional data to confirm the complex stoichiometry in solution could be provided by measuring the molecular mass by mass photometry. Consistently, in our kinase assays and during the sample preparation for phospho-sites Mass-Spectrometry, we observed that when the cyclin B1 PBP was mutated, p21 seemed less stable and shortly after the start of the kinase reaction, the complex looked sub-stoichiometric on an SDS-PAGE gel. We never observed a change in elution profiles during size-exclusion chromatography between the WT and the mutants, but we only performed this comparison between unphosphorylated complexes. Mass photometry would be a well-suited technique to detect differences in complex stoichiometry, caused by T-loop displacement that we can observe in our HDX-MS results when the PBP is mutated. When the CDK1 T-loop and L12 helix are depleted (Arg106-Val164) in human breast cancer cells, neither cyclin B1 nor p21 are capable of binding to CDK1¹⁸⁹. During my thesis, I tried to assemble a CCCP21 complex with CDK1 mutated on its Thr161 (replaced by an Ala), however it was impossible to form this complex even in the presence of a phosphomimetic Thr57 to glutamate mutation that should stabilize this interaction. Moreover, the CCCP21 kinase activity changes when the PBP is mutated and the region facing the T-loop undergoes specific rearrangements. These observations imply a potential relationship between the T-loop and the positively charged patch on cyclin B1 that the loop is facing. It is known that cell cycle CDKs share high homology between family members, they all share the same mechanism involving their activation by phosphorylation of their T-loop by the CAK complex (Figure 34. A) and they are all p21 targets. However, cyclin B1 differs from other cyclins especially regarding its electrostatic surface charges of its PBP and surrounding positively charged patch. When the PBP is mutated, our HDX-MS results show a protection of the T-loop and the cyclin B1 positively charged patch highlighting a change in the structural organization in that region, potentially due to changes in the T-loop and the cyclin B1 PBP. The results of the activity kinase assays combined with the HDX-MS question if what we observe is due to p21 only, a potential relationship between CDK1 T-loop and cyclin B1 PBP only or both at the same time. We believe that a good approach to answer this question would be to test if we can repeat the same results by modifying CDK1 T-loop distance to cyclin B1 PBP without interfering with its activity and Thr160 phosphorylation (Figure 34. B). We would like to know if, by pushing the T-loop away from the cyclin B1 PBP but maintaining CDK1 activity, p21 can still inhibit the complex, and if we find the results are comparable in a TEM or EFE background.

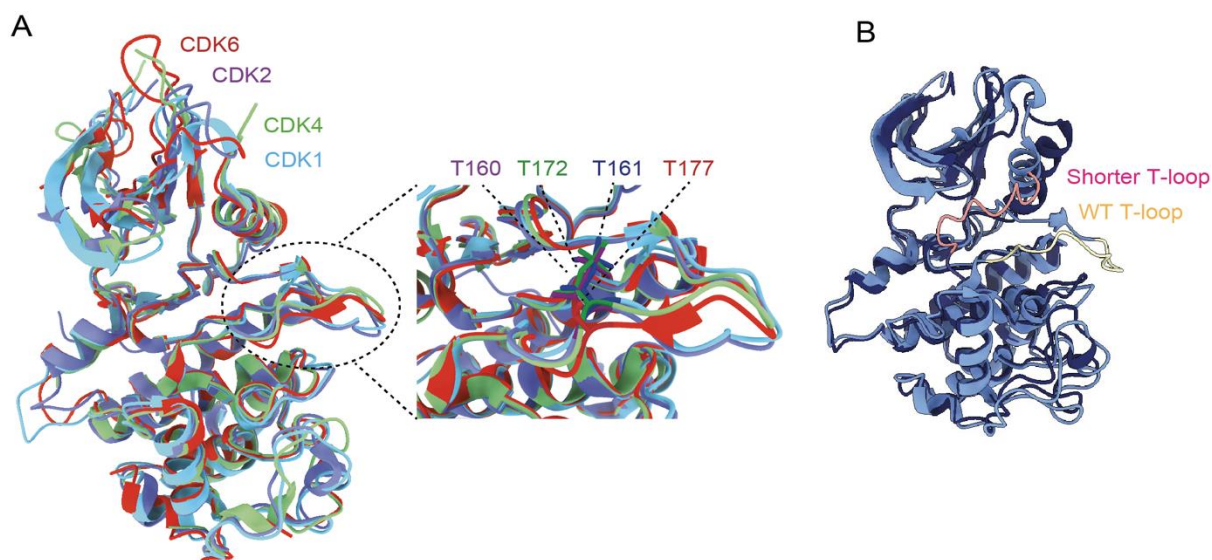


Figure 34. A possible implication of a specific interaction between CDK1 T-loop and the facing cyclin B1 positively charged patch into CCC regulation by p21

A, Superimposed structures of CDK1 (PDB:7NJ0⁶⁴), 2 (PDB:1QMZ¹⁸⁸), 4 (PDB:7SJ3¹²⁹) and 6 (PDB:1JOW¹⁹⁰) with a zoom on their activated T-loop and phosphorylated conserved threonines (T161, T160, T172 and T177). **B**, AF2 prediction of CDK1 with a designed shorter T-loop in dark blue compared to the structure of WT active CDK1 (PDB:7NJ0⁶⁴) in light blue.

5. A proposed model for p21-mediated regulation of cyclin B1/CDK1 during mitosis

Based on the results and conclusions made during this thesis project work, we would like to propose a model for p21-mediated regulation of cyclin B1/CDK1 during the different steps of mitosis (Figure 35). This model is speculative and will serve us as a foundation to continue our research to precisely understand the underlying molecular mechanisms and to shed light on the mysterious diverse roles of p21 in regulating CDK1's activity.

We propose a dual role of p21 on cyclin B1/CDK1 complexes depending on the cell cycle state (Figure 35 A and B). First, at the end of the G2 phase, CCC complexes are inhibited by two phosphorylations on Thr14 and Tyr15 executed by the kinases Wee1/Myt1. Dephosphorylation of those residues by Cdc25 happen and p21 binding to CCC is enhanced, thereby protecting the unphosphorylated region through its K domain during the G2/M transition checkpoint. The first RxL motif in the Cy1 domain bind with high affinity to cyclin B1 and the 3₁₀ helix will enter and block the CDK1 ATP-binding pocket, therefore inhibiting CDK1 kinase activity. The CDK1 T-loop remains unphosphorylated and p21 enhances the interaction between CDK1 and cyclin B1 by connecting them to each other and generating a potential interface between the CDK1 T-loop and the cyclin B1 PBP. These steps represent the first inhibitory role of p21 towards the CCC complex and happens during G2/M transition and until all CCC are bound to one p21 (Figure 35. A). We believe as well that this complex

formation is necessary to tighten the interaction between CDK1 and cyclin B1 until their disassembly during anaphase.

At the end of G2/M transition checkpoint and, therefore at the entrance of M phase, the remaining cyclin A2/CDK2 complexes phosphorylate p21 on Thr57 (part of the K domain) to promote a structural rearrangement of p21 (Figure 35. B). As a consequence of the rapprochement between cyclin B1 and CDK1 during the previous steps, cyclin B1 PBP and the positively charged patch change their structural organization in the complex. We believe that this allows p21 to bind close to the T-loop preventing the loop to fold into a fully active conformation upon phosphorylation of Thr161 by the CAK during pro-metaphase. Simultaneously, NRTKs phosphorylate p21 Tyr77 to displace the 3_{10} helix. The combination of this event with the phosphorylation of Thr161 by the CAK allows the formation of a partially active CDK1 capable of phosphorylating defined substrates with the co-factor CKS1 bound. The second RxL motif located in the Cy2 domain competes with the first motif for binding and residues at the C-terminus can be phosphorylated. Finally, Tyr151 is phosphorylated, releasing Cy2 from cyclin B1. During that period, many p21 residues are phosphorylated like Ser310 known to be important for the following APC/C-mediated degradation of p21. Slowly, cyclin B1 starts to be phosphorylated. The dissociation and degradation of p21 at the end of prometaphase allows the full activation of the CDK1 T-loop in metaphase and therefore of the complex. CDK1 is ready to play its central role in initiating anaphase, eventually resulting in chromosomes segregation. The cyclin B1 PBP is released from its close proximity to the CDK1 T-loop and can mediate other protein-protein interactions, such as binding to the protease separase. This is our proposed second role of p21 during mitosis, postulating that p21 is not only acting as an inhibitor of CDK1, but also has a potential role as a mediator and regulator of the CCC complex activity.

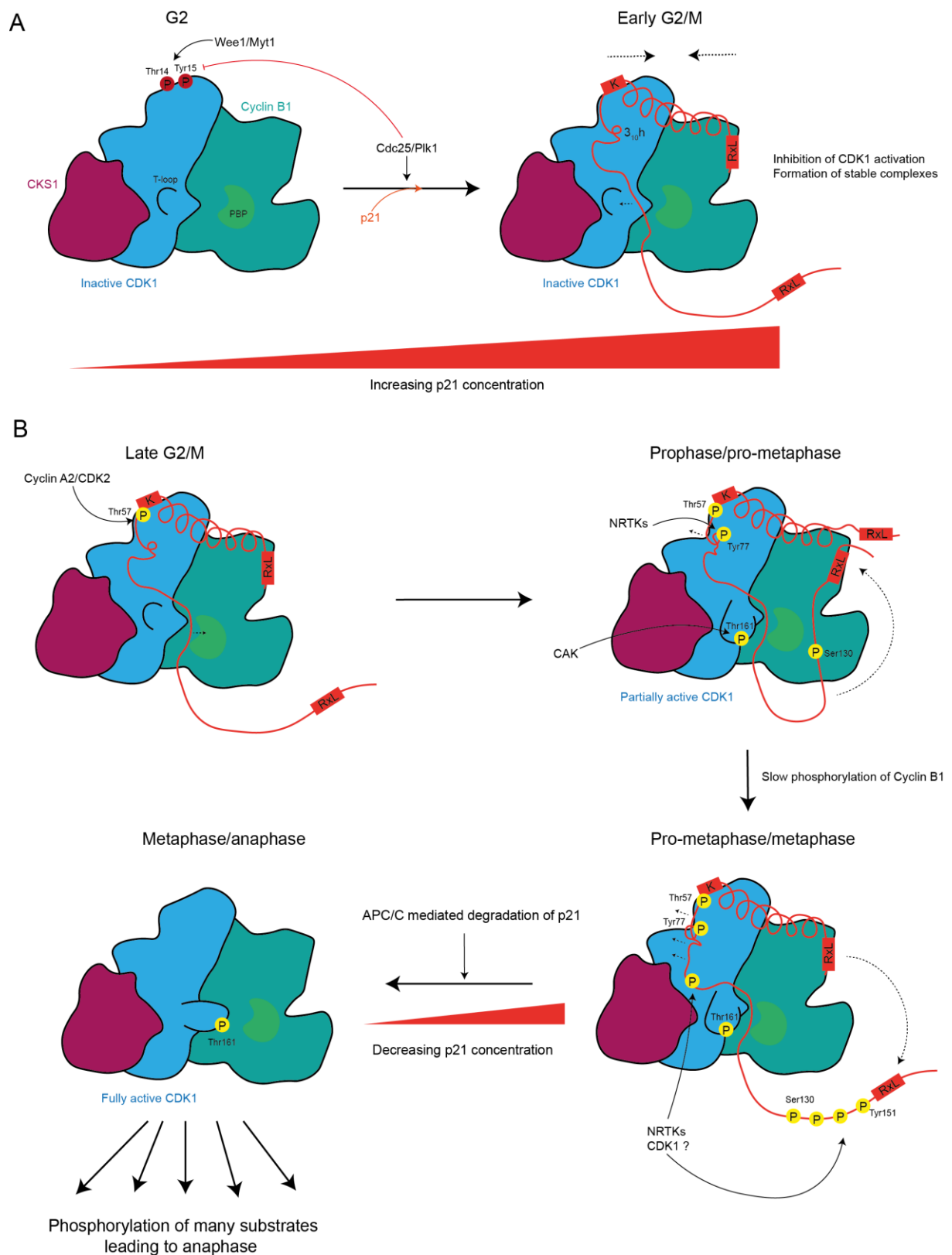


Figure 35. Proposed model based on our discoveries and existing literature

A, The first role of p21 during the preparation for M phase entrance during G2/M transition checkpoint. **B**, The second role of p21 during mitosis preparing the CCC complex for its fully activation during anaphase.

6. Perspectives and openings into the understanding of p21 degradation during mitosis

One aspect of the CCCP21 complex dynamics during mitosis that we did not study during my thesis was the degradation of p21 mediated by the APC/C complex⁸⁶. To date, there are no structural data of p21 in complex with any E3 ubiquitin ligase available. In this thesis I show that p21 does not interact with CKS1, so a mechanism similar to p27-degradation¹¹⁴ can be excluded. Existing literature suggests that p21 can bind directly to the co-activator subunit of the APC/C, CDC20, active in prometaphase, and to the APC3 subunit of the APC⁸⁶. This binding is mediated through a destruction box located in the C-terminal region of p21 (residues 86 to 89). It is interesting to notice here, that p21 remains bound to CCC and continues to inhibit CDK1 in prometaphase, raising the possibility that a large complex of CCCP21 bound to the APC/C is forming in cells. A former Ph.D. student in my research group, Anna Höfler, showed during her thesis the importance of cyclin B1 PBP for APC/C activity and degradation of cyclin B1 during anaphase rising the question if a similar mechanism is necessary for p21's degradation as well. Moreover, phosphorylations of Tyr77 and Ser130 were shown to be important for p21's degradation¹⁹¹ making us wonder if these are needed to expulse p21 from CCC or to stabilize the binding with APC/C and mediate the subsequent ubiquitination of p21. These open questions are once again establishing the potential grandiose importance of cyclin B1 PBP and its central role into the orchestration of CDK1's kinase activity through mitosis.

MATERIAL AND METHODS

1. List of reagents

REAGENT	SUPPLIER	REFERENCE
HEPES	Fisher bioreagents	7365-45-9
NACL	Fisher bioreagents	7647-14-5
GLYCEROL	Roth	3783.1
BETA-MERCAPTOETHANOL	Roth	4227.1
IMIDAZOLE	Sigma-Aldrich	102613382
DESTHIOBIOTIN	Iba	1000-2137
MALTOSE	Sigma-Aldrich	BCBX1741
NAOH	ThermoScientific	206060010
TCEP	ThermoFisher Scientific	T2556
IPTG	Fluorochem	M02726
GRAPHENE-OXIDE	Aldrich	MKCG1271
CHAPS	PanReac AppliChem / ITW reagents	A1099
CHAPSO	Apollo Scientific	BIC0121
TRITON 100X	Merck	1039039
LMNG	Anatrace	4218540
KCL	Roth	6781-2
ATP	Acros Organics	102800500
KANAMYCIN	Roth	T832.2
GENTAMYCIN	PanReac AppliChem / ITW reagents	8N014262
AMPICILLIN	Sigma	01930-25G
SF-900 III SFM SERUM FREE MEDIUM COMPLETE	Gibco	12658-027

Table 3. List of reagents used in this study

2. List of protein sequences

All protein sequences used during my thesis were retrieved from the freely accessible database UniProt (<https://www.uniprot.org/>).

PROTEIN NAME	GENE NAME	UNIPROT CODE	ORGANISM
CYCLIN-DEPENDENT KINASE INHIBITOR 1 (P21)	CDKN1A	P38936	<i>Homo sapiens</i> (human)

CYCLIN-DEPENDENT KINASE REGULATORY SUBUNIT 1 (CKS1)	CKS1B	P61024	<i>Homo sapiens</i> (human)
G2/MITOTIC-SPECIFIC CYCLIN-B1 (CYCLIN B1)	CCNB1	P14635	<i>Homo sapiens</i> (human)
CYCLIN-DEPENDENT KINASE 1 (CDK1)	CDK1	P06493	<i>Homo sapiens</i> (human)
SEPARIN (SEPARASE)	ESPL1	Q14674	<i>Homo sapiens</i> (human)
SECURIN	PTTG1	O95997	<i>Homo sapiens</i> (human)

Table 4. List of protein sequences used for protein expression in this study

These sequences were gene-optimized depending on the chosen expression system. Details can be found in sections 3.2 and 3.3.

<u>PROTEIN NAME</u>	<u>GENE NAME</u>	<u>UNIPROT CODE</u>	<u>ORGANISM</u>
CYCLIN-DEPENDENT KINASE 2 (CDK2)	CDK2	P24941	<i>Homo sapiens</i> (human)
CYCLIN-DEPENDENT KINASE 4 (CDK4)	CDK4	P11802	<i>Homo sapiens</i> (human)
CYCLIN-DEPENDENT KINASE 6 (CDK6)	CDK6	Q00534	<i>Homo sapiens</i> (human)
G2/MITOTIC-SPECIFIC CYCLIN-B2 (CYCLIN B2)	CCNB2	O95067	<i>Homo sapiens</i> (human)
G2/MITOTIC-SPECIFIC CYCLIN-B3 (CYCLIN B3)	CCNB3	Q8WWL7	<i>Homo sapiens</i> (human)
CYCLIN-A2	CCNA2	P20248	<i>Homo sapiens</i> (human)
G1/S-SPECIFIC CYCLIN-D1 (CYCLIN D1)	CCND1	Q6FI00	<i>Homo sapiens</i> (human)
G1/S-SPECIFIC CYCLIN-E1 (CYCLIN E1)	CCNE1	P24864	<i>Homo sapiens</i> (human)
CYCLIN-H	CCNH	P51956	<i>Homo sapiens</i> (human)
CYCLIN-DEPENDENT KINASE REGULATORY SUBUNIT 2 (CKS2)	CKS2	P33552	<i>Homo sapiens</i> (human)
CYCLIN-DEPENDENT KINASE INHIBITOR 1B (P27)	P27 KIP1	O43806	<i>Homo sapiens</i> (human)

Table 5. List of protein sequences used for sequence alignments in this study

All sequence alignments were performed with Clustal Omega Multiple Sequence Alignment software (EMBL, <https://www.ebi.ac.uk/Tools/msa/clustalo/>).

3. Cloning, protein expression and purification

All DNA constructs were designed and analyzed on the cloud-based platform Benchling (<https://www.benchling.com/>).

3.1. List of Polymerase Chain Reaction (PCR) protocols used during this study

STEP	TEMPERATURE (°C)	TIME
INITIAL DENATURATION	95	1 minute
DENATURATION	95	30 seconds
ANNEALING	53-57	1 minute
EXTENSION	72	1 minute/kb
ADDITIONAL EXTENSION	72	5 minutes
STORAGE	4	Indefinite

Table 6. Typical PCR protocol for gene amplification and plasmid linearization.

In orange are annotated the iterative steps that will undergo 25 cycles. The annealing temperature is always between 53 to 57°C depending on the primers (typically 15-20 bases long) melting temperature.

STEP	TEMPERATURE (°C)	TIME
INITIAL DENATURATION	95	1 minute
DENATURATION	95	30 seconds
ANNEALING	57	30 seconds
EXTENSION	72	1 minute/kb
ADDITIONAL EXTENSION	72	10 minutes
STORAGE	4	Indefinite

Table 7. Typical PCR protocol for direct-site mutagenesis.

In orange are annotated the iterative steps that will undergo 30 cycles. The annealing temperature is always between 57°C as the primers are always designed long (over 30 bases) and complementary to each other.

3.2. Bacteria

Codon-optimized genes for human p21 and CKS1 (sequences were synthesized by GeneScript) were amplified by PCR using primers with suitable overhangs and cloned into a petM41 plasmid (Addgene, #38334) using the Gibson assembly cloning protocol (NEB). The p21 T57E mutant and p21 1-35, 36-121 and 121-164 pieces were obtained by modifying directly the wild-type sequence by direct-site mutagenesis PCR using Phusion high-fidelity polymerase (NEB). DNA was isolated following the NucleoSpin Plasmid kit protocol (Macherey & Nagel) and validated through Sanger sequencing (Microsynth AG). The plasmid constructs for expression of N-terminal MBP/6his-tagged CKS1, p21, p21 T57E and p21 pieces were transformed into competent BL21 (DE3) E.Coli bacterial cells. 1L cell cultured supplemented with kanamycin (0,05mg/ml) were grown, after induction with 0.5 mM IPTG, 3 hours at 25°C before being harvested and frozen in liquid N2 for later use.

3.3. Insect cells

Codon-optimized genes for human CDK1, cyclin B1 and CKS1 (sequences synthesized by GeneScript) were amplified by PCR using primers with suitable overhangs and cloned into a gentamycin-resistant (0,007mg/ml) pF1 plasmid (Addgene, #177330) using the Gibson assembly cloning protocol (NEB). The EFE, TEM and RLR mutations on cyclin B1 were obtained by modifying directly the wild-type sequence by direct-site mutagenesis PCR using Phusion high-fidelity polymerase (NEB). DNA was isolated following the NucleoSpin Plasmid kit protocol (Macherey & Nagel) and validated through Sanger sequencing. Cyclin B1, the cyclin B1/CDK1 (CC) complex as well as the CCC complex were expressed in *Spodoptera frugiperda* pupal ovarian tissue cell line 9 (Sf9) insect cells (Invitrogen) using baculoviruses expression system. To do so, competent DH10 Multibac bacterial cells were transformed with the generated pF1 plasmid constructs. DNA was isolated following the NucleoSpin Plasmid kit protocol (Macherey & Nagel) except for the elution step which was performed by precipitation with 70% ethanol. Purified bacmids were resuspended in 100µl Sf900 medium (ThermoFisher Scientific) and mixed with 5µl transfection reagent (Fugene Promega) and added to 1.0×10^6 of adherent Sf9 cells per ml in 6-well plates in a total of 2ml. After 72h incubation at 27 °C, cells were centrifugated and the supernatant (P1) was harvested, filtered and inoculated to 3.0×10^6 Sf9 cells per ml in suspension in a total of 50ml. After 72h at 27 °C at 110 rev/min, the cells were centrifugated and the supernatant (P2) was harvested, filtered and stored at 4°C for further use. To perform large-scale protein expression, a new round of baculovirus amplification in 3.0×10^6 Sf9 cells per ml in suspension in a total of 50ml is necessary with P2 from which P3 is collected. Typically, 25 ml recombinant P3 baculovirus was used to infect 500 ml of Sf9 insect cells at a cell density of roughly 2.0×10^6 cells per ml. The cells were incubated for 72 h at 27 °C at 110 rev/min, harvested at a cell viability rate of 85–90%, flash-frozen in liquid nitrogen and stored at –80 °C. In the case of cyclin B1 and the CC complex, cyclin B1 was recombinantly expressed with a N-terminal strep-tag and CDK1 with a N-terminal MBP-tag. For CCC, cyclin B1 is expressed with a N-terminal strep-tag and CDK1 with a C-terminal 6his-tag.

4. Protein purification

All purifications were performed in our research laboratory with three purification systems: an ÄKTA Pure chromatography system (Cytiva), an ÄKTA go protein purification system (Cytiva) and an ÄKTA Fast Protein Liquid Chromatography (FPLC) system (Amersham Biosciences).

4.1. Purification of the CC and CCC complexes

Purification of all proteins and protein complexes was performed at 4 °C. Cells expressing the CC complex were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, and 10mM Imidazole) containing protease inhibitor cocktail tablets (PIC) (complete EDTA-free; Roche Diagnostics), and 5 units/ml supernuclease (Novagen), subsequently sonicated and centrifuged for 1 h at 20,000g. For the CC complex, the soluble fraction was applied onto 1 ml of HisPur Ni-NTA resin (Thermo Scientific) beads and under rotation at 4°C for 2h. The beads were washed with wash buffer (50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, 50mM Imidazole). Proteins were eluted with elution buffer (50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, 500mM Imidazole). The elution was applied onto 200uL of Strep-Tactin Superflow high-capacity resin (iba) and the beads were washed with the final buffer (20mM HEPES-NaOH pH 8, 150mM NaCl, 0,5mM TCEP). Proteins were eluted with a final buffer containing 2.5mM desthiobiotin and were loaded on a Superose 6 Increase 10/300 GL (GE Healthcare Life Sciences) size-exclusion column as a final purification step. The excess of CDK1 was pooled and concentrated as well for further experiments. For the CCC complex, the soluble fraction was slowly (0.5 ml/min flow rate) applied to two 5 ml StrepTactin Superflow Cartridge (Qiagen) after the cells were lysed as described before with a lysis buffer containing 50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, and 5% glycerol. The column was washed with lysis buffer and proteins were eluted with lysis buffer containing 2.5mM desthiobiotin and loaded onto a 5ml HisTrap HP column (GE Healthcare Life Sciences). The complex was eluted in lysis buffer containing 500mM Imidazole. Peak fractions were pooled, concentrated and loaded on a Superdex 200 Increase 10/300 GL (GE Healthcare Life Sciences) size-exclusion column as a final purification step. The same protocol was used when the cyclin B1 was mutated into TEM, EFE and RLR PBP mutants. The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

4.2. Purification of CKS1

Cells expressing CKS1 were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 7.5, 300mM NaCl, 0,5mM TCEP) containing protease inhibitor cocktail tablets (PIC) (complete EDTA-free; Roche Diagnostics), and 5 units/ml supernuclease (Novagen), and

subsequently sonicated and centrifuged for 1 h at 20,000g. The soluble fraction was loaded onto a 5ml HisTrap HP column (GE Healthcare Life Sciences) and the protein eluted following a gradient with a lysis buffer containing from 50mM Imidazole (buffer A) to 500 mM Imidazole (buffer B). The peak fractions were pooled and dialyzed overnight at 4°C against 20mM HEPES-NaOH pH 7.5, 150mM NaCl, 0,5mM TCEP supplemented with TEV protease to cut the affinity tags. The dialyzed protein was concentrated and loaded on a Superdex 75 Increase 10/300 GL (GE Healthcare Life Sciences) size-exclusion column as a final purification step. As the MBP-tag was not used during this procedure, after TEV protease removal, it was pooled, concentrated and stored to be used as a control during our pull-down experiments. The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

4.3. Purification of cyclin B1

Cells expressing the structured part of cyclin B1 protein (144-433) were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 7, 300mM NaCl, 2mM B-mercaptoethanol, and 10mM Imidazole) containing protease inhibitor cocktail tablets (PIC) (complete EDTA-free; Roche Diagnostics), and 5 units/ml supernuclease (Novagen), subsequently sonicated and centrifuged for 1 h at 20,000g. The soluble fraction was applied onto 1 ml of HisPur Ni-NTA resin (Thermo Scientific) beads and under rotation at 4°C for 2h. The beads were washed with wash buffer (50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, 50mM Imidazole). Proteins were eluted with elution buffer (50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, 500mM Imidazole). The peak fractions were pooled and dialyzed overnight at 4°C against 20mM HEPES-NaOH pH 7.5, 150mM NaCl, 0,5mM TCEP supplemented with TEV protease to cut the affinity tags. The dialyzed protein was concentrated and loaded on a Superdex 75 Increase 10/300 GL (GE Healthcare Life Sciences) size-exclusion column as a final purification step. The same protocol was used when the cyclin B1 was mutated into TEM, EFE and RLR PBP mutants. The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

4.4. Purification of p21 1-35, 36-120 and 121-164

Cells expressing CKS1 were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 7.5, 300mM NaCl, 0,5mM TCEP) containing protease inhibitor cocktail tablets (PIC) (complete

EDTA-free; Roche Diagnostics), and 5 units/ml supernuclease (Novagen), and subsequently sonicated and centrifuged for 1 h at 20,000g. The soluble fraction was loaded onto a 5ml HisTrap HP column (GE Healthcare Life Sciences) and the protein eluted following a gradient with a lysis buffer containing from 50mM Imidazole (buffer A) to 500 mM Imidazole (buffer B). The peak fractions were pooled and concentrated and loaded on a Superdex 200 Increase 10/300 GL (GE Healthcare Life Sciences) size-exclusion column as a final purification step. The N-terminal MBP remained untouched as it was used to perform the pull-down assays. The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

4.5. Purification of separase

The human separase C2029D attached to Securin lacking the first 141 residues used during my thesis has been cloned, expressed, purified and provided by my colleague Dr. Jun Yu. All the details of his protocols can be found in his publication⁶⁴.

4.6. Purification of p21 and p21 T57E and CCCP21 complex assembly

Purification of full-length p21 constructs was always associated to CCCP21 complex assembly during this study. As a first step to CCCP21 complex assembly, the CCC complex was purified as described in the section 4.1, but the purification was paused after the elution on the StrepTactin Superflow Cartridge (Qiagen).

As p21 is a very unstable and unstructured protein, it needs to rapidly bind its partners to avoid protein aggregation and precipitation. Cells expressing p21 and p21 T57E were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 7.5, 300mM NaCl, 0.5 mM TCEP and 10% glycerol) containing protease inhibitor cocktail tablets (PIC) (complete EDTA-free; Roche Diagnostics) and 5 units/ml supernuclease (Novagen) and subsequently lysed with a Avestin Emulsiflex C3 high pressure homogenizer (ATA scientific instruments) at 4°C and centrifuged for 1 h at 20,000g. The soluble fraction was loaded onto a 5mL MBPTrap HP column (GE Healthcare Life Sciences). After extensive column wash with a washing buffer (50mM HEPES-NaOH pH 7.5, 300mM NaCl, 0.5 mM TCEP and 5% glycerol), the protein was eluted with a lysis buffer supplemented with 10mM maltose. Immediately after, the pending CCC complex was mixed with p21 and p21 T57E under rotation for an hour at 4°C. The mix was loaded onto a 5mL StrepTactin Superflow Cartridge (Qiagen), and the formed complex was eluted with lysis buffer containing 2.5mM desthiobiotin and supplemented with

TEV protease to cut the affinity tags overnight at 4°C without any rotation to avoid protein aggregation. The complex was after concentrated and loaded on a Superdex 200 Increase 10/300 GL (GE Healthcare Life Sciences) size-exclusion column as a final purification step with a gel filtration buffer (50mM HEPES-NaOH pH 8, 150mM NaCl and 0.5 mM TCEP). The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue. This complex was used for further EM, MS and activity analysis.

The same was performed with the TEM and EFE CCC mutants. All analytical size exclusion chromatography analysis were performed using a Superose 6 5/150 GL column (Cytiva). To test the effect of the usage of different salts in our buffer condition, this final step was performed using a buffer containing 150mM KCl instead of NaCl.

5. Pull-down assays

5.1. CC and separase

Purified CC (WT, EFE, TEM and RLR) protein complexes (N-terminal MBP/6his-tag) were mixed with securin (lacking the first 141 residues) and separase (C2029D) and incubated with 20µl amylose beads (Iba) for 1h at 4°C in 20mM HEPES-NaOH pH 7.5, 150mM NaCl, 0,5mM TCEP. The beads were extensively washed with wash buffer (20mM HEPES-NaOH pH 7.5, 300mM NaCl, 0,5mM TCEP) before the protein being eluted using the same buffer supplemented with 10mM maltose, analyzed on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

5.2. CCC and CDK1 and p21 1-35, 36-120 and 121-164

Purified p21 protein constructs (N-terminal MBP-His tag) were mixed with the different CCC constructs (WT, TEM and EFE) and incubated with 20µl amylose beads (Iba) for 1h at 4°C in 20mM HEPES-NaOH pH 7.5, 150mM NaCl, 0,5mM TCEP. The beads were extensively washed with was buffer (20mM HEPES-NaOH pH 7.5, 300mM NaCl, 0,5mM TCEP) before the protein being eluted using the same buffer supplemented with 10mM maltose, analyzed on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

6. Radioactive kinase activity assays

Each reaction (total volume, 10 μ l) contained 1,5 μ g recombinant CCC protein complexes (with or without p21) or 350 ng CDK1–cyclin B recombinant human protein (Thermo Fisher PV3292), 0.5 mM or 5mM cold ATP and 5 or 50 μ Ci γ^{32} P-ATP in kinase buffer (50 mM HEPES pH 7.5, 10 mM $MgCl_2$, 1 mM EGTA, 0.01% Brij-35) supplemented with PhosSTOP (Sigma-Aldrich) and Complete EDTA-free proteases inhibitors (Sigma-Aldrich). The samples were incubated at 30°C for every time point studied. The reactions were stopped by adding 3 $\times$ Laemmli sample buffer and boiled at 95°C for 5 min. The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad). The gel was stained with Coomassie Blue, dried and exposed to a phosphorimager screen (GE Healthcare). The results were analysed with a Typhoon FLA 9500 (GE Healthcare).

7. AlphaFold predictions

The structure predictions made with AlphaFold2 were performed with Google Colab and default settings¹⁹². The obtained five predictions were compared and the one with highest accuracy value was analyzed using UCSF ChimeraX¹⁹³.

8. Electron microscopy

8.1. List of EM grids used during this study

EM GRID TYPE	SUPPLIER	CATALOGUE NUMBER
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	Quantifoil micro tools via Electron Microscopy Sciences (EMS)	Q350AR1.3
R 2/2 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	Quantifoil micro tools via EMS	Q350AR2
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH COPPER GRID	Quantifoil micro tools via EMS	Q350CR1.3
R 2/2 HOLEY CARBON FILM ON A 300-MESH COPPER GRID	Quantifoil micro tools via EMS	Q350CR2
R 0.6/1 HOLEY CARBON FILM ON A 300-MESH COPPER GRID	Quantifoil micro tools via EMS	Q350CR-06
R 1.2/1.3 HOLEY GOLD FILM ON A 300-MESH GOLD GRID (ULTRAUFOIL)	Quantifoil micro tools via EMS	Q350AR13A

R 2/2 HOLEY CARBON FILM ON A 300-MESH COPPER GRID WITH, UTC 2NM CARBON FILM	Quantifoil micro tools via EMS	Q350CR1.3-2nm
R 1.2/1.3 C-FLAT™ HOLEY CARBON FILM ON A 300-MESH COPPER GRID	EMS	C313-50
R 2/2 C-FLAT™ HOLEY CARBON FILM ON A 300-MESH COPPER GRID	EMS	C312-50
HEXAUFOL, R 0.3/0.6 HOLEY GOLD FILM ON A 600-MESH GOLD GRID	Christopher J. Russo, LMB	
LACEY CARBON FILM ON A 200-MESH COPPER GRID	Agar scientific via EMS	LC200-Cu
NEGATIVE STAIN CARBON FILM ON A 400-MESH COPPER GRID	EMS	CF400-Cu

Table 8. List of EM grid types and manufacturers used during this study

Before plunge-freezing procedures, EM grids are glow discharged using a GloQube Plus Glow Discharge system (Quorum, EMS). Typically, EM grids are applied a negative charge by glow-discharging them at 40mA during 20 to 60 seconds for cryoEM and 40 seconds for negative staining. Each of those grids have been used to analyze CCCP21 particles and the details of the screened conditions can be found in Table 2.

8.2. Graphene-oxide grid preparation

Graphene-oxide support layers were prepared on Quantifoil R 1.2/1.3 Holey carbon film on a 300-mesh gold grid (Quantifoil micro tools) following a well-established protocol¹⁶¹. Typically, the graphene-oxide solution was diluted 10x before 5µl was absorbed on a glow-discharged grid for 30 seconds. The grid was washed with three drops of water before being dried for few minutes. Directly after the grids are ready and the sample can be applied and blotted depending on the tested conditions.

8.3. Negative staining

Grid preparation and acquisition. Typically, a 5uL drop of 50uM CCCP21 T57E complex (tagged proteins) was absorbed on a glow-discharged carbon-coated copper grid (Electron Microscopy Sciences) (Table 8.). The grid was washed with two consecutive drops of gel filtration buffer (50mM HEPES-NaOH pH 8, 150mM NaCl and 0.5 mM TCEP) and stained with two consecutive drops of 1% uranyl-acetate. After a final stain solution absorption step with Whatman paper, the grid was air-dried for few minutes before being imaged at room

temperature on a Talos L120C G2 (S) Transmission Electron Microscope (TEM) (Thermo Fisher Scientific). The microscope is equipped with a LaB6 filament and is operating at a 120keV accelerating voltage. Images of the complex were recorded at a x120.000 magnification (pixel size 1.998Å) at a defocus value at -1.5 uM.

Image processing. Two datasets were collected at two protein concentrations, 0.025mg/ml and 0.05mg/ml. Respectively, 230 and 130 micrographs were collected. All data processing steps were performed using CryoSPARC v 3.3.1. The raw micrographs were imported, and particles were directly picked using the blob picker tool. After extracting the particles, two rounds of 2D classifications were performed, and the selected particles were used as a template to improve automated particle picking. Two more rounds of 2D classification were performed before analyzing the 2D classes.

8.4. CryoEM of CCCP21 with tags

Grid preparation and acquisition. A 5uL drop of 50uM CCCP21 T57E complex (tagged proteins) was applied onto glow-discharged grids (Table 8) and blotted following the details in the results section (Table 1.) before being plunged-frozen in liquid ethane using an EM GP2 automatic plunge freezer (Leica). The grids were stored in liquid nitrogen until further use. A total of 432 movies were recorded on a Talos Arctica cryo-TEM (Thermo Fisher Scientific) equipped with a Falcon 3EC Directed Electron Detector (Thermo Fisher Scientific) operating at 200 keV at a magnification of 150.000x resulting in a super-resolution pixel size of 0.998Å. A total of 40e-/Å² was distributed over 44 frames. The movies were collected using the automated data collection software EPU (Thermo Fisher Scientific) with a set of defocus range from -1uM to -2.5uM.

Image processing. Raw movies from this first dataset were aligned and dose-weighted using MorionCor2 implemented in RELION. The resulting micrographs were imported into CryoSPARC v 3.3.1 and the CTF parameters were estimated using Gctf. Particles were picked with Topaz extract with a Topaz pre-trained model and extracted. With several rounds of 2D classifications, the best 2D classes were picked and analyzed.

8.5. CryoEM of CCCP21 without tags

Grid preparation and acquisition. A 5uL drop of 50uM CCCP21 T57E complex was applied onto a glow-discharged 300 square mesh Quantifoil R 1.2/1.3 Holey carbon film on a 300-mesh gold grid (Quantifoil Micro Tools) and immediately back blotted for 1.5s with 1mm

additional movement (90% humidity at 14°C) before being plunged-frozen in liquid ethane using an EM GP2 automatic plunge freezer (Leica). The grids were stored in liquid nitrogen until further use. A total of 500 movies were recorded on a Falcon 3EC Directed Electron Detector (Thermo Fisher Scientific) on a Talos Arctica cryo-TEM (Thermo Fisher Scientific) operating at 200 keV at a magnification of 150.000x resulting in a super-resolution pixel size of 0.998Å. A total of 40e-/Å² was distributed over 44 frames. The movies were collected using the automated data collection software EPU (Thermo Fisher Scientific) with a set of defocus range of -1uM to -2.5uM.

Image processing. Raw movies from this first dataset were aligned and dose-weighted using MorionCor2 implemented in RELION. The resulting micrographs were imported into CryoSPARC v 3.3.1 and the CTF parameters were estimated using Gctf. Particles were picked with Topaz extract with a Topaz pre-trained model before being extracted. With several rounds of 2D classifications, the best classes were picked and displayed.

8.6. CryoEM of the first big CCCP21 dataset

Grid preparation and acquisition. For large data collection, the same blotting and plunging conditions were used as in section 8.5 and stored in liquid nitrogen until further use. The previously already collected 500 micrographs dataset was pooled to the four more collected here for a total of 5639 recorded movies. Data acquisition was performed in the exact same way as in section 8.5.

Image processing. Raw movies from this first dataset were aligned and dose-weighted using MorionCor2 implanted in. RELION. The micrographs were imported into CryoSPARC v 3.3.1 and the CTF parameters were estimated using Gctf. Particles were picked with Topaz extract with a Topaz pre-trained model before being extracted. With several rounds of 2D classifications, the best classes were used to train Topaz for automated particle picking until satisfaction. The resulting trained model was used for all further cryo data processing. After an initial round of 3D refinement on CryoSPARC v 3.3.1, the result was transferred back on RELION 3.3.1 for Bayesian polishing and CTF refinement before being transferred back to CryoSPARC v 3.3.1. Several rounds of 3D refinement (heterogenous refinement and non-uniform refinement) as well as 3D classifications were necessary to reach, after 3D sharpening, a final 7.4Å resolution EM density volume.

8.7. CryoEM of the second big dataset

Grid preparation and acquisition. For this second large data collection, the same blotting and plunging conditions were used as in sections 8.5 and 8.6 and stored in liquid nitrogen until further use. The grids were screened on a 200kv Talos Arctica cryo-TEM (Thermo Fisher Scientific) before performing the data collection on a 300kv Titan Krios cryo-TEM (Thermo Fisher Scientific) equipped with a Selectris energy filter (Thermo Fisher Scientific). A total of 16.598 movies were recorded on a Falcon 4i Directed Electron Detector (Thermo Fisher Scientific) at a magnification of 150.000x resulting in a super-resolution pixel size of 0.726Å. A total of 50e⁻/Å² was distributed over 44 frames. The movies were collected using the automated data collection software EPU (Thermo Fisher Scientific) with a set of defocus range of -0,5uM to -1,6uM.

Image processing. The dataset was divided into three parts (8600, 4085 and 3913 micrographs each) for processing simplification. Raw movies from this dataset were aligned and dose-weighted using MorionCor2 implanted in RELION. The micrographs were imported into CryoSPARC v 3.3.1 and the CTF parameters were estimated using Gctf. 1000 particles were manually picked and used as a template for picking other particles. Several rounds of 2D classifications were necessary to obtain visible CCCP21 particles and these particles were used to reconstruct a 3D volume at a limited final resolution.

8.8. Cryo-EM of CCCP21 with the Center for Electron Microscopy

Grid preparation and acquisition. For each grid prepared, 4 ul of purified CCCP21T57E complex (0.1 to 0,3 mg/ml) were applied onto a glow-discharged set of grids described in the results and immediately blotted following the conditions in Table 9 and plunged-frozen in liquid ethane using an EM GP2 automatic plunge freezer (Leica). The grids were screened on a 200kv Talos Arctica cryo-TEM (Thermo Fisher Scientific) and stored then sent in liquid nitrogen to the Center for Electron Microscopy of Thermo Fisher Scientific in Eindhoven. The data collections were performed on a 300kv Titan Krios cryo-TEM (Thermo Fisher Scientific) equipped with a Selectris energy filter (Thermo Fisher Scientific). Movies were recorded on a Falcon 4 Directed Electron Detector (Thermo Fisher Scientific) at a magnification of 180.000x resulting in a super-resolution pixel size of 0.59Å. A total of 70e⁻/Å² was distributed over 50 frames. The movies were collected using the automated data collection software EPU (Thermo Fisher Scientific) with a set of defocus range of -0,8uM to -2uM.

<u>EM GRID TYPE</u>	<u>PROTEIN CONCENTRATION</u>	<u>BLOTTING CONDITIONS</u>	<u>NUMBER OF MOVIES COLLECTED</u>
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,3mg/ml	1,5s blotting time 1mm extra movement, 14°C, 100% humidity	1000
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	1,5s blotting time 1mm extra movement, 14°C, 100% humidity	825
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	2s blotting time 1mm extra movement, 14°C, 100% humidity	1000
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,3mg/ml	2s blotting time 1mm extra movement, 14°C, 100% humidity	960
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	1,5s blotting time 1mm extra movement, 14°C, 100% humidity	825
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	1s blotting time 1mm extra movement, 14°C, 100% humidity	1000
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	1s blotting time 1mm extra movement, 14°C, 100% humidity	900
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	1s blotting time 1mm extra movement, 14°C, 100% humidity	815
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	2s blotting time 1mm extra movement, 14°C, 100% humidity	1165
R 1.2/1.3 HOLEY CARBON FILM ON A 300-MESH GOLD GRID	0,2mg/ml	1,5s blotting time 1mm extra movement, 14°C, 100% humidity	1000

Table 9. Details of the EM grid types, blotting and plunging conditions used to prepare the grids imaged by the Center for Electron Microscopy of Thermo Fisher Scientific in Eindhoven

The EM grids used here are described in Table 6.

Image processing. Raw movies were imported on CryoSPARC v 4.4.0 before being aligned and dose-weighted using the implanted Patch Motion Correction tool and the CTF estimated with the implanted Patch CTF tool. The micrographs were manually curated to exclude the

empty holes, the carbon edges and the damaged parts of the grids. The particles were picked using the implanted Blob Picker tool and several runs of 2D classifications were performed to analyze the quality of the particles.

8.9. CryoEM to improve the quality of the particles

Detergents. For each grid prepared, 4 μ l of purified CCCP21 complex (0.1 mg/ml) was applied onto a glow-discharged Quantifoil R 1.2/1.3 Holey carbon film on a 300-mesh gold grid (Quantifoil Micro Tools) blotted similarly as in 8.4. CHAPSO and CHAPS in different ranges of diluted 8mM CMC (0,1 to 1X) was added directly to the drop on the grid a few seconds before the blotting step. Triton X-100 (0,1x CMC of 0,24mM), Tween 20 0,01 % (CMC 0,06mM) and LMNG (0,01x CMC of 0,01mM) were added to the protein solution before the pipetting step on the grid prior to the blotting step. The grids were screened on a 200kv Talos Arctica cryo-TEM (Thermo Fisher Scientific) equipped with a Falcon 4i Directed Electron Detector (Thermo Fisher Scientific). A total of 535 movies were recorded at a magnification of 150.000x resulting in a super-resolution pixel size of 0.97Å on a grid prepared with CHAPSO (0.1x CMC).

Salts. For each grid prepared, 4 μ l of purified CCCP21 complex (0.1 to 0.2 mg/ml) previously purified in 150mM KCl was applied onto glow-discharged Quantifoil R 1.2/1.3 Holey carbon film on a 300-mesh gold grid (Quantifoil Micro Tools) and immediately back blotted for 1.5s with 1mm additional movement (90% humidity at 14°C) before being plunged-frozen in liquid ethane using an EM GP2 automatic plunge freezer (Leica). The grids were screened and analyzed similarly as for the detergents section. The same is applicable for the data collection of 1816 movies.

EM grid type. Different EM grid types, blotting and plunging conditions were extensively screened (Table 2). For each condition, 4 μ l of CCCP21 complex was applied onto a glow-discharged EM grid, blotted and plunged-frozen in liquid ethane using an EM GP2 automatic plunge freezer (Leica). The grids were screened and analyzed as previously described. A total of 589 and 573 were recorded on Quantifoil R1.2/1.3 UltrAuFoil and Quantifoil R 1.2/1.3 Holey carbon film on a 300-mesh copper grids respectively (Quantifoil Micro Tools).

Image processing. Raw movies were imported on CryoSPARC v 3.3.1 and later on v 4.4.0 and the image processing was performed in the exact same way as for in the previous section 8.5.

9. HDX-MS

These experiments have been conducted by Dr. Oscar Vadas.

HDX-MS experiments were performed at the UniGe Protein Platform (University of Geneva, Switzerland) following a well-established protocol with minimal modifications¹⁹⁴. Details of reaction conditions and all data are presented in HDX tables. Two conditions were tested and compared: 1) CCC complexes and i2) CCCP21 complexes. HDX reactions were done in 50 μ l volumes with a final protein concentration of 2.2 μ M. Briefly, picomoles of protein (Table 10-14) were pre-incubated in 10 μ l final volume for 5 min at 0 and 4°C before initiating deuteration reaction.

Deuterium exchange reaction was initiated by adding 40 μ l of D₂O exchange buffer (20 mM Tris pH 8, 150 mM NaCl, 0.5 mM TCEP) to the protein mixture. Reactions were carried-out on ice for 3 incubation times (3 s, 30 s, 300 s) and terminated by the sequential addition of 20 μ l of ice-cold quench buffer 1 (4M Gdn-HCl, 1M NaCl, 100mM NaH₂PO₄ pH 2.4, 1 % formic acid (FA), 100mMTCEP). Samples were immediately frozen in liquid nitrogen and stored at -80 °C for up to two weeks. All experiments were done in triplicates.

To quantify deuterium uptake into the protein, samples were thawed and injected in a UPLC system immersed in ice with 0.1 % FA as liquid phase. The protein was digested via two immobilized pepsin columns (Thermo #23131), and peptides were collected onto a VanGuard precolumn trap (Waters). The trap was subsequently eluted, and peptides separated with a C18, 300Å, 1.7 μ m particle size Fortis Bio 100 x 2.1mm column over a gradient of 8 – 30 % buffer C over 20 min at 150 μ l/min (Buffer B: 0.1% formic acid; buffer C: 100% acetonitrile). Mass spectra were acquired on an Orbitrap Velos Pro (Thermo), for ions from 400 to 2200 m/z using an electrospray ionization source operated at 300 °C, 5 kV of ion spray voltage. Peptides were identified by data-dependent acquisition of a non-deuterated sample after MS/MS and data were analyzed by Mascot. All peptides analyzed are shown in HDX Tables. Deuterium incorporation levels were quantified using HD examiner software (Sierra Analytics), and quality of every peptide was checked manually. Results are presented as percentage of maximal deuteration compared to theoretical maximal deuteration. Changes in deuteration level between two states were considered significant differences if present for a given peptide both at percentage level and number of deuterons (see HDX setup Tables 10-14 below).

<u>DATASET</u>	<u>STATE 1</u>	<u>STATE 2</u>
DESCRIPTION	CKS1_CCC	CKS1_CCCP21
REACTION VOLUME	50µl	50µl
% D ₂ O IN THE REACTION	76%	76%
D ₂ O INCUBATION TIME (SECONDS)	3,30,300, , ,	3,30,300, , ,
TEMPERATURE	0°C	0°C
CONTROL SAMPLE	Non-deuterated (ND) CKS1_CCC	Non-deuterated (ND) CKS1_CCC
QUENCH BUFFER	4 M Gdn-HCl, 1 M NaCl, 100 mM NaH ₂ PO ₄ pH 2.4, 1 % formic acid (FA), 100 mM TCEP	4 M Gdn-HCl, 1 M NaCl, 100 mM NaH ₂ PO ₄ pH 2.4, 1 % formic acid (FA), 100 mM TCEP
QUENCH BUFFER VOLUME	20µl	20µl
NUMBER OF PEPTIDES ANALYZED	31	31
SEQUENCE COVERAGE	92,4%	92,4%
REPLICATES	1,1,1,0,0,0	1,1,1,0,0,0
CRITERIA FOR HDX RATE DIFFERENCE	Difference of HDX level at a given timepoint >15% and >1.5Da and p value of student t-test is <0.01	Difference of HDX level at a given timepoint >15% and >1.5Da and p value of student t-test is <0.01
PROTEIN AMOUNT (PMOL)	109	109

Table 10. HDX setup table for CKS1 measurements in CCC and CCCP21

<u>DATASET</u>	<u>STATE 1</u>	<u>STATE 2</u>
DESCRIPTION	CDK1_CCC	CDK1_CCCP21
REACTION VOLUME	50µl	50µl
% D ₂ O IN THE REACTION	76%	76%
D ₂ O INCUBATION TIME (SECONDS)	3,30,300, , ,	3,30,300, , ,
TEMPERATURE	0°C	0°C
CONTROL SAMPLE	Non-deuterated (ND) CDK1_CCC	Non-deuterated (ND) CDK1_CCC

QUENCH BUFFER	4 M Gdn-HCl, 1 M NaCl, 100 mM NaH ₂ PO ₄ pH 2.4, 1 % formic acid (FA), 100 mM TCEP	4 M Gdn-HCl, 1 M NaCl, 100 mM NaH ₂ PO ₄ pH 2.4, 1 % formic acid (FA), 100 mM TCEP
QUENCH BUFFER VOLUME	20µl	20µl
NUMBER OF PEPTIDES ANALYZED	111	111
SEQUENCE COVERAGE	91,4%	91,4%
REPLICATES	1,1,1,0,0,0	1,1,1,0,0,0
CRITERIA FOR HDX RATE DIFFERENCE	Difference of HDX level at a given timepoint >15% and >1.5Da and p value of student t-test is <0.01	Difference of HDX level at a given timepoint >15% and >1.5Da and p value of student t-test is <0.01
PROTEIN AMOUNT (PMOL)	109	109

Table 11. HDX setup table for CDK1 measurements in CCC and CCCP21

<u>DATASET</u>	<u>STATE 1</u>	<u>STATE 2</u>
DESCRIPTION	Cyclin B1_CCC	Cyclin B1_CCCP21
REACTION VOLUME	50µl	50µl
% D ₂ O IN THE REACTION	76%	76%
D ₂ O INCUBATION TIME (SECONDS)	3,30,300, , ,	3,30,300, , ,
TEMPERATURE	0°C	0°C
CONTROL SAMPLE	Non-deuterated (ND) Cyclin B1_CCC	Non-deuterated (ND) Cyclin B1_CCC
QUENCH BUFFER	4 M Gdn-HCl, 1 M NaCl, 100 mM NaH ₂ PO ₄ pH 2.4, 1 % formic acid (FA), 100 mM TCEP	4 M Gdn-HCl, 1 M NaCl, 100 mM NaH ₂ PO ₄ pH 2.4, 1 % formic acid (FA), 100 mM TCEP
QUENCH BUFFER VOLUME	20µl	20µl
NUMBER OF PEPTIDES ANALYZED	121	121
SEQUENCE COVERAGE	78,2%	78,2%
REPLICATES	1,1,1,0,0,0	1,1,1,0,0,0
CRITERIA FOR HDX RATE DIFFERENCE	Difference of HDX level at a given timepoint >15% and >1.5Da and p value of student t-test is <0.01	Difference of HDX level at a given timepoint >15% and >1.5Da and p value of student t-test is <0.01
PROTEIN AMOUNT (PMOL)	109	109

Table 12. HDX setup table for cyclin B1 measurements in CCC and CCCP21

<u>DATASET</u>	<u>STATE 1</u>	<u>STATE 2</u>	<u>STATE 3</u>	<u>STATE 4</u>	<u>STATE 5</u>	<u>STATE 6</u>
DESCRIPTION	CDK1_CCC	CDK1_CCCP21	CDK1_CCC(EF E)P21	CDK1_CCC(TE M)P21	CDK1_CCC(EF E)	CDK1_CCC(TE M)
CONTROL SAMPLE	Non-deuterated (ND)	Non-deuterated (ND)	Non-deuterated (ND)	Non-deuterated (ND)	Non-deuterated (ND)	Non-deuterated (ND)
	CDK1_CCC	CDK1_CCC	CDK1_CCC	CDK1_CCC	CDK1_CCC	CDK1_CCC
D2O INCUBATION TIME (SECONDS)	3,30,300, , ,	3,30,300, , ,	3,30,300, , ,	3,30,300, , ,	3,30,300, , ,	3,30,300, , ,
REPLICATES	3,3,3,0,0,0	2,3,3,0,0,0	1,1,1,0,0,0	1,1,1,0,0,0	0,1,1,0,0,0	0,1,1,0,0,0
SEQUENCE COVERAGE	91,4%	91,4%	91,4%	91,4%	91,4%	91,4%
NUMBER OF PEPTIDES ANALYZED	97	97	97	97	97	97
REACTION VOLUME	50µl	50µl	50µl	50µl	50µl	50µl
% D2O IN THE REACTION	75,9%	75,9%	75,9%	75,9%	75,9%	75,9%
TEMPERATURE	4°C	4°C	4°C	4°C	4°C	4°C
PROTEIN AMOUNT (PMOL)	109	91	91	73	152	152
PROTEIN CONCENTRATION IN REACTION (µM)	2.2	1.8	1.8	1.5	1.8	1.8
PROTEIN CONCENTRATED MIX VOLUME (µM)	10	10	10	10	10	10
QUENCH VOLUME (µL)	20	20	20	20	20	20
DEUTERATED BUFFER	20mM Tris pH8, 150mM NaCl, 0,5mM TCEP	20mM Tris pH8, 150mM NaCl, 0,5mM TCEP	20mM Tris pH8, 150mM NaCl, 0,5mM TCEP	20mM Tris pH8, 150mM NaCl, 0,5mM TCEP	20mM Tris pH8, 150mM NaCl, 0,5mM TCEP	20mM Tris pH8, 150mM NaCl, 0,5mM TCEP
QUENCH BUFFER	Same as Table 12	Same as Table 12	Same as Table 12	Same as Table 12	Same as Table 12	Same as Table 12
CRITERIA FOR HDX RATE DIFFERENCE	Same as Table 12	Same as Table 12	Same as Table 12	Same as Table 12	Same as Table 12	Same as Table 12

Table 13. HDX setup table for CDK1 measurements in CCC and CCCP21 TEM and EFE mutants compared to WT

DATASET	STATE 1	STATE 2	STATE 3
DESCRIPTION	Cyclin B1_CCC	Cyclin B1_CCC(EFE)	Cyclin B1_CCC(TEM)
CONTROL SAMPLE	Non-deuterated (ND) cyclin B1_CCC	Non-deuterated (ND) cyclin B1_CCC	Non-deuterated (ND) cyclin B1_CCC
D2O INCUBATION TIME (SECONDS)	30, 300	30, 300	30, 300
REPLICATES	1,1	1,1	1,1
SEQUENCE COVERAGE	73,8%	91%	91%
NUMBER OF PEPTIDES ANALYZED	114	114	114
REACTION VOLUME	50µl	50µl	50µl
% D2O IN THE REACTION	75,9%	75,9%	75,9%
TEMPERATURE	4°C	4°C	4°C
PROTEIN AMOUNT (PMOL)	109	152	152
PROTEIN CONCENTRATION IN REACTION (µM)	2.2	1.8	1.8
PROTEIN CONCENTRATED MIX VOLUME (µM)	10	10	10
QUENCH VOLUME (µL)	20	20	20
DEUTERATED BUFFER	Same as Table 13	Same as Table 13	Same as Table 13
QUENCH BUFFER	Same as Table 12	Same as Table 12	Same as Table 12
CRITERIA FOR HDX RATE DIFFERENCE	Same as Table 12	Same as Table 12	Same as Table 12

Table 14. HDX setup table for cyclin B1 measurements in CCC TEM and EFE mutants compared to WT

10. LC-MS/MS

Prior to the MS measurement, 20 µg of WT CCCP21 were mixed with 5 mM cold ATP in kinase buffer (50 mM HEPES pH 7.5, 10 mM MgCl₂, 1 mM EGTA, 0.01% Brij-35) supplemented with PhosSTOP (Sigma-Aldrich) and Complete EDTA-free proteases inhibitors (Sigma-Aldrich). The samples were incubated at 30°C for every time point studied. The reactions were stopped by adding 3× Laemmli sample buffer and boiled at 95°C for 5 min. The samples were then resolved on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad). The gel was stained with Coomassie Blue. P21 and cyclin B1 protein bands were cut at 0 and 30 minutes after kinase reaction and sent in ddH₂O to the Proteomics Core Facility in EMBL, Heidelberg.

In-gel digestion. Bands corresponding to the protein of interest were cut from the gel and in-gel digestion with trypsin (Promega) was performed essentially as previously described¹⁹⁵. Peptide extraction was done by sonication for 15 minutes, followed by centrifugation and supernatant collection. A solution of 50:50 water: acetonitrile, 1 % formic acid was used for a second extraction. The supernatants of both extractions were pooled and dried in a vacuum concentrator. Peptides were dissolved in 10 µL of reconstitution buffer (96:4 water: acetonitrile, 1% formic acid) and analyzed by LC-MS/MS.

Mass spectrometry. For LC-MS/MS measurement, an Orbitrap Fusion Lumos instrument (Thermo) coupled to an UltiMate 3000 RSLC nano LC system (Dionex) was used. Peptides were concentrated on a trapping cartridge (µ-Precolumn C18 PepMap 100, 5 µm, 300 µm i.d. x 5 mm, 100 Å) with a constant flow of 0.05% trifluoroacetic acid in water at 30 µL/min for 6 minutes. Subsequently, peptides were eluted and separated on the analytical column (nanoEase™ M/Z HSS T3 column 75 µm x 250 mm C18, 1.8 µm, 100 Å, Waters) using a gradient composed of Solvent A (3% DMSO, 0.1% formic acid in water) and solvent B (3% DMSO, 0.1% formic acid in acetonitrile) with a constant flow of 0.3 µL/min. The percentage of solvent B was stepwise increased from 2% to 8% in 6 min, to 25% in 6 min, to 40% in 3 min, to 85% in 8 min and back to 2% in 6 min. The outlet of the analytical column was coupled directly to the mass spectrometer using the nanoFlex source equipped with a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (MS Vii). Instrument parameters: spray voltage of 2.4 kV; positive mode; capillary temperature 275°C; Mass range 375-1200 m/z (Full scan) in profile

mode in the Orbitrap with resolution of 120000; Fill time 50 ms with the AGC target set to standard. Data dependent acquisition (DDA) mode, MS/MS scans were acquired in the Orbitrap with a resolution of 15000, with a fill time of up to 54 ms and a AGC target of 400%. A normalized collision energy of 34 was applied (HCD). MS2 data was acquired in profile mode.

Data processing. The raw mass spectrometry data was processed with MaxQuant (v2.3.0.0) (PMID: 19029910) and searched against the uniprot-proteome UP000000625 (E.coli, 4450 entries, May 2022) database and an inhouse SF21 database containing the sequence of the proteins of interest and common contaminants. The data was searched with the following modifications: Carbamidomethyl (C) as fixed modification, acetylation (Protein N-term), oxidation (M) and phosphorylation (STY) as variable modifications. The default mass error tolerance for the full scan MS spectra (20 ppm) and for MS/MS spectra (0.5 Da) was used. A maximum number of 3 missed cleavages was allowed. For protein identification, a minimum of 2 unique peptides with a peptide length of at least seven amino acids and a false discovery rate below 0.01 were required on the peptide and protein level. Match between runs was used with default parameters. The msms.txt output file of MaxQuant was used for MS1 Filtering with Skyline (v21.1.0.278)¹⁹⁶.

APPENDIX. Structural analysis of the Cell Migration-Inducing and hyaluronan-binding Protein (CEMIP)

During my thesis, I worked on a structural biology project in collaboration with other research groups in Switzerland and Germany. Because the work I performed on this project has been substantial enough to participate to the research findings, but as I am not the main investigator, we chose to describe my scientific implication in an appendix and not in a second chapter in my thesis. The project, Structural analysis of the Cell Migration-Inducing and hyaluronan-binding Protein (CEMIP), was performed in collaboration with the research group of Prof. Dr. Leonardo Scapozza from the University of Geneva and was part of the thesis project of Dr. Sofia Spataro.

1. Introduction

The Cell Migration-Inducing and hyaluronan-binding Protein (CEMIP), also known as KIAA1199 or HYBID, is a 153kDa hyaluronidase which role is to depolymerize one of the most major components of the extracellular matrix, the glycosaminoglycan (GAG) Hyaluronic Acid (HA)^{197,198}. HA has a central role in providing a structural support and regulating the cell behavior in the interstitial extracellular matrix found in all human adult tissues and is one of the most common viscous GAG in vertebrate tissues¹⁹⁹. The expression of HA in mammals is performed by a family of enzymes called hyaluronan synthase, capable of synthesizing HA linear polymers with molecular masses from 1000 to 10.000 kDa^{200,201}. Quickly after its synthesis, HA is depolymerized by a family of enzymes called hyaluronidases²⁰² in small fragments with molecular masses from 10 to 100kDa. These enzymes show different expression and depolymerization patterns and are not all involved into global HA degradation *in vivo*. Organized events of HA polymerization and depolymerization are essential for the structural role of the extracellular matrix but also for very diverse other biological processes from cell signaling and migration, tissue morphogenesis to inflammation and cancer pathogenesis. Moreover, the role of HA is dictated by its molecular weight, as large HA fragments are involved into anti-angiogenic and anti-inflammatory functions whereas small HA fragments do the opposite²⁰³. The skin tissue in particular contains the highest amount of HA (50%) and its depolymerization is also faster here than in any other tissues²⁰⁴. Therefore,

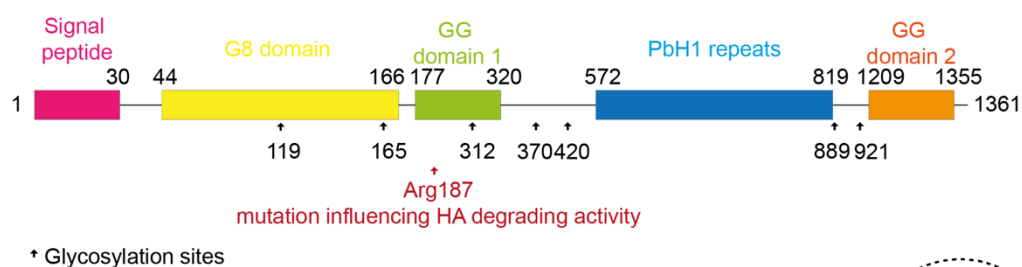
understanding the protein machinery behind this dynamic HA turnover in different tissues is critical.

CEMIP was described for the first time in 2013 as a protein involved in the specific HA degradation pathway in skin and synovial fibroblasts¹⁹⁸. This unique HA depolymerization mechanisms involves the interaction between clathrin-coated vesicles with CEMIP and the subsequent degradation of HA in early endosomes resulting in small HA fragments released in the extracellular space. What remains unclear is the molecular mechanism behind CEMIP's potential direct involvement in the enzymatic degradation of HA. It has been extensively showed that CEMIP is involved in numerous malignant and fibrotic diseases and for some of them, the importance of HA depolymerization has not be defined yet. In skin fibroblasts, mis-regulation of CEMIP can lead to wrinkling and sagging²⁰⁵ whereas in bones, it is associated to more severe phenotypes like osteoarthritis and rheumatoid arthritis^{206,207}. CEMIP has been shown to be involved into malignant tumor progress in model animals, but it remains unclear if, in humans, it is due to its link to HA depolymerization. To understand better the function of CEMIP in human cells, getting insight on its structural organization is necessary. Unfortunately, structural information about CEMIP are inexistent and my task in this project was to shed light on CEMIP's 3D structural organization. The main investigator of this project, Dr Sofia Spataro, had the tedious task to understand the biological role of CEMIP and to link the structure to its functions in human cells. All her findings and conclusions are available in her thesis.

2. Description of the protein and its AlphaFold structure prediction

CEMIP is a 153kDa protein composed of 1361 residues and divided into different domains with, for most of them, elusive and discussed roles (Figure 36. A). To help us visualize the organization and conformation of these domains into the protein, we used the structure prediction tool AlphaFold2 (AF2)^{150,192} (Figure 36. B). The predicted structure of CEMIP was performed with very high confidence by AF2 and we believe that its real structure does not differ drastically from the predicted one. First, on its N-terminal, CEMIP bears a 30 residues long signal sequence necessary for its excretion out of the endoplasmic reticulum²⁰⁸ but also for the N-glycosylation of the protein and its enzymatic activity²⁰⁹.

A



B

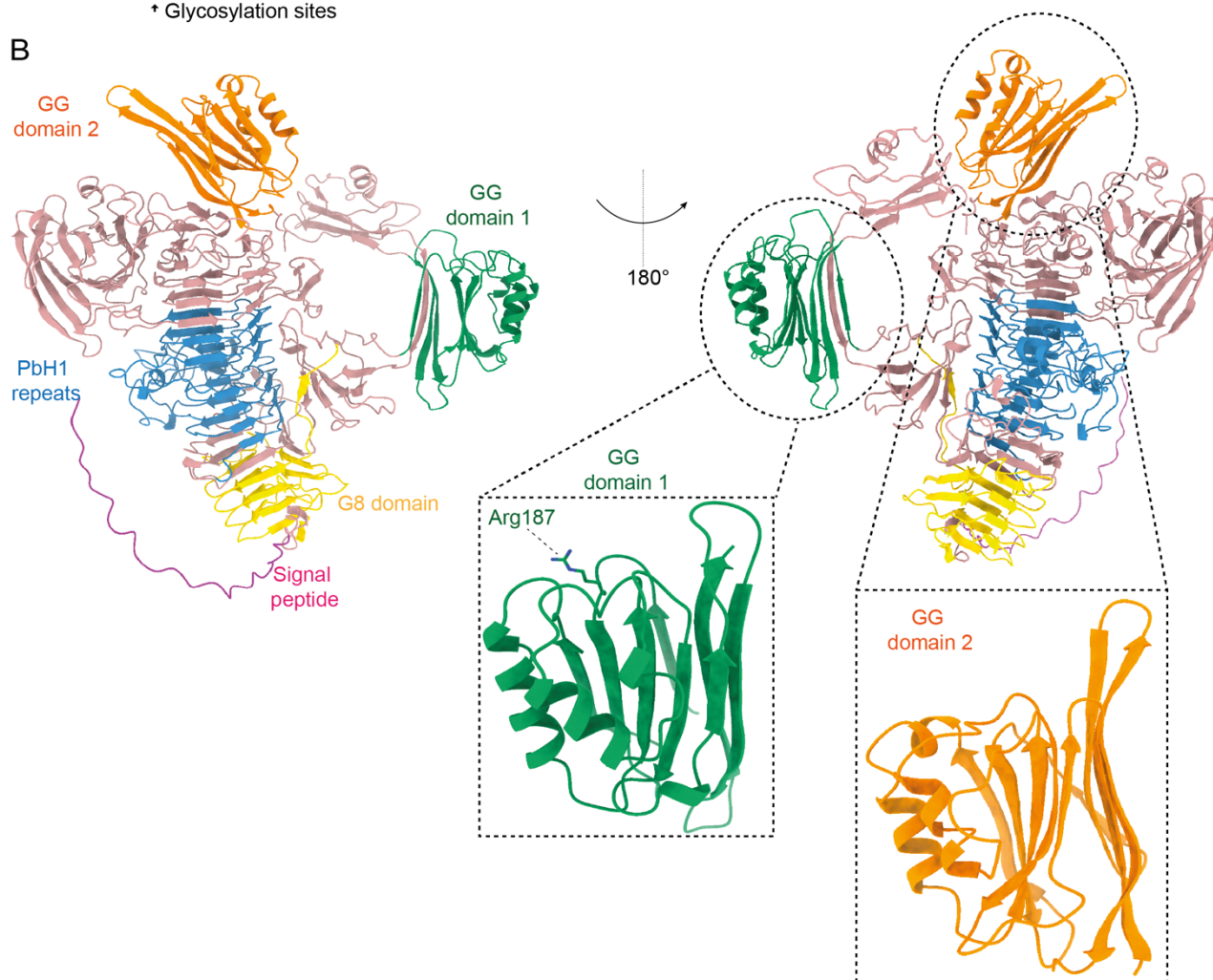


Figure 36. AF2 prediction of CEMIP structure

A, Schematic representation of CEMIP and its different domains boundaries where the export signal sequence (pink), the G8 domain (yellow), the first GG domain (green), the Pbh1 repeats (blue) and the second GG domain (orange). The black arrows represent identified glycosylation sites on their numerated residue. The red arrow represents the residue Arg187 in the first GG domain, residue known to play a role in the enzymatic activity of CEMIP on HA. **B**, AlphaFold2 prediction of the full-length CEMIP from two different sides. The same color codes as in A are used here to represent the different domain boundaries. The first and the second GG domains in green and orange respectively can be observed more in details in the zooms. The residue Arg187 in the GG domain 1 is also annotated in the structure.

Right after can be found the G8 domain in which are located eight well-conserved glycine residues. This domain is organized into five consecutive β strands folding as the start of the β roll core of the protein. From this G8 domain emerges the first GG domain. GG domains exist in a multitude of proteins and are composed of seven anti-parallel β strands and two α helices²¹⁰ but their precise functions in cells is still not well understood. In the case of CEMIP, a mutation in the residue Arg187 present in the first GG domain inhibits the degradation of HA in cells suggesting that GG domains may be the key structural features for CEMIP enzymatic activity^{198,211}. The next identified domain contains a series of PbH1 (Parallel β Helix 1) repeats found in enzymes capable of degrading polysaccharides. PbH1 domains fold into parallel β strands only to form a β roll structure²¹². This “stacks” of β strands starting from the G8 domain and continuing after the PbH1 repeats are predicted to compose the main stable core of CEMIP. They form a relatively “simple” protein folding mechanism questioning the evolution of these structures and their potential common biological functions. Finally, on the C-terminal part of CEMIP, a second GG domain can be found. As stated before, very little is known about the specific functions of these domains and to understand the function of CEMIP on HA degradation, validating structural information experimentally is crucial. At the beginning of my thesis, the main interest of this research project was on the first GG domain as it is in this one that a mutation has been identified to directly impact CEMIP’s enzymatic activity on HA. We decided therefore to solve the structure of the first GG domain by X-ray crystallography.

3. Attempts to determine the crystal structure of the first GG domain

3.1. Expression and purification of GG domain 1

CEMIP is a highly glycosylated human protein, and because of this, we chose to express the first GG domain in human HEK393 cells. The design of the GG domain 1 construct was performed before the apparition of AlphaFold and was based on homologies found with other GG domains in other proteins. Predictions of the secondary structures was performed on this region and based on these results; we chose to clone a recombinant 145 residues construct of GG domain with a C-terminal 6his-tag expressing a 19kDa protein (Figure 37). If we now show on the available AF2 prediction the boundaries of this construct, we can see that it does not include three of the seven domain β strands and the N-terminal part of the protein starts

in the β roll. Nevertheless, the protein was expressed and purified under a soluble form and called ig GG domain 1. It eluted as a single monomeric peak after gel filtration but on SDS-PAGE gel, we can observe the omnipresence of two close bands witnessing the potential degradation of the protein in solution. We could harvest a substantial amount of pure GG domain 1 and could concentrate it up to 22mg/ml. With this protein, we set up different crystallization plates with various buffer conditions stimulating the formation of protein crystals. We tried different concentrations of GG domain 1 and different ratios protein/buffer but could never obtain protein crystals. Nevertheless, this soluble and stable protein was used to perform functional assays by the other collaborators on this project.

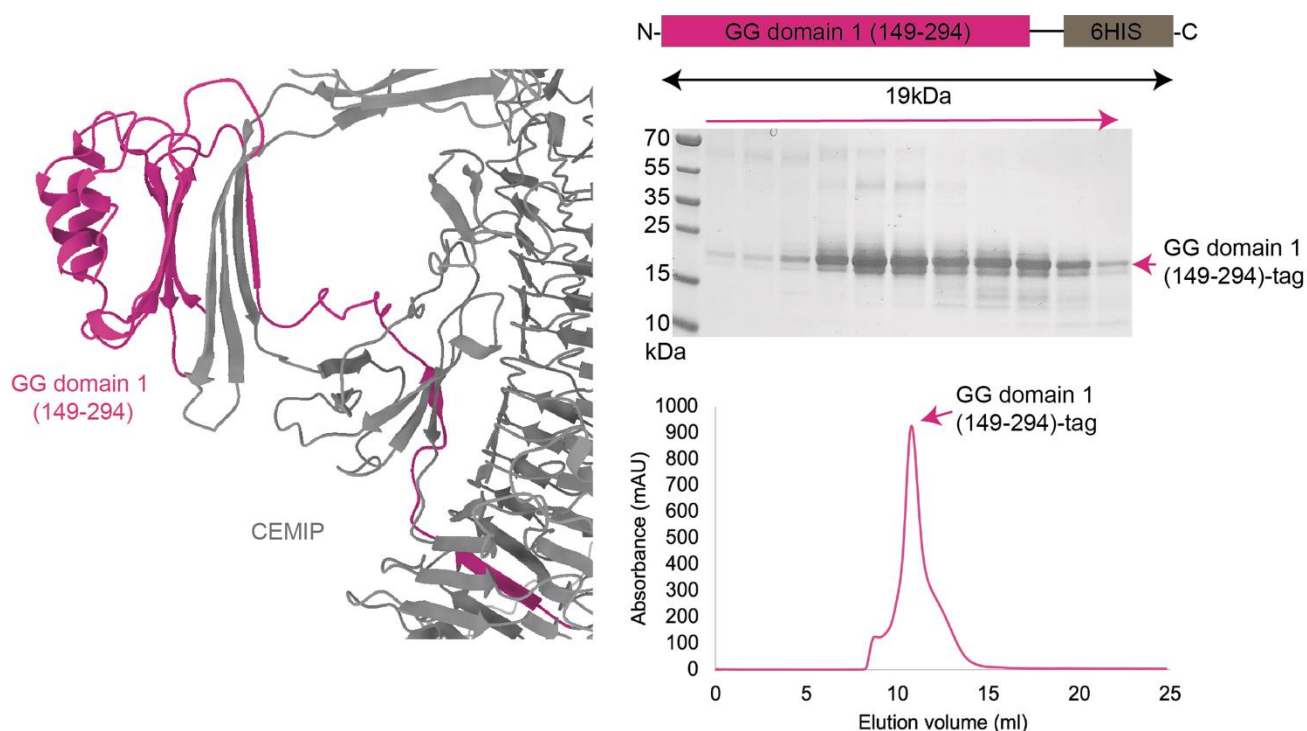


Figure 37. Design and purification of the first construct of GG domain 1

On the left is a part of the AF2 prediction of CEMIP with in pink the boundaries of the expressed GG domain 1 (149-294). On the top right is a schematic representation of the expressed 19kDa protein with its C-terminal 6HIS tag. The GG domain 1 was expressed in human cells and purified with nickel affinity beads followed by a size exclusion chromatography. The elution fractions were loaded on an SDS-PAGE gel, visible below where the protein is pointed with a pink arrow. The size exclusion chromatography elution profile of GG domain 1 is on the bottom right. The protein elutes at around 11ml on our Superdex increase 75 10/300 GL column.

3.2. Improving the protein crystallization chances

The difficulty to crystalize GG domain 1 protein made us wonder if the construct was optimal for crystallization. As noticed before, this construct englobes a part of the protein which was not predicted by AF2 to be part of the first GG domain and that could lead to

protein flexibility and reduced chances to form organized crystals. Many strategies have been developed to overcome protein crystallization issues as only 15-20% of purified proteins can form a crystal from which a structure can be determined. It is known that stable structured globular proteins have more chances to crystallize than unstructured proteins and one technique that showed remarkable success into promoting the crystallization of proteins is called *in-situ* proteolysis²¹³. The technique takes advantage of the specific proteolytic enzymatic activity of peptidases to cut fragments of a protein of interest resulting in a small, more stable protein. In dusty amounts, these proteases are incubated with the protein for a short period of time before the set up a crystallization conditions, leaving enough time for the protease to cut the easily accessible unstructured regions first. Because of the extremely small amounts of proteases in the crystal drop, the chances that this protein forms protein crystals are almost inexistant. Therefore, we wanted to perform such experiment on the GG domain 1 and observe if, after a short period of time with trace amounts of four common peptidases (elastase, trypsin, subtilisin and GluC), we could obtain a more compact form of the protein (Figure 38. A). We decided to work at a high GG domain 1 protein concentration (22mg/ml) and to try four different dilutions of our proteases (1:50, 1:100, 1:500 and 1:1000), each of them at an original concentration of 1mg/ml. We mixed the proteins together and observed on an SDS-PAGE gel the results after 30 minutes digestion at room temperature. As a first observation, elastase did not digest the GG domain 1 whereas all the three others did. Interestingly, after 30 minutes and at 1:50 and 1:100 ratios, the endoprotease GluC digested GG domain 1 into a stable form visible at a molecular mass of 12-13kDa (Figure 38. A). It is also to notice that after 2 hours of reaction, the protein completely precipitated highlighting the temporal importance of this technique. Consequently, we used the same protease to perform our *in-situ* proteolysis experiment in diverse crystallization conditions. In a set 100µl protein drop on crystal plates, 2µl of GluC (1mg/ml) was added and microscopic protein crystals formed in 20mM MES pH9, 200mM ammonium sulfate and 30% PEG5000 (Figure 38. B). Unfortunately, those crystals are too small to be analyzed and we are still working on increasing their size in the perspective to resolve a high-resolution structure of the GluC-digested GG domain 1.

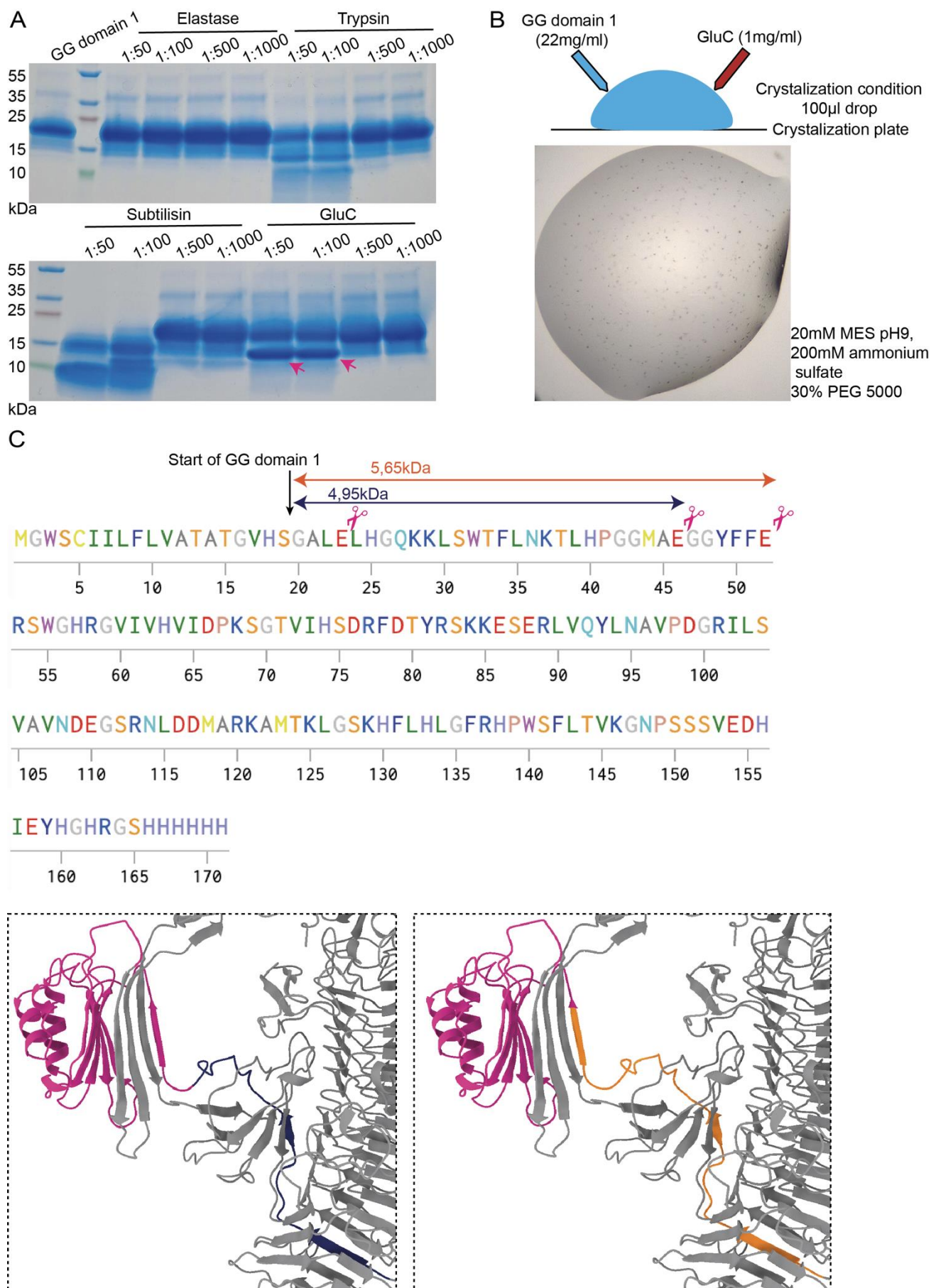


Figure 38. In-situ limited proteolysis of GG domain 1

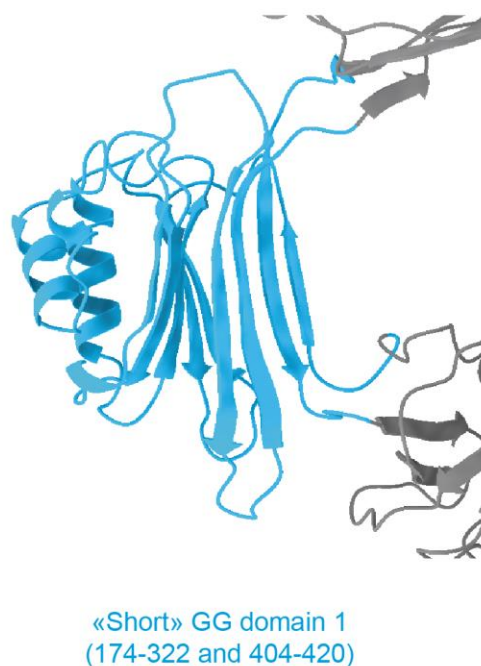
A, Followed by a 30 minutes proteolysis reaction with different proteases at room temperature, the digested GG domain 1 protein was loaded on a SDS-PAGE gel displayed here. Different concentrations of the four proteases (elastase, trypsin, subtilisin and GluC) were tested (1:50, 1:100, 1:500 and 1:1000) and the digested results were compared to the undigested GG domain in the first lane from the left of the top panel. The pink arrows point at an observed stabilized digested fragment of GG domain 1 that led us to decide to set up *in-situ* proteolysis crystallization assays with the GluC endoproteidase. **B**, On the top panel is a schematic explanation of the crystallization set up. Each 100µl drop contains 50µl of GG domain 1 at 22mg/ml and 50µl of a tested crystallization condition to which was added 2µl of GluC at 1mg/ml. On the lower panel is a picture of one condition in which microscopic protein crystals grew. **C**, On the top panel is displayed the amino-acid sequence of the GG domain 1 with its first residue being Gln20. The pink scissors represent the favored cutting sites of GluC, on Gln residues of the unstructured parts in the N-terminal region. The molecular weight of potentially digested regions was calculated and displayed in blue and orange above the sequence. The same colors were used to show these regions on the AF2 prediction of CEMIP on the lower panels.

The endoprotease GluC preferentially cleaves the C-terminal peptide bonds of glutamic acid residues (Glu, E) and if we look at the amino-acid sequence of our GG domain 1 protein, we can see the presence of three Glu residues in the N-terminal part of the protein (Figure 38. C). We know that the observed digested stable form of the GG domain 1 has an observable molecular mass around 12-13 kDa on our SDS-PAGE gel. Two of the three Glu residues are located in a way that, if cut by GluC, the digested regions would be 4,95 and 5,65kDa resulting in a shorter GG domain 1 with a molecular mass potentially corresponding to the one observed on our SDS-PAGE gel. If we now correspond these calculations to the AF2 prediction of CEMIP, the potential cut region of GG domain 1 would be the first β strand in the β roll and could be extended to the following β strand. These results show us that the protease GluC most likely stabilize the GG domain 1 by digesting the regions that do not belong to the AF2 predicted first GG domain.

As we discussed before, the design of the GG domain 1 was performed before the development of AlphaFold and our previous results highlight the importance of the protein stability to form crystals. Once a AF2 predicted structure of CEMIP became available, we started to design new constructs to express and purify the first GG domain. Two new constructs were designed, a “short” and a “long” GG domain 1 (Figure 39. A). The “short” GG domain 1 includes the residues 174-322 and 404-420. CEMIP domains boundaries predictions showed that the first GG domain is located between the residues 177 and 320 but if we look at the AF2 prediction, we can see that a β strand predicted to be part of the GG domain is not included in this sequence and it is because between the predicted GG domain sequence and this β sheets, an 82 residues-long sequence is present and predicted to fold independently to the GG domain. The “short” GG domain 1 construct excludes those 82 residues and directly links the last β strand to the rest of the GG domain. The “long” GG domain 1 construct does the opposite by including those 82 residues. As the GG domain 1

showed excellent protein expression and good stability in solution, we decided to try the expression of our new constructs in bacteria as we believed that we didn't need the sophistication of human cells expression machinery for these small proteins. The proteins were recombinantly expressed with a C-terminal 6HIS tag. Unfortunately, neither "short" or "long" GG domain 1 were soluble and all the protein was found in the insoluble fraction (pellet) (Figure 39. B).

A



B

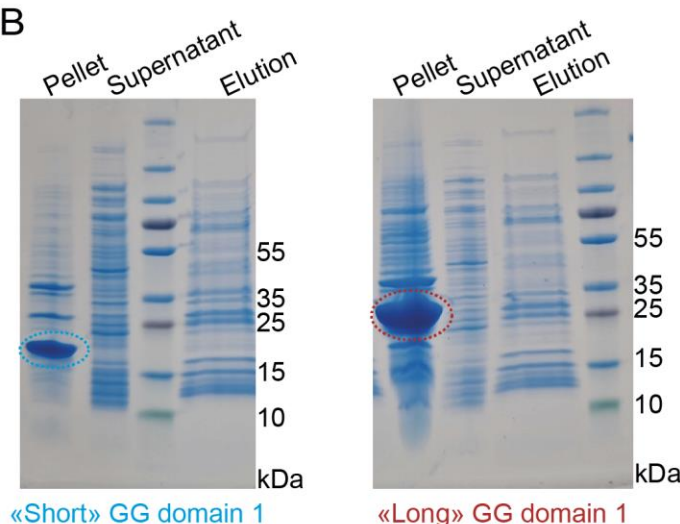


Figure 39. Improving the protein crystallization chances with different GG domain 1 constructs

A, The "short" and "long" GG domain 1 newly designed constructs are displayed here on the AF2 predictions of CEMIP in blue and red respectively. **B**, The two new GG domain constructs were expressed in bacteria and purified. The results of these purifications were loaded on SDS-PAGE gels visible here. The insoluble (pellet), soluble (supernatant) and elution fractions are displayed here. Using the same color code as in A, the expressed GG domain 1 constructs are circled in the insoluble fractions and cannot be found in either the soluble or the elution fractions

4. Electron Microscopy

AF2 predicted CEMIP to share high structural similarity to the human transmembrane hyaluronidase TMEM2, from which the structure was resolved recently^{214,215}. To understand the function of the full length CEMIP in humans, we decided to also try to elucidate the structure of the full-length CEMIP. In human cells, CEMIP is found in relatively low amounts enhancing the difficulty reach concentrations high enough for crystallization assays. Therefore, we decided to use cryoEM that requires less protein. CEMIP has a relatively high molecular mass of 153kDa and a predicted stable globular shape which makes it a good candidate for the technique. Nevertheless, the protein is mainly composed of β sheets that are not well defined at medium resolution like α helices which imposes us to reach high resolution to resolve precisely the structure of the full-length CEMIP.

4.1. Attempts to express and purify the full-length CEMIP

Our first goal was to express and purify full length CEMIP by ourselves in our insect-cells expressing vectors. We decided to exclude the first unstructured 30 residues comprising the signal export sequences but for the rest of the protein, we kept it as it was. We recombinantly expressed CEMIP 30-1131 in insect cells with a N-terminal double streptavidin tag (2xStrep) (Figure 40. A). Mass-spectrometry validated our CEMIP expression and validation with a 54% coverage of the protein. Nevertheless, the protein is not expressed in high quantity, and we could never obtain a pure sample. Moreover, we could never perform a final size exclusion chromatography, CEMIP seems to attach the gel filtration resin or to disassemble during the procedure and that is why we only displayed here the affinity column results. Another issue is that, even after extensive ATP wash of the affinity column, the contaminants always remain. Mass-Spectrometry showed the dominant presence of folding chaperones leading us to think that the protein is not well folded into our insect cells expression system. In parallel to this experiment, CEMIP was expressed and purified by a company (Genscript) with high purity but still at low concentration (Figure 40. B). We tried to concentrate the protein, but it aggregates very easily so we decided to work with the available concentration of 0,05mg/ml.

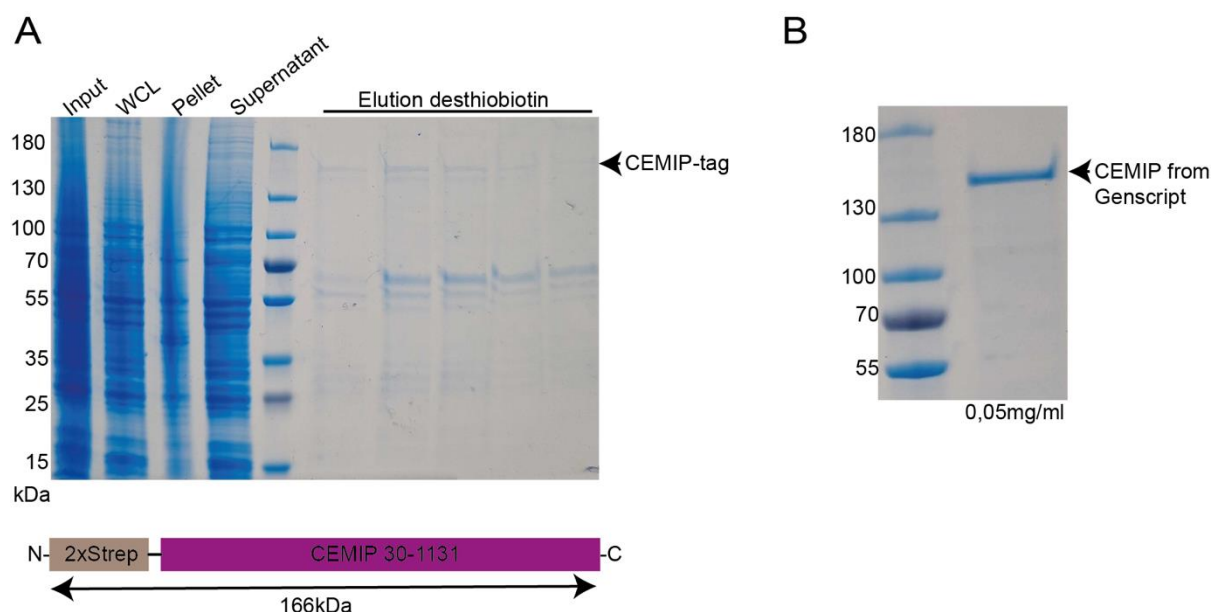


Figure 40. Expression and purification of the full-length CEMIP

A, CEMIP (30-1131) was expressed recombinantly with a N-terminal 2xStrep tag forming a 166kDa protein as displayed in the schematic representation of the protein. This protein was expressed in insect cells and different steps, and the results of the purification were loaded on an SDS-PAGE gel displayed here. The input represents the insect cells before lysis where the whole cell lysis (WCL) is after lysis mechanism. The insoluble (pellet) and soluble (supernatant) of the WCL were loaded as well but CEMIP was not expressed in enough high quantity to distinguish it from other proteins in these fractions. The results of the elution from the affinity streptavidin column with desthiobiotin show that CEMIP was successfully purified but only in small quantities and always with contaminating proteins. **B**, the full-length CEMIP was expressed and purified by Genscript and we loaded the protein on a SDS-PAGE gel at the given 0,05mg/ml concentration. The protein is relatively pure and migrates at the expected molecular weight.

4.2. Negative staining

To first analyze the structural organization of CEMIP, we decided to perform negative staining experiments. As the studied protein doesn't need to be embedded into an ice layer but applied to a carbon surface at room temperature, the working protein concentrations can be diminished compared to cryoEM. Therefore, we could work at our actual concentration of 0,05mg/ml and lower. CEMIP is found to be well dispersed on the EM grid and is well stained (Figure 41). There was very little amount of protein aggregates allowing us to collect 320 micrographs and to pick and extract many particles. We selected around 20.000 and performed an 3D *ab-initio* reconstruction of CEMIP at low resolution. What we can see is a small protein with a size corresponding to the AF2 prediction of CEMIP harboring a main core and two arms which could be linked to the AF2 predicted central β roll and the two emerging GG domains. To know more about the details of the structure of CEMIP, we needed to push the final resolution by working in cryogenic conditions.

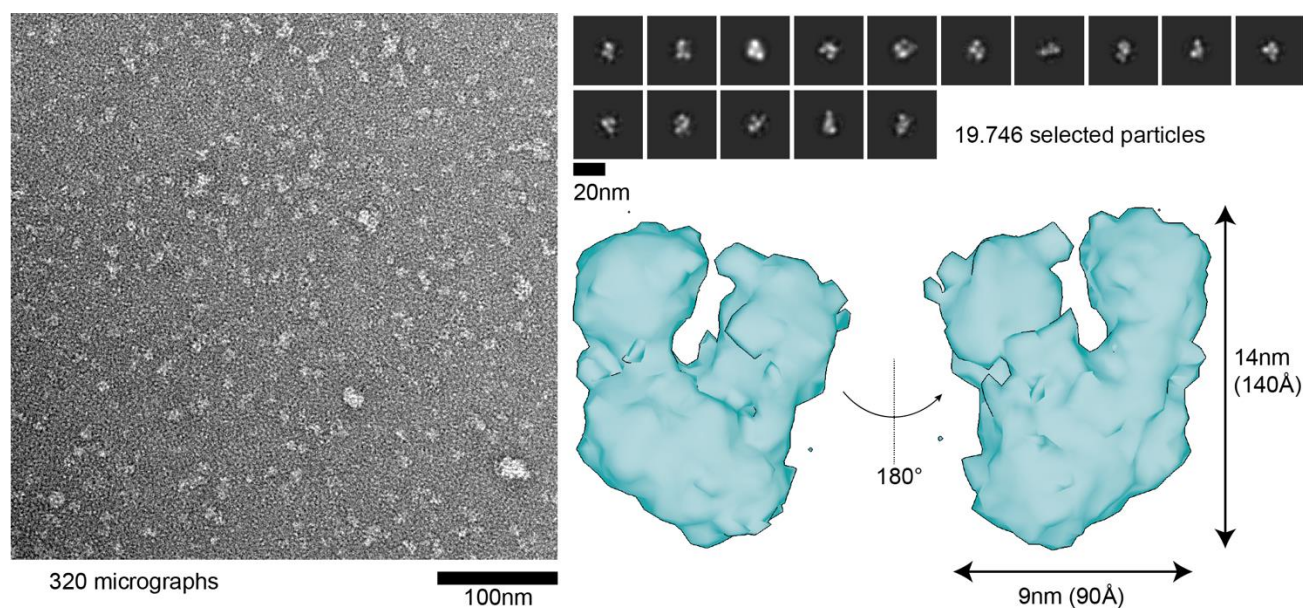


Figure 41. Negative staining of CEMIP

On the right panel is a representative electron micrograph of negative stained full-length CEMIP proteins. Particles are visible in white where the background is stained in dark grey. On the top right panel are the extracted and averaged in 2D classes the CEMIP particles picked from the 320 collected micrographs. 19.746 particles were selected and used to form an ab-initio 3D reconstruction of CEMIP displayed in the lower right panel in different orientations. The dimensions of this volume were measured with a 14nm height and 9nm length corresponding to what is expected from CEMIP if the AF2 prediction is correct.

4.3. CryoEM

As previously mentioned, our CEMIP protein was expressed and produced by an industrial company (Genscript) and not by us and was delivered to us frozen at a concentration of 0,05mg/ml. The formation of a nice thin layer of ice comprising a single layer of protein in different orientations in the holes of a cryoEM grid requires a certain protein concentration. Moreover, protein like to attach to the holey carbon film on the grid and sometimes escape from the ice layer reducing the number of particles present in the ice. For these reasons, we could not work at the original given concentration of CEMIP and needed to increase it to obtain good, plunged EM grids. We concentrated CEMIP to 0,2mg/ml and prepared Quantifoil 300 mesh copper with a carbon layer 1,2/1,3 and imaged our EM grids at a lower magnification than usual (pixel size of 1.998Å) (Figure 42. A). What we observed was the apparition of organized protein aggregates with particles in the exact same orientation where the ice was the thinnest in the hole. We could pick and extract some of those particles but all of them showed very similar orientations and multiple particle orientations in ice are required to reconstruct a high-resolution 3D model of the studied

protein. CEMIP is a difficult protein to embed in pure ice and we believe that this semi-organized layer of protein is due to the close proximity of CEMIP to the water/air interface.

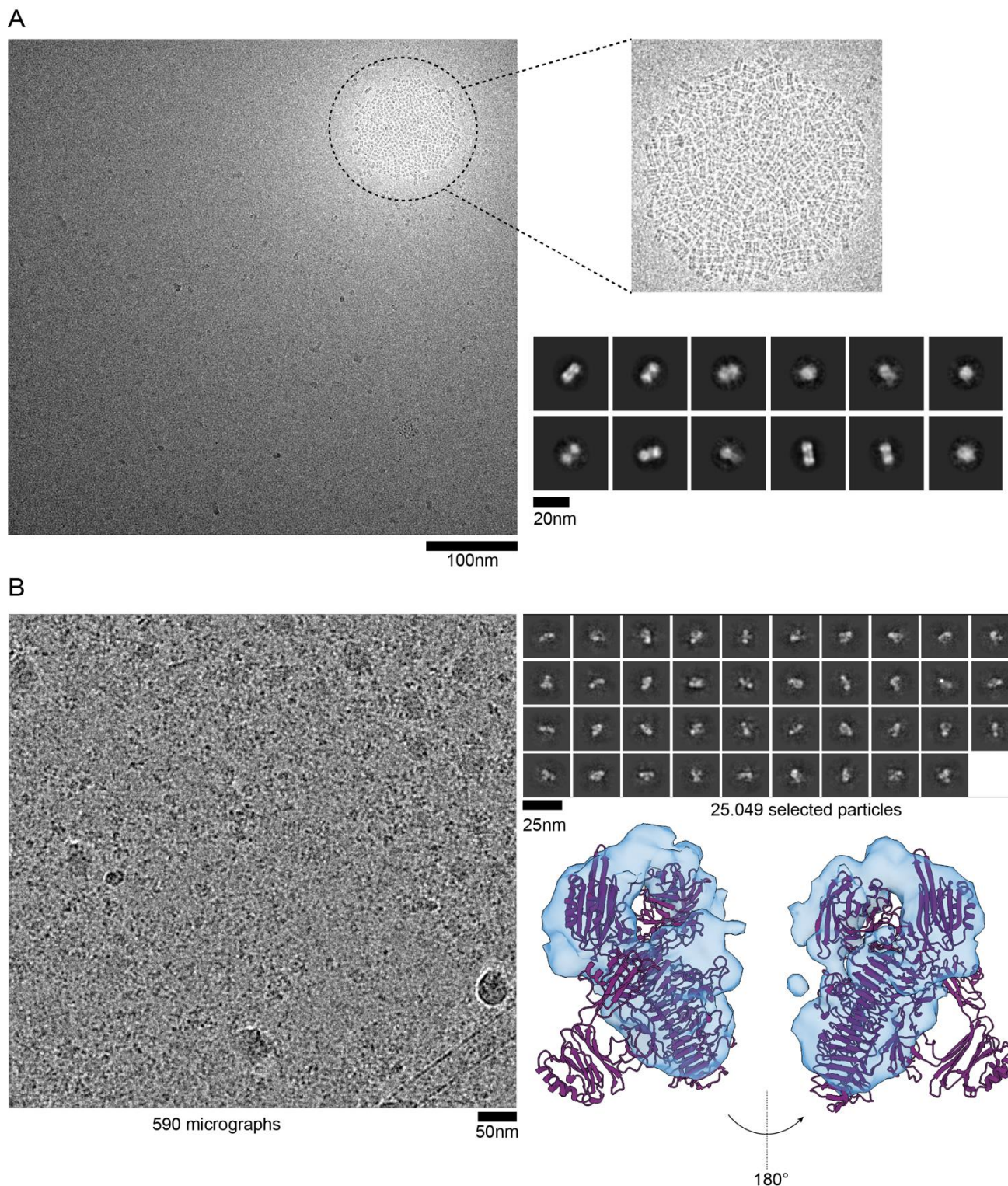
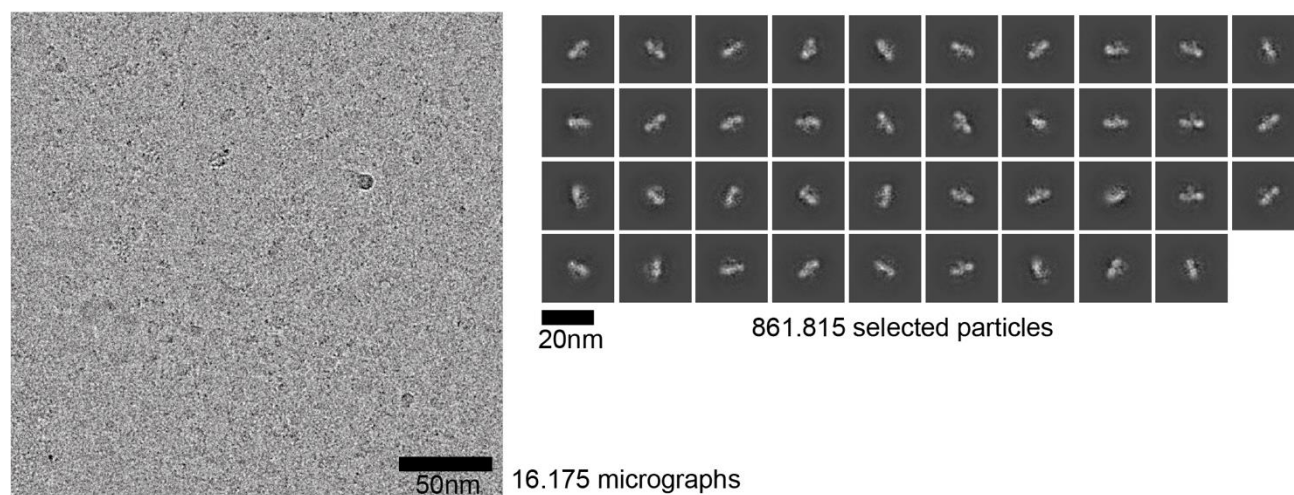


Figure 42. Graphene oxide-coated EM grids improve the repartition of CEMIP particles into the ice
A, On the left is a representative cryo-electron micrograph on CEMIP particles embedded into ice. Particles are nearly absent in the dark grey areas where the ice is thicker and are assembled in the light grey area where the

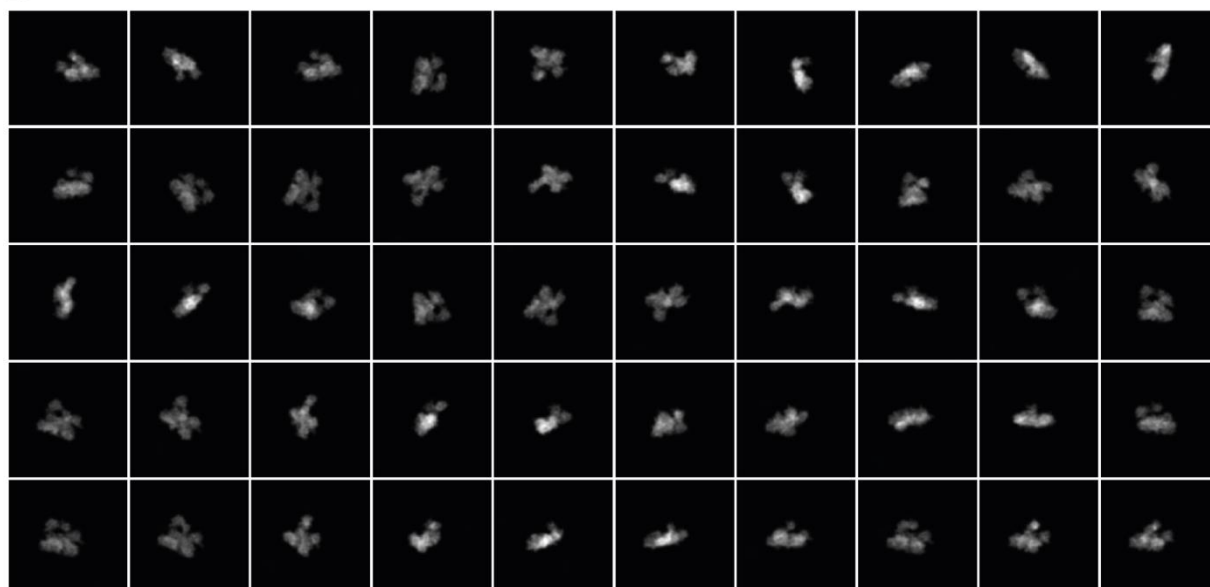
ice is the thinnest is semi-organized aggregates. On the top right panel is a zoom on this observed aggregation and it can be observed that all particles are orientated in the same way. Particles could be nevertheless picked, extract and averaged into 2D classes visible on the bottom right panel but most of them show the same orientation limiting the possibility to reconstruct a 3D volume of CEMIP. **B**, On the left is a representative cryo-electron micrograph of CEMIP particles embedded into ice and sited on a graphene oxide support layer. A dataset of 590 micrograph was collected on our 200kv Talos Arctica microscope from which CEMIP particles could be picked, extracted and averaged into 2D classes visible on the top right panel. The selected 25.049 particles were used to reconstruct a low-resolution 3D *ab-initio* volume of CEMIP displayed on the bottom right panel in different orientations. The AF2 prediction of CEMIP was fitted manually into the volume emphasizing the limited poor resolution and the missing regions of CEMIP due to their high dynamics in solution.

These first cryoEM results informed us on the challenging behavior of CEMIP in pure ice and on the necessity to use other techniques to attract the protein into ice. For this purpose, we prepared grids with a graphene-oxide support layer. With the help of support layers, we can reduce the protein concentration which allow us to not concentrate CEMIP and use it at its original 0,05mg/ml concentration (Figure 42. B). What we observed was an improved distribution of particles in ice with the notable presence of protein aggregates. Nevertheless, we managed to pick and extract single particles and average them into 2D classes. Particles alignment resulted into blurry 2D classes with a preferential orientation of T-shaped particles, similar to what we observed in pure ice (Figure 42. A). Out of the 25.049 selected particles on the 590 collected micrographs, we could reconstruct a low-resolution *ab-initio* 3D model of CEMIP resembling the one obtained with negative staining. If we try to fit the AF2 prediction of CEMIP in this model, we can see that not all the predicted structure matches with the volume showing us that, either the AF2 prediction is not correct or, what is more likely to happen, the protein region where the first GG domain is located is too flexible compared to the main core of CEMIP and therefore, its movements in ice makes its alignment challenging excluding it in the final reconstruction. Nevertheless, this small data collection on graphene oxide-coated grids showed us that support layers improved the repartition of CEMIP particles in ice and allowed us to work with low protein concentrations. Therefore, we decided to pursue with this approach and to collect a bigger dataset.

A



B



C

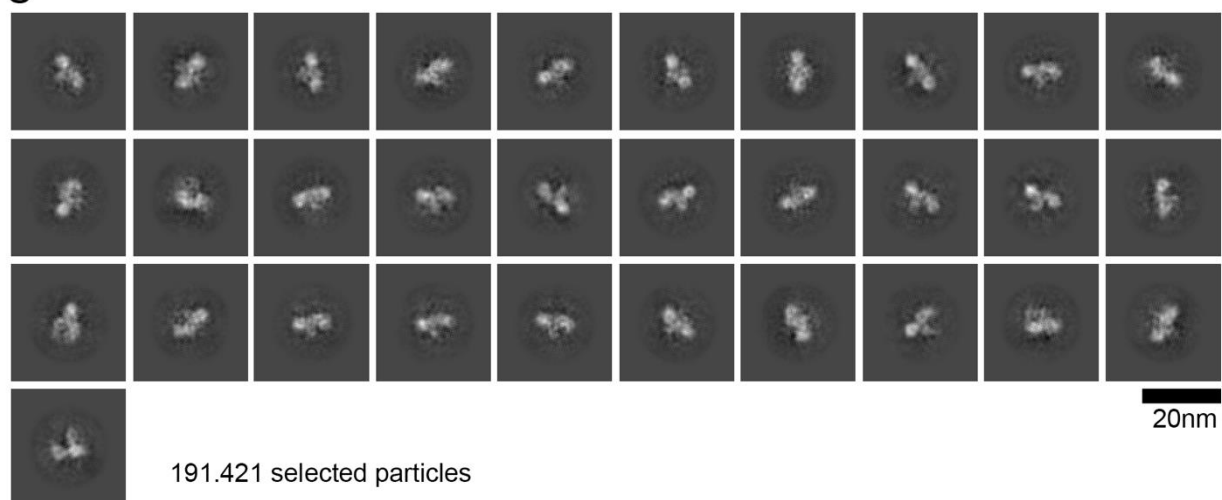


Figure 43. Collection of a big dataset on a 300kv Titan Krios microscope

A, On the left panel is a representative cryo-electron micrograph of CEMIP particles embedded into ice and sited on a graphene oxide support layer. 16.175 micrographs were recorded, and 861.815 selected particles were picked, extracted and averaged into 2D classes visible on the left panel. **B**, AF2 predicted CEMIP structure volume was used to generate 2D templates of the particle. These templates were used to pick with more accuracy CEMIP particles on the micrographs and improve the chances to pick rare views. **C**, The template generated in B helped picked the particles averaged in 2D classes displayed here.

At the time we decided to collect a higher number of micrographs, we were not equipped yet with a high-end image detector on our 200kev Talos Arctica microscope and decided to collect on a 300kev Titan Krios microscope in the DCI Lausanne EM facility. We prepared similar graphene oxide-coated grids as before with the same CEMIP concentration and collected 16.175 micrographs out of the best prepared EM grid from which we could pick and extract 861.815 particles as a start (Figure 43. A). The micrographs do not show the same contrast as the ones imaged on our Talos Arctica and the 2D averaged particles show a preferential orientation in ice. To improve the implanted template picking tool in CryoSPRAC outcome, we decided to generate 2D templates based on the AF2 prediction of CEMIP. With this approach, we can generate backprojected 2D classes from the AF2 predicted 3D volume. These 2D classes were used as a template to pick similar particles in ice (Figure 43. B). If we first observe these templates, we can see that we can find similar averaged particles in our results like the T and Y-shaped classes, but we also get classes that we do not find in our 2D classes like the X-shaped ones. Interestingly, after using these templates to pick more precisely CEMIP particles, we still could not obtain X-shape 2D classes (Figure 43. C). Unfortunately, we could not obtain better 2D classes than what was obtained on our 200kev Arctica microscope and could not improve the final resolution of the 3D volume of CEMIP. CEMIP is a challenging protein to structurally study and we cannot exclude that the problems we are encountering are not mainly because we are not working with freshly expressed protein. Indeed, freshly purified protein shows greater results than stored frozen protein and we would like to reproduce these experiments on freshly expressed protein in human cells. Moreover, we would like to improve the amount of protein purified to be able to use pure ice and compare the results with graphene oxide grids.

5. Discussion

5.1. Crystal structure of the first GG domain

During my thesis, we tried with our collaborators to obtain a high-resolution structure of the first GG domain of CEMIP. This domain showed before its potential participation into the catalytic role of CEMIP on HA and its mutation led to catastrophic functional consequences in different diseases implicating HA depolymerization²¹¹. We chose crystallography to analyze this domain because of its small molecular mass and its predicted globular structure making it a good candidate for forming organized protein crystals. Before the accession of AlphaFold into standard structure prediction analysis pipeline, we designed a GG domain 1 construct based on secondary structure organization predictions and obtained from protein purification a stable soluble protein that we used extensively to perform crystallization plates and to try to form protein crystals. Despite our numerous attempts to improve the crystallization conditions by performing *in-situ* limited proteolysis and by designing new constructs of the GG domain 1 based on the AF2 structure predictions, we could never get suitable crystals for X-ray analysis in a synchrotron facility. Nevertheless, understanding what makes the protein soluble and stable is the key to overcome this issue. We expressed our new constructs in bacteria where the expression of the GG domain 1 was performed in human cells, an already possible explanation on the lack of stability of the new constructs in solution. The capacity of bacteria to rapidly express large amounts of proteins at a reasonable price was for us an important point to consider and these expression systems are normally used to express small, simply folded proteins like the GG domain 1. Nevertheless, we cannot exclude post-translational modifications like glycosylations (impossible to perform by bacteria) are required for the proper protein folding and stability. We would need to repeat those experiments with human cells-expressed recombinant protein.

5.2. Cryo-EM analysis of the full-length CEMIP

During my thesis, we tried with our collaborators to get the structure of the full-length CEMIP in order to understand its structural organization and link it to the functions observed for HA depolymerization. We tried to express and purify our own protein in our insect cells expression system, but we could not purify a soluble form of CEMIP and decided to pursue

our experiments with an industrially produced CEMIP. The two main issues that we encountered were the original protein concentration and the almost impossibility to increase it. Moreover, the available volume of sample and the great sensitivity of CEMIP to form aggregates led us to work at the original concentration of 0,05mg/ml. We could perform a first negative staining visualization of the protein without difficulties, but our cryoEM experiments were impacted by these issues. Because of the concentration, we opted for a graphene oxide-coated EM grid option attracting particles in ice and is supposed to help randomize particles distribution as well. We did observe a major improvement in particles distribution, and we could pick and extract particles and average them into 2D classes but noticed that particles mainly show a preferential orientation in ice. Moreover, we could not observe protein features in our 2D classes, but this is because CEMIP is mainly composed of β sheets, less visible at low resolution. The selected particles could not produce a high-resolution 3D volume of CEMIP but was enough to tell us that, based on the AF2 prediction, not the entire protein can be averaged and reconstructed, and we believe that this is because of the molecular dynamics of these regions. The logical continuity of this project is to express and purify CEMIP protein in human cells to improve the quality of the sample but also to increase the protein concentration. TMEM2 (106-1383), a transmembrane hyaluronidase sharing a very similar domains organization than CEMIP was recently crystalized²¹⁵ and the obtained structure confirmed the high accuracy of the AlphaFold prediction of the same protein which could lead us to believe that the same would result from the structure of CEMIP. TMEM2, as well as CEMIP, cannot depolymerize modified HA but, in contrary to CEMIP, TMEM2 has an enzymatic activity on HA *in vitro* even though the two proteins have almost an identical structure. The difference between the two could reside in their subcellular localization as CEMIP is not a membrane protein and could require the binding to a partner to activate its enzymatic activity.

5.3. Potential future structural analysis of CEMIP interactions with HA and with other protein partners

Sofia Spataro in her Ph.D. thesis explained that CEMIP shows an enzymatic activity on HA only in cells and not *in vitro*. it is therefore conceivable that CEMIP can bind HA but not depolymerize it without another cellular signal that is missing in *in vitro* experiments. It would therefore be interesting to identify this unknown protein partner and study its complex

formation. Long HA polymers form a viscous gel at high concentration, but short polymers of HA can be diluted and mixed with CEMIP to form complexes. Interestingly, sulfated HA was showed to be an inhibitor of CEMIP enzymatic activity in cells but what is not known if it is because CEMIP and the HA cannot interact anymore or if the sulfation blocks structurally the cutting activity of CEMIP²¹⁶. It would be therefore interesting to study the interactions between the two components and uncover the specific enzymatic activity of CEMIP. Even though the molecular regulation of CEMIP in cells is still not completely understood, interactions with other proteins involved in different signaling pathways in cells have been identified and could be used to resolve the structure but also understand the molecular dynamics of CEMIP. CEMIP was not shown to interact with the cellular membrane but interacts with different membrane proteins like the Epidermal Growth Factor Receptor (EGFR) and the transmembrane receptor Plexin-A2 (PLXNA2)²¹⁷ and to form stable complexes involved in EGFR signaling pathway. The molecular mechanism of HA depolymerization by CEMIP is still not completely understood and the existing literature and our results point at the fact that CEMIP potentially needs an interacting partner to be stabilized *in vitro* and to mediate its enzymatic activity. Identifying those protein partners would be interesting to stabilize CEMIP *in vitro* and obtain its high-resolution structure.

Another approach to improve the stability of CEMIP in ice with a protein interactor would be with the addition of monoclonal fragments antigens binding (Fab). Fab fragments have a molecular weight of 50kDa and are capable, through their antigen binding site, to bind to a specific target protein and to form rigid and stable complexes²¹⁸. Specific antibodies against human CEMIP exist and Fab fragments can be isolated with simple protease papain digestion using immobilized papain agarose beads. The isolated Fab fragments can be after mixed with the target protein and analyzed with cryoEM.

6. Material and methods

I would like to mention once more that this project has been conducted in the collaboration with the research laboratory of Prof. Dr. Leonardo Scapozza from the University of Geneva, the research group of Andrea Cavalli from the Institute for research in biomedicine in Bellinzona and the research laboratory of Jonathan Sleeman from the University of Heidelberg.

6.1. AlphaFold predictions

CEMIP full-length structure prediction is available on the AlphaFold Protein Structure Database (AlphaFoldDB: Q3WUJ3) and was analyzed using UCSF ChimeraX¹⁹³.

6.2. Cloning and protein expression

All DNA constructs were designed and analyzed on the cloud-based platform Benchling (<https://www.benchling.com/>).

Human cells. The human GG domain 1 (149-294) was amplified from the human CEMIP gene (UniProt code Q8WUJ3 and sequence synthesized by Genscript) by PCR using primers with suitable overhangs and cloned into a prp plasmid (Addgene, #41841) with a CMV promoter and puromycin resistance using the Gibson assembly cloning protocol (NEB). This procedure was performed by Sofia Spataro, at the time Ph.D. student in the research laboratory of Prof. Dr. Leonardo Scapozza at the University of Geneva. 10mg of plasmid was amplified using a plasmid giga kit (Qiagen) and used to transfect HEK293 cells by the Protein Production and Structure core facility in EPFL. The expressed protein in cells was returned to us for purification.

Bacteria. Codon-optimized genes for GG domain 1 “short” (174-322 404-420) and “long” (174-420) (sequence synthesized by Genscript) were amplified by PCR using primers with suitable overhangs and cloned into a pet24a+ plasmid (Addgene, #69749-3) using the Gibson assembly cloning protocol (NEB). DNA was isolated following the NucleoSpin Plasmid kit protocol (Macherey & Nagel) and validated through Sanger sequencing (Microsynth AG). This procedure was performed by the research team of Prof. Dr. Andrea Cavalli at the IRB Bellinzona. The further expression and purification of the proteins were performed in by me. The constructs for expression of N-terminal 6his-tagged proteins were transformed into competent BL21 (DE3) E.Coli. 1L cell cultured were grown, after induction with 0.5 mM IPTG, 3 hours at 25°C before being harvested and frozen in liquid N2 for later use.

Insect cells. The human CEMIP gene (sequence synthesized by Genscript) was amplified by PCR using primers with suitable overhangs and cloned into a pACEbac plasmid (Geneva Biotech) using the Gibson assembly cloning protocol (NEB) and were expressed in *Spodoptera frugiperda* pupal ovarian tissue cell line 9 (Sf9) insect cells (Invitrogen) using

baculoviruses expression system. To do so, competent DH10 Multibac bacterial cells were transformed with the generated pF1 plasmid constructs. DNA was isolated following the NucleoSpin Plasmid kit protocol (Macherey & Nagel) except for the elution step which was performed by precipitation with 70% ethanol. Purified bacmids were resuspended in 100µl Sf900 medium (ThermoFisher Scientific) and mixed with 5µl transfection reagent (Fugene Promega) and added to 1.0×10^6 of adherent Sf9 cells per ml in 6-well plates in a total of 2ml. After 72h incubation at 27 °C, cells were centrifugated and the supernatant (P1) was harvested, filtered and inoculated to 3.0×10^6 Sf9 cells per ml in suspension in a total of 50ml. After 72h at 27 °C at 110 rev/min, the cells were centrifugated and the supernatant (P2) was harvested, filtered and stored at 4°C for further use. To perform large-scale protein expression, a new round of baculovirus amplification in 3.0×10^6 Sf9 cells per ml in suspension in a total of 50ml is necessary with P2 from which P3 is collected. Typically, 25 ml recombinant P3 baculovirus was used to infect 500 ml of Sf9 insect cells at a cell density of roughly 2.0×10^6 cells per ml. The cells were incubated for 72 h at 27 °C at 110 rev/min, harvested at a cell viability rate of 85–90%, flash-frozen in liquid nitrogen and stored at –80 °C.

6.3. Protein purification

GG domain 1. Purification of all proteins and protein complexes was performed at 4 °C. Cells expressing the GG domain 1 were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 8, 300mM NaCl and 10mM Imidazole) containing protease inhibitor cocktail tablets (PIC) (complete EDTA-free; Roche Diagnostics), and 5 units/ml supernuclease (Novagen), subsequently sonicated and centrifuged for 1 h at 20,000g. The soluble fraction was applied (1ml/min flow rate) onto a 5 ml HisTrap HP excel affinity column (Cytiva). The beads were washed with wash buffer (50mM HEPES-NaOH pH 8, 300mM NaCl and 50mM Imidazole). Proteins were eluted with elution buffer (50mM HEPES-NaOH pH 8, 300mM NaCl and 500mM Imidazole) before being concentrated and applied on a Superdex 75 Increase 10/300 GL (GE Healthcare Life Sciences)) size-exclusion column as a final purification step with a final buffer (20mM HEPES-NAOH pH 8 and 150mM NaCl). The samples were applied on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue. Pure GG domain 1 was concentrated at 10mg/ml and stored at -80°C for further experiments.

“Short” and “long” GG domain 1. The purification was performed in the exact same way as for the GG domain 1.

Full-length CEMIP. Cells expressing CEMIP were slowly thawed in cold lysis buffer (50mM HEPES-NaOH pH 8, 300mM NaCl and 2mM B-mercaptoethanol) containing protease inhibitor cocktail tablets (PIC) (complete EDTA-free; Roche Diagnostics), and 5 units/ml supernuclease (Novagen), subsequently sonicated and centrifuged for 1 h at 20,000g. The soluble fraction was slowly (0.5 ml/min flow rate) applied to two 5 ml StrepTactin Superflow Cartridge (Qiagen) after the cells were lysed as described before with a lysis buffer containing 50mM HEPES-NaOH pH 8, 300mM NaCl, 2mM B-mercaptoethanol, and 5% glycerol. The column was washed with lysis buffer and proteins were eluted with lysis buffer containing 2.5mM desthiobiotin and loaded onto a 5ml HisTrap HP column (GE Healthcare Life Sciences). The samples were applied on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue.

6.4. Protein crystallization and in-situ limited proteolysis

Snap-frozen purified CEMIP GG domain 1 was thawed and concentrated to different ranges of protein concentrations (15, 18 and 22mg/ml) and a range of commercially available crystallization screens (Molecular dimensions) were set up using a Mosquito nanoliter liquid handler (STP Labtech).

To identify a good protease candidate to perform *in-situ* proteolysis, 10µl of GG domain 1 at 22mg/ml was digested at room temperature by four different proteases (elastase, trypsin, subtilisin and GluC) at 1mg/ml. Different dilutions of those proteases were tested (1:50, 1:100, 1:500 and 1:1000) and the digestion lasted for 30 minutes before being quenched by the addition of loading dye. The results of each digestion were loaded on a 4-20% gradient poly-acrylamide mini-PROTEAN TGX precast gel (Bio-Rad) and stained with Coomassie Blue. As we analyzed that GluC was the best candidate, we set up crystallization screens of 100µl drops with a 50:50 dilution of GG domain 1 (22mg/ml) and the crystallization condition supplemented with 2µl of GluC. Each drop was set up using a Mosquito nanoliter liquid handler (STP Labtech). Very small crystals were obtained at room temperature using 20mM MES pH9, 200mM ammonium sulfate and 30% PEG5000 (Molecular dimensions).

6.5. Negative staining

Grid preparation and acquisition. Typically, a drop of 5 μ L of full-length CEMIP at 0,05mg/ml (Genscript) was absorbed on a glow-discharged (GloQube, Quorum, EMS) carbon-coated copper grid (Electron Microscopy Sciences). The grid was washed with two drops of gel filtration buffer and stained with two drops of 1% uranyl-acetate. After absorbing the staining solution with Whatman paper, the grid was air-dried for few minutes before being imaged at room temperature on a Talos L120C G2 (S) Transmission Electron Microscope (TEM) (Thermo Fisher Scientific). The microscope is equipped with a LaB6 filament and is operating at a 120kV accelerating voltage. Images of the complex were recorded at a x120.000 magnification (pixel size 1.998Å) at a defocus value at -1.5 μ M.

Image processing. A dataset of 320 micrographs was collected and all processing steps were performed using CryoSPARC v 3.3.1. The raw micrographs were imported, and particles were directly picked using the blob picker tool. After extracting the particles, two rounds of 2D classifications were performed, and the selected particles were used as a template to pick particles. Two more rounds of 2D classification were performed before analyzing the best 2D classes. A total of selected 19.746 particles were used to generate an ab-initio 3D reconstruction of CEMIP. The obtained volume was analyzed using UCSF ChimeraX.

6.6. CryoEM

Before plunge-freezing procedures, EM grids are glow discharged using a GloQube Plus Glow Discharge system (Quorum, EMS). Typically, EM grids are applied a negative charge by glow-discharging them at 40mA during 20 to 60 seconds for cryoEM and 40 seconds for negative staining.

6.6.1. Graphene oxide-coated EM grids preparation

Graphene-oxide support layers were prepared on Quantifoil R 1.2/1.3 Holey carbon film on a 300-mesh gold grid (Quantifoil micro tools) following a well-established protocol¹⁶¹. Typically, the graphene-oxide solution was diluted 10x before 5 μ l was absorbed on a glow-discharged grid for 30 seconds. The grid was washed with three drops of water before being dried for few minutes. Directly after the grids are ready and the sample can be applied and blotted depending on the tested conditions.

6.6.2. First small data collections

Grid preparation and acquisition. 5uL of full-length CEMIP at 0,05mg/ml were applied onto glow-discharged Lacey carbon films on copper TEM grids (agar scientific) and graphene oxide coated grids as described before. For the Lacey grids, following a 2 to 4s blotting, the grids were plunged-frozen in liquid ethane using an EM GP2 automatic plunge freezer (Leica). For the graphene oxide grids, the protein drop was incubated at least 30 seconds on the grid before being blotted for 1 to 3 seconds followed by their plunge-freezing in liquid ethane using an EM GP2 automatic plunge freezer (Leica) The grids were stored in liquid nitrogen until further use. The grids were analyzed on a equipped with a Falcon 3EC Directed Electron Detector (Thermo Fisher Scientific) operating at 200 keV.

Image processing. A first dataset of 520 movies was collected on a Lacey grid. All processing steps were performed using CryoSPARC v 3.3.1. Raw movies from this first dataset were aligned and dose-weighted using MorionCorr2. The micrographs were imported into CryoSPARC v 3.3.1 and the CTF parameters were estimated using Gctf. Particles were picked with Topaz extract with a Topaz pre-trained model before being extracted. With several rounds of 2D classifications, the best classes were used to select the best 16.960 particles. A second dataset of 590 movies was collected on a graphene oxide grid. Raw movies from this first dataset were imported into CryoSPARC v 3.3.1 and were aligned and dose-weighted using the implanted Patch Motion Correction tool and the CTF parameters were estimated using the implanted Patch CTF tool. Particles were directly picked using the blob picker tool. After extracting the particles, several rounds of 2D classifications were performed, and the selected particles were used as a template to pick particles. Two more rounds of 2D classification were performed before analyzing the 2D classes displaying 25.049 particles. These last ones were used to reconstruct an ab-initio 3D model of CEMIP analyzed using UCSF ChimeraX.

6.6.3. Big data collection

Grid preparation and acquisition. The grid preparation was the same as described in the previous section. The grids were screened on a 200kev Talos Arctica cryo-TEM (Thermo Fisher Scientific) and stored then sent in liquid nitrogen to the Dubochet Center of Imaging facility in Lausanne. The data collections were performed on a 300kv Titan Krios cryo-TEM (Thermo Fisher Scientific) equipped with a Selectris energy filter (Thermo Fisher Scientific).

Movies were recorded on a Falcon 4 Directed Electron Detector (Thermo Fisher Scientific) at a magnification of 150.000x resulting in a super-resolution pixel size of 0.86Å. A total of 50e⁻/Å² was distributed over 44 frames. The movies were collected using the automated data collection software EPU (Thermo Fisher Scientific) with a set of defocus range of -0,8uM to -2uM.

Image processing. A dataset of 16.175 movies was collected. All processing steps were performed using CryoSPARC v 4.4.0. Raw movies from this first dataset were aligned and dose-weighted using MorionCorr2. The micrographs were imported into CryoSPARC v 4.4.0 and the CTF parameters were estimated using the implanted Patch CTF tool. Particles were template-picked using a generated template 2D classes model from the AF2 prediction of CEMIP. After extracting the particles, several rounds of 2D classifications were performed, and 191.420 “good” particles were selected and classified in a final 2D classification.

REFERENCES

1. Hooke R. *Micrographia.*; 1665.
2. Van Leeuwenhoek A. *Letter 35 of March 3.*; 1682.
3. Flemming W. Ueber die ersten Entwicklungserscheinungen am Ei der Teichmuschel. *Arch für Mikroskopische Anat.* 1874;10:257-292.
4. Flemming W. Studien in der Entwicklungsgeschichte der Najaden. *Sitzungsber Kaiserl Akad Wiss.* 1875;71:81-212.
5. Flemming W. Beobachtungen über die Beschaffenheit des Zellkerns. *Arch mikrosk Anat.* 1876;13:693-717.
6. Strasburger EA. *On Cell Formation and Cell Division.*; 1876.
7. Wilson EB. *The Cell in Development and Heredity.*; 1925.
8. Weismann A. *Das Keimplasma; Eine Theorie Der Vererbung.*; 1892.
9. Howard A, Pelc SR. Synthesis of nucleoprotein in bean root cells. *Nature.* 1951;167(4250):599-600. doi:10.1038/167599a0
10. Howard A, Pelc SR. Synthesis of Deoxyribonucleic Acid in Normal and Irradiated Cells and Its Relation to Chromosome Breakage. *Heredity (Edinb).* 1953;6:261-173.
11. Lajtha LG. On the concept of cell cycle. *J Cell Comp Physiol.* 1963;62:SUPPL1:143-5. <https://pubmed.ncbi.nlm.nih.gov/14067857/>
12. Paul MR, Hochwagen A, Ercan S. Condensin action and compaction. *Curr Genet.* 2019;65(2):407-415. doi:10.1007/s00294-018-0899-4
13. Tatsuya H. Chromosome condensation, cohesion and separation. *DNA Repair (Amst).* 2000;69:115-144.
14. Silkworth WT, Nardi IK, Paul R, Mogilner A, Cimini D. Timing of centrosome separation is important for accurate chromosome segregation. *Mol Biol Cell.* 2012;23(3):401-411. doi:10.1091/mbc.E11-02-0095
15. Inoué S, Sato H. Cell motility by labile association of molecules: The nature of mitotic spindle fibers and their role in chromosome movement. *J Gen Physiol.* 1967;(50):259-292. doi:10.1142/9789812790866_0025
16. Mcintosh JR. Mitosis. *Cold Spring Harb Perspect Biol.* 2016;(8):a023218.
17. Chou YH, Ngai KL, Goldman R. The regulation of intermediate filament reorganization in mitosis. *J Biol Chem.* 1991;266(12):7325-7328. doi:10.1016/s0021-9258(20)89448-4
18. Cheeseman IM, Desai A. Molecular architecture of the kinetochore-microtubule

- interface. *Nat Rev Mol Cell Biol.* 2008;9(1):33-46. doi:10.1038/nrm2310
19. Cheeseman IM. The Kinetochore. *Cold Spring Harb Perspect Biol.* 2014;6(7). doi:10.1101/cshperspect.a015826
 20. Shah J V., Cleveland DW. Waiting for Anaphase. *Cell.* 2000;103(7):997-1000. doi:10.1016/s0092-8674(00)00202-6
 21. Rieder CL, Salmon ED. Motile kinetochores and polar ejection forces dictate chromosome position on the vertebrate mitotic spindle. *J Cell Biol.* 1994;124(3):223-233. doi:10.1083/jcb.124.3.223
 22. Hauf S, Waizenegger IC, Peters JM. Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science (80-).* 2001;293(5533):1320-1323. doi:10.1126/science.1061376
 23. Mitchison T, Evans L, Schulze E, Kirschner M. Sites of microtubule assembly and disassembly in the mitotic spindle. *Cell.* 1986;45(4):515-527. doi:10.1016/0092-8674(86)90283-7
 24. Maddox P, Desai A, Oegema K, Mitchison TJ, Salmon ED. Poleward microtubule flux is a major component of spindle dynamics and anaphase A in mitotic *Drosophila* embryos. *Curr Biol.* 2002;12(19):1670-1674. doi:10.1016/S0960-9822(02)01183-1
 25. Sherr CJ. Cancer Cell Cycles. *Science (80-).* 1996;274(5293):1672-1677. doi:10.1126/science.274.5293.1672
 26. Lim S, Kaldis P. Cdks, cyclins and CKIs: Roles beyond cell cycle regulation. *Dev.* 2013;140(15):3079-3093. doi:10.1242/dev.091744
 27. Morgan DO. *The Cell Cycle Principles of Control.*; 2007.
 28. Jeronimo C, Collin P, Robert F. The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain. *J Mol Biol.* 2016;428(12):2607-2622. doi:10.1016/j.jmb.2016.02.006
 29. Guen VJ, Gamble C, Lees JA, Colas P. The awakening of the CDK10/Cyclin M protein kinase. *Oncotarget.* 2017;8(30):50174-50186. doi:10.18632/oncotarget.15024
 30. Pak V, Eifler TT, Jäger S, Krogan NJ, Fujinaga K, Peterlin BM. CDK11 in TREX/THOC regulates HIV mRNA 3' end processing. *Cell Host Microbe.* 2015;18(5):560-570. doi:10.1016/j.chom.2015.10.012
 31. Nurse P, Thuriaux P, Nasmyth K. Genetic control of the cell division cycle in the fission yeast *Schizosaccharomyces pombe*. *MGG Mol Gen Genet.* 1976;146(2):167-178. doi:10.1007/BF00268085
 32. Hartwell LH, Mortimer RK, Culotti J, Culotti M. Genetic control of the cell division

- cycle in yeast: V. Genetic analysis of cdc mutants. *Genetics*. 1973;74(2):267-286.
doi:10.1093/genetics/74.2.267
33. Durkacz B, Carr A, Nurse P. Transcription of the cdc2 cell cycle control gene of the fission yeast *Schizosaccharomyces pombe*. *EMBO J*. 1986;5(2):369-373.
doi:10.1002/j.1460-2075.1986.tb04221.x
 34. Simanis V, Nurse P. The cell cycle control gene cdc2+ of fission yeast encodes a protein kinase potentially regulated by phosphorylation. *Cell*. 1986;45(2):261-268.
doi:10.1016/0092-8674(86)90390-9
 35. Moreno S, Hayles J, Nurse P. Regulation of p34cdc2 protein kinase during mitosis. *Cell*. Published online 1989. doi:10.1016/0092-8674(89)90850-7
 36. Lee MG, Nurse P. Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2. *Nature*. 1987;327(6117):31-35.
doi:10.1038/327031a0
 37. Labbe JC, Leet MG, Nurset P, Picard A, Doree M. Activation at M-phase of a protein kinase encoded by a starfish homologue of the cell cycle control gene cdc2+. *Nature*. 1987;(335):251-254.
 38. Gautier J, Norbury C, Lohka M, Nurse P, Maller J. Purified maturation-promoting factor contains the product of a *Xenopus* homolog of the fission yeast cell cycle control gene cdc2+. *Cell*. 1988;54(3):433-439. doi:10.1016/0092-8674(88)90206-1
 39. Nurse P. Universal control mechanism regulating cell cycle timing of M-phase. *Nature*. 1990;344(April):503-508.
<https://link.springer.com/content/pdf/10.1038/344503a0.pdf>
<https://www.nature.com/articles/344503a0.pdf>
 40. Hanks SK. Homology probing: Identification of cDNA clones encoding members of the protein-serine kinase family. *Proc Natl Acad Sci U S A*. 1987;84(2):388-392.
doi:10.1073/pnas.84.2.388
 41. Matsushime H, Ewen ME, Strom DK, et al. Identification and properties of an atypical catalytic subunit (p34PSK-J3/cdk4) for mammalian D type G1 cyclins. *Cell*. 1992;71(2):323-334. doi:10.1016/0092-8674(92)90360-O
 42. Meyerson M, Harlow E. Identification of G 1 Kinase Activity for cdk6, a Novel Cyclin D Partner. *Mol Cell Biol*. 1994;14(3):2077-2086. doi:10.1128/mcb.14.3.2077-2086.1994
 43. Fesquet D, Labbe JC, Derancourt J, et al. The MO15 gene encodes the catalytic subunit of a protein kinase that activates cdc2 and other cyclin-dependent kinases

- (CDKs) through phosphorylation of Thr161 and its homologues. *EMBO J.* 1993;12(8):3111-3121. doi:10.1002/j.1460-2075.1993.tb05980.x
44. Fisher RP, Morgan DO. A novel cyclin associates with M015/CDK7 to form the CDK-activating kinase. *Cell.* 1994;78(4):713-724. doi:10.1016/0092-8674(94)90535-5
 45. Evans T, Rosenthal ET, Youngblom J, Distel D, Hunt T. Cyclin: A protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell.* 1983;33(2):389-396. doi:10.1016/0092-8674(83)90420-8
 46. Murray AW, Kirschner MW. Cyclin synthesis drives the early embryonic cell cycle. *Nature.* 1989;58(58):99-104. <https://www.nature.com/articles/339275a0>
 47. Murray AW, Solomon MJ, Kirschner MW. The role of cyclin synthesis and degradation in the control of maturation promoting factor activity. *Nature.* 1989;(339):280-286.
 48. Swenson KI, Farrell KM, Ruderman J V. The clam embryo protein cyclin A induces entry into M phase and the resumption of meiosis in *Xenopus* oocytes. *Cell.* 1986;47(6):861-870. doi:10.1016/0092-8674(86)90801-9
 49. Minshull J, Pines J, Golsteyn R, et al. The role of cyclin synthesis, modification and destruction in the control of cell division. *J Cell Sci.* 1989;94(SUPPL. 12):77-97. doi:10.1242/jcs.1989.supplement_12.8
 50. Draetta G, Luca F, Westendorf J, Brizuela L, Ruderman J, Beach D. cdc2 protein kinase is complexed with both cyclin A and B: Evidence for proteolytic inactivation of MPF. *Cell.* 1989;56(5):829-838. doi:10.1016/0092-8674(89)90687-9
 51. Glotzer M, Murray AW, Kirschner MW. Cyclin is degraded by the ubiquitin pathway. *Nature.* Published online 1991. doi:10.1038/349132a0
 52. Sherr CJ, Roberts JM. Living with or without cyclins and cyclin-dependent kinases. *Genes Dev.* 2004;18(22):2699-2711. doi:10.1101/gad.1256504
 53. Malumbres M, Barbacid M. Mammalian cyclin-dependent kinases. *Trends Biochem Sci.* 2005;30(11):630-641. doi:10.1016/j.tibs.2005.09.005
 54. Larochelle S, Merrick KA, Terret ME, et al. Requirements for Cdk7 in the Assembly of Cdk1/Cyclin B and Activation of Cdk2 Revealed by Chemical Genetics in Human Cells. *Mol Cell.* 2007;25(6):839-850. doi:10.1016/j.molcel.2007.02.003
 55. Larochelle S, Batliner J, Gamble MJ, et al. Dichotomous but stringent substrate selection by the dual-function Cdk7 complex revealed by chemical genetics. *Nat Struct Mol Biol.* 2006;13(1):55-62. doi:10.1038/nsmb1028
 56. Russo AA, Jeffrey PD, Pavletich NP. Structural basis of cyclin-dependent kinase

activation by phosphorylation. *Nat Struct Biol*. Published online 1996.

doi:10.1038/nsb0896-696

57. Jeffrey PD, Russo AA, Polyak K, et al. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature*. 1995;376(6538):313-320.
doi:10.1038/376313a0
58. Wood DJ, Endicott JA. Structural insights into the functional diversity of the CDK–cyclin family. *Open Biol*. 2018;8(9). doi:10.1098/rsob.180112
59. Brown NR, Lowe ED, Petri E, Skamnaki V, Antrobus R, Johnson L. Cyclin B and Cyclin A Confer Different Substrate Recognition Properties on CDK2. *Cell Cycle*. 2007;6(11):1350-1359. doi:10.4161/cc.6.11.4278
60. Tatum NJ, Endicott JA. Chatterboxes: the structural and functional diversity of cyclins. *Semin Cell Dev Biol*. 2020;107(March):4-20. doi:10.1016/j.semcdb.2020.04.021
61. Gallant P, Nigg EA. Identification of a novel vertebrate cyclin: Cyclin B3 shares properties with both A- and B-type cyclins. *EMBO J*. 1994;13(3):595-605.
doi:10.1002/j.1460-2075.1994.tb06297.x
62. Petri ET, Errico A, Escobedo L, Hunt T, Basavappa R. The crystal structure of human cyclin B. *Cell Cycle*. Published online 2007. doi:10.4161/cc.6.11.4297
63. Bourne Y, Watson MH, Arvai AS, Bernstein SL, Reed SI, Tainer JA. Crystal structure and mutational analysis of the *Saccharomyces cerevisiae* cell cycle regulatory protein Cks1: implications for domain swapping, anion binding and protein interactions. *Structure*. 2000;8(8):841-850. doi:10.1016/S0969-2126(00)00175-1
64. Yu J, Raia P, Ghent CM, et al. Structural basis of human separase regulation by securin and CDK1–cyclin B1. *Nature*. 2021;596(7870):138-142. doi:10.1038/s41586-021-03764-0
65. Pines J. Cell cycle: Reaching for a role for the Cks proteins. *Curr Biol*. 1996;6(11):1399-1402. doi:10.1016/S0960-9822(96)00741-5
66. Patra D, Wang SX, Kumagai A, Dunphy WG. The *Xenopus* Suc1/Cks protein promotes the phosphorylation of G2/M regulators. *J Biol Chem*. 1999;274(52):36839-36842. doi:10.1074/jbc.274.52.36839
67. Arvai AS, Bourne Y, Mickey MJ, Tainer JA. Crystal structure of the human cell cycle protein CksHs1: Single domain fold with similarity to kinase N-lobe domain. *J Mol Biol*. 1995;249(5):835-842. doi:10.1006/jmbi.1995.0341
68. Kõivomägi M, Örd M, Iofik A, et al. Multisite phosphorylation networks as signal processors for Cdk1. *Nat Struct Mol Biol*. 2013;20(12):1415-1424.

doi:10.1038/nsmb.2706

69. Maiani E, Milletti G, Nazio F, et al. AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity. *Nature*. 2021;592(7856):799-803. doi:10.1038/s41586-021-03422-5
70. Chaikovsky AC, Li C, Jeng EE, et al. The AMBRA1 E3 ligase adaptor regulates the stability of cyclin D. *Nature*. 2021;592(7856):794-798. doi:10.1038/s41586-021-03474-7
71. Simoneschi D, Rona G, Zhou N, et al. CRL4AMBRA1 is a master regulator of D-type cyclins. *Nature*. 2021;592(7856):789-793. doi:10.1038/s41586-021-03445-y
72. Siu KT, Rosner MR, Minella AC. An integrated view of cyclin E function and regulation. *Cell Cycle*. 2012;11(1):57-64.
73. Nakamaya K, Nagahama H, Minamishima YA, et al. Skp2-Mediated Degradation of p27 Regulates Progression into Mitosis. *Dev Cell*. 2004;(6):661-672. doi:10.1007/s13277-015-4151-2
74. Vigneron S, Sundermann L, Labbé JC, et al. Cyclin A-cdk1-Dependent Phosphorylation of Bora Is the Triggering Factor Promoting Mitotic Entry. *Dev Cell*. 2018;45(5):637-650.e7. doi:10.1016/j.devcel.2018.05.005
75. de Castro IJ, Gil RS, Ligammari L, Di Giacinto ML, Vagnarelli P. CDK1 and PLK1 co-ordinate the disassembly and re-assembly of the nuclear envelope in vertebrate mitosis. *Oncotarget*. 2018;9(8):7763-7773. doi:10.18632/oncotarget.23666
76. Fourest-Lieuvin A, Peris L, Gache V, et al. Microtubule Regulation in Mitosis: Tubulin Phosphorylation by the Cyclin-dependent Kinase Cdk1. *Mol Biol Cell*. 2006;17:1041-1050. doi:10.1091/mbc.E05
77. Haneke K, Schott J, Lindner D, et al. CDK1 couples proliferation with protein synthesis. *J Cell Biol*. 2020;219(3). doi:10.1083/jcb.201906147
78. Fujimitsu K, Grimaldi M, Yamano H. Cyclin-dependent kinase 1 – dependent activation of APC/C ubiquitin ligase. Published online 2016.
79. Stemmann O, Zou H, Gerber SA, Gygi SP, Kirschner MW. Dual inhibition of sister chromatid separation at metaphase. *Cell*. 2001;107(6):715-726. doi:10.1016/S0092-8674(01)00603-1
80. Brandeis M, Rosewell I, Carrington M, et al. Cyclin B2-null mice develop normally and are fertile whereas cyclin B1-null mice die in utero. *Proc Natl Acad Sci U S A*. 1998;95(8):4344-4349. doi:10.1073/pnas.95.8.4344
81. Draviam VM, Orrechia S, Lowe M, Pardi R, Pines J. The localization of human cyclins

- B1 and B2 determines CDK1 substrate specificity and neither enzyme requires MEK to disassemble the Golgi apparatus. *J Cell Biol.* 2001;152(5):945-958. doi:10.1083/jcb.152.5.945
82. Geley S, Kramer E, Gieffers C, Gannon J, Peters JM, Hunt T. Anaphase-promoting complex/cyclosome-dependent proteolysis of human cyclin A starts at the beginning of mitosis and is not subject to the spindle assembly checkpoint. *J Cell Biol.* 2001;153(1):137-147. doi:10.1083/jcb.153.1.137
 83. Zhang S, Tischler T, Barford D. Cyclin A2 degradation during the spindle assembly checkpoint requires multiple binding modes to the APC/C. *Nat Commun.* 2019;10(1). doi:10.1038/s41467-019-11833-2
 84. Kabeche L, Compton DA. Cyclin A regulates kinetochore microtubules to promote faithful chromosome segregation. *Nature.* 2013;502(7469):110-113. doi:10.1038/nature12507
 85. Morgan DO. Principles of CDK regulation. *Nature.* 1995;374(6518):131-134. doi:10.1038/374131a0
 86. Amador V, Ge S, Santamaría PG, Guardavaccaro D, Pagano M. APC/CCdc20 Controls the Ubiquitin-Mediated Degradation of p21 in Prometaphase. *Mol Cell.* 2007;27(3):462-473. doi:10.1016/j.molcel.2007.06.013
 87. Smits VAJ, Klompmaier R, Vallenius T, Rijksen G, Mäkelä TP, Medema RH. p21 Inhibits Thr161 phosphorylation of Cdc2 to enforce the G2 DNA damage checkpoint. *J Biol Chem.* 2000;275(39):30638-30643. doi:10.1074/jbc.M005437200
 88. Gorr IH, Boos D, Stemmann O. Mutual inhibition of separase and Cdk1 by two-step complex formation. *Mol Cell.* 2005;19(1):135-141. doi:10.1016/j.molcel.2005.05.022
 89. Ciosk R, Zachariae W, Michaelis C, Shevchenko A, Mann M, Nasmyth K. An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. *Cell.* 1998;93(6):1067-1076. doi:10.1016/S0092-8674(00)81211-8
 90. Sullivan M, Morgan DO. Finishing mitosis, one step at a time. *Nat Rev Mol Cell Biol.* 2007;8(11):894-903. doi:10.1038/nrm2276
 91. Echard A, O'Farrell PH. The degradation of two mitotic cyclins contributes to the timing of cytokinesis. *Curr Biol.* Published online 2003. doi:10.1016/S0960-9822(03)00127-1
 92. Lozano JC, Vergé V, Schatt P, Juengel JL, Peaucellier G. Evolution of cyclin B3 shows an abrupt three-fold size increase, due to the extension of a single exon in

- placental mammals, allowing for new protein-protein interactions. In: *Molecular Biology and Evolution*. ; 2012. doi:10.1093/molbev/mss189
93. Li Y, Wang L, Zhang L, et al. Cyclin b3 is required for metaphase to anaphase transition in oocyte meiosis I. *J Cell Biol*. 2019;218(5):1553-1563. doi:10.1083/jcb.201808088
 94. Xiong Y, Zhang H, Beach D. D type cyclins associate with multiple protein kinases and the DNA replication and repair factor PCNA. *Cell*. 1992;71(3):505-514. doi:10.1016/0092-8674(92)90518-H
 95. Wade Harper J, Adami GR, Wei N, Keyomarsi K, Elledge SJ. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*. 1993;75(4):805-816. doi:10.1016/0092-8674(93)90499-G
 96. Gu Y, Turck CW, Morgan DO. Inhibition of CDK2 activity in vivo by an associated 20K regulatory subunit. *Nature*. 1993;366(6456):707-710. doi:10.1038/366707a0
 97. Xiong V, Hannon GJ, Zhang H, Beach D. p21 is a universal inhibitor of cyclin kinases. *Nature*. 1993;366(December):701-704.
 98. El-Deiry WS, Tokino T, Velculescu VE, et al. WAF1, a potential mediator of p53 tumor suppression. *Cell*. 1993;75(4):817-825. doi:10.1016/0092-8674(93)90500-P
 99. Abbas T, Dutta A. P21 in cancer: Intricate networks and multiple activities. *Nat Rev Cancer*. 2009;9(6):400-414. doi:10.1038/nrc2657
 100. Moldovan GL, Pfander B, Jentsch S. PCNA, the Maestro of the Replication Fork. *Cell*. 2007;129(4):665-679. doi:10.1016/j.cell.2007.05.003
 101. Coqueret O, Gascan H. Functional interaction of STAT3 transcription factor with the cell cycle inhibitor p21(WAF1/CIP1/SDI1). *J Biol Chem*. 2000;275(25):18794-18800. doi:10.1074/jbc.M001601200
 102. Zhang Y, Fujita N, Tsuruo T. Caspase-mediated cleavage of p21(Waf1/Cip1) converts cancer cells from growth arrest to undergoing apoptosis. *Oncogene*. 1999;18(5):1131-1138. doi:10.1038/sj.onc.1202426
 103. Kreis NN, Louwen F, Yuan J. The multifaceted p21 (Cip1/Waf1/CDKN1A) in cell differentiation, migration and cancer therapy. *Cancers (Basel)*. 2019;11(9):14-16. doi:10.3390/cancers11091220
 104. Su JY, Rempel RE, Erikson E, Maller JL. Cloning and characterization of the *Xenopus* cyclin-dependent kinase inhibitor p27(XIC1). *Proc Natl Acad Sci U S A*. 1995;92(22):10187-10191. doi:10.1073/pnas.92.22.10187
 105. Lane ME, Sauer K, Wallace K, Jan YN, Lehner CF, Vaessin H. Dacapo, a cyclin-

- dependent kinase inhibitor, stops cell proliferation during *Drosophila* development. *Cell*. 1996;87(7):1225-1235. doi:10.1016/S0092-8674(00)81818-8
106. Hong Y, Roy R, Ambros V. Developmental regulation of a cyclin-dependent kinase inhibitor controls postembryonic cell cycle progression in *Caenorhabditis elegans*. *Development*. 1998;125(18):3585-3597. doi:10.1242/dev.125.18.3585
 107. Wang, H., Fowke, L. & Crosby WA. A plant cyclin-dependent kinase inhibitor gene. *Nature*. 1997;386:451.
 108. Macleod KF, Sherry N, Hannon G, et al. p53-dependent and independent expression of p21 during cell growth, differentiation, and DNA damage. *Genes Dev*. 1995;9(8):9335-9344.
 109. Martín-Caballero J, Flores JM, García-Palencia P, Serrano M. Tumor Susceptibility of p21 Waf1/Cip1-deficient Mice 1. *Cancer Res*. 2001;61:6234-6238.
 110. Asada M, Yamada T, Ichijo H, et al. Apoptosis inhibitory activity of cytoplasmic p21(Cip1/WAF1) in monocytic differentiation. *EMBO J*. 1999;18(5):1223-1234. doi:10.1093/emboj/18.5.1223
 111. Liu Y, Yeh N, Zhu XH, et al. Somatic cell type specific gene transfer reveals a tumor-promoting function for p21Waf1/Cip1. *EMBO J*. 2007;26(22):4683-4693. doi:10.1038/sj.emboj.7601886
 112. Balbín M, Hannon GJ, Pendás AM, et al. Functional analysis of a p21(WAF1,CIP1,SDI1) mutant (Arg94 → Trp) identified in a human breast carcinoma. Evidence that the mutation impairs the ability of p21 to inhibit cyclin-dependent kinases. *J Biol Chem*. 1996;271(26):15782-15786. doi:10.1074/jbc.271.26.15782
 113. Blagosklonny M V, Wu GS, Omura S, El-deiry WS. Proteasome-Dependent Regulation of p21 WAF1 / CIP1 Expression several regulatory proteins . The degradation of these short-lived proteins by the proteasome of p21 expression . We found that LC induced p21 , by both transcriptional and posttranscrip- p21 I. *Biochem Biophys Res Commun*. 1996;569:564-569.
 114. Bornstein G, Bloom J, Sitry-Shevah D, Nakayama K, Pagano M, Hershko A. Role of the SCFSkp2 ubiquitin ligase in the degradation of p21Cip1 in S phase. *J Biol Chem*. 2003;278(28):25752-25757. doi:10.1074/jbc.M301774200
 115. Abbas T, Sivaprasad U, Terai K, Amador V, Pagano M, Dutta A. PCNA-dependent regulation of p21 ubiquitylation and degradation via the CRL4Cdt2 ubiquitin ligase complex. *Genes Dev*. 2008;22(18):2496-2506. doi:10.1101/gad.1676108
 116. Harper JW, Elledge SJ, Keyomarsi K, et al. Inhibition of cyclin-dependent kinases by

- p21. *Mol Biol Cell*. 1995;6(4):387-400. doi:10.1091/mbc.6.4.387
117. Aprelikova O, Xiong Y, Liu ET. Both p16 and p21 families of cyclin-dependent kinase (CDK) inhibitors block the phosphorylation of cyclin-dependent kinases by the CDK-activating kinase. *J Biol Chem*. 1995;270(31):18195-18197. doi:10.1074/jbc.270.31.18195
 118. Cheng M, Olivier P, Diehl JA, et al. The p21(Cip1) and p27(Kip1) CDK “inhibitors” are essential activators of cyclin D-dependent kinases in murine fibroblasts. *EMBO J*. 1999;18(6):1571-1583. doi:10.1093/emboj/18.6.1571
 119. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc Natl Acad Sci U S A*. 1996;93(21):11504-11509. doi:10.1073/pnas.93.21.11504
 120. Kreis NN, Louwen F, Yuan J. Less understood issues: P21Cip1 in mitosis and its therapeutic potential. *Oncogene*. 2014;34(14):1758-1767. doi:10.1038/onc.2014.133
 121. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. *Genome Inform Ser Workshop Genome Inform*. 2000;11(May 2014):161-171.
 122. Junjie Chen, Peter K. Jackson MWK& AD. Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA. *Nature*. 1995;374:386-388.
 123. Lacy ER, Filippov I, Lewis WS, et al. p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat Struct Mol Biol*. 2004;11(4):358-364. doi:10.1038/nsmb746
 124. Russo AA, Jeffrey PD, Patten AK, Massagué J, Pavletich NP. Crystal structure of the p27(Kip1) cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature*. 1996;382(6589):325-331. doi:10.1038/382325a0
 125. Wang Y, Fisher JC, Mathew R, et al. Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21. *Nat Chem Biol*. 2011;7(4):214-221. doi:10.1038/nchembio.536
 126. Chen J, Saha P, Kornbluth S, Dynlacht BD, Dutta A. Cyclin-Binding Motifs Are Essential for the Function of p21 CIP1 . *Mol Cell Biol*. 1996;16(9):4673-4682. doi:10.1128/mcb.16.9.4673
 127. Wohlschlegel JA, Dwyer BT, Takeda DY, Dutta A. Mutational Analysis of the Cy Motif from p21 Reveals Sequence Degeneracy and Specificity for Different Cyclin-Dependent Kinases. *Mol Cell Biol*. 2001;21(15):4868-4874.

doi:10.1128/mcb.21.15.4868-4874.2001

128. Dash BC, El-Deiry WS. Phosphorylation of p21 in G₂/M Promotes Cyclin B-Cdc2 Kinase Activity . *Mol Cell Biol*. 2005;25(8):3364-3387. doi:10.1128/mcb.25.8.3364-3387.2005
129. Gharbi SI, Pelletier LA, Espada A, et al. Crystal structure of active CDK4-cyclin D and mechanistic basis for abemaciclib efficacy. *npj Breast Cancer*. 2022;8(1):126. doi:10.1038/s41523-022-00494-y
130. Rowland RJ, Heath R, Maskell D, et al. Cryo-EM structure of SKP1-SKP2-CKS1 in complex with CDK2-cyclin A-p27KIP1. *Sci Rep*. 2023;13(1):1-11. doi:10.1038/s41598-023-37609-9
131. Guiley KZ, Stevenson JW, Lou K, et al. p27 allosterically activates cyclin-dependent kinase 4 and antagonizes palbociclib inhibition. *Science (80-)*. 2019;366(6471). doi:10.1126/science.aaw2106
132. Gulbis JM, Kelman Z, Hurwitz J, O'Donnell M, Kuriyan J. Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell*. 1996;87(2):297-306. doi:10.1016/S0092-8674(00)81347-1
133. Rodríguez-Vilarrupla A, Díaz C, Canela N, Rahn HP, Bachs O, Agell N. Identification of the nuclear localization signal of p21cip1 and consequences of its mutation on cell proliferation. *FEBS Lett*. 2002;531(2):319-323. doi:10.1016/S0014-5793(02)03549-4
134. Zhu H, Nie L, Maki CG. Cdk2-dependent inhibition of p21 stability via a C-terminal cyclin-binding motif. *J Biol Chem*. 2005;280(32):29282-29288. doi:10.1074/jbc.M407352200
135. Bunz F, Dutriaux A, Lengauer C, et al. Requirement for p53 and p21 to sustain G₂ arrest after DNA damage. *Science (80-)*. 1998;282(5393):1497-1501. doi:10.1126/science.282.5393.1497
136. Peng CY, Graves PR, Thoma RS, Wu Z, Shaw AS, Piwnicka-Worms H. Mitotic and G₂ checkpoint control: Regulation of 14-3-3 protein binding by phosphorylation of Cdc25c on serine-216. *Science (80-)*. 1997;277(5331):1501-1505. doi:10.1126/science.277.5331.1501
137. Lopez-Girona A, Furnari B, Mondesert O, Russell P. Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. *Nature*. 1999;397(6715):172-175. doi:10.1038/16488
138. Kreis NN, Friemel A, Zimmer B, et al. Mitotic p21Cip1/CDKN1A is regulated by cyclin-dependent kinase 1 phosphorylation. *Oncotarget*. 2016;7(31):50215-50228.

doi:10.18632/oncotarget.10330

139. Ubersax JA, Woodbury EL, Quang PN, et al. Targets of the cyclin-dependent kinase Cdk1. *Nature*. 2003;425(6960):859-864. doi:10.1038/nature02062
140. Colleoni B, Paternot S, Pita JM, et al. JNKs function as CDK4-activating kinases by phosphorylating CDK4 and p21. *Oncogene*. 2017;36(30):4349-4361. doi:10.1038/onc.2017.7
141. Child ES, Mann DJ. The intricacies of p21 phosphorylation: Protein/protein interactions, subcellular localization and stability. *Cell Cycle*. 2006;5(12):1313-1319. doi:10.4161/cc.5.12.2863
142. Henriques J, Lindorff-Larsen K. Protein Dynamics Enables Phosphorylation of Buried Residues in Cdk2/Cyclin-A-Bound p27. *Biophys J*. 2020;119(10):2010-2018. doi:10.1016/j.bpj.2020.06.040
143. Rath SL, Senapati S. Mechanism of p27 Unfolding for CDK2 Reactivation. *Sci Rep*. 2016;6(April):1-12. doi:10.1038/srep26450
144. Tsytlonok M, Sanabria H, Wang Y, et al. Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation. *Nat Commun*. 2019;10(1). doi:10.1038/s41467-019-09446-w
145. Swadling JB, Warnecke T, Morris KL, Barr AR. Conserved Cdk inhibitors show unique structural responses to tyrosine phosphorylation. *Biophys J*. 2022;121(12):2312-2329. doi:10.1016/j.bpj.2022.05.024
146. Huang Y, Yoon MK, Otieno S, Lelli M, Kriwacki RW. The activity and stability of the intrinsically disordered Cip/Kip protein family are regulated by non-receptor tyrosine kinases. *J Mol Biol*. 2015;427(2):371-386. doi:10.1016/j.jmb.2014.11.011
147. Asfaha JB, Örd M, Carlson CR, Faustova I, Loog M, Morgan DO. Multisite phosphorylation by Cdk1 initiates delayed negative feedback to control mitotic transcription. *Curr Biol*. 2022;32(1):256-263.e4. doi:10.1016/j.cub.2021.11.001
148. Huis in 't Veld PJ, Wohlgemuth S, Koerner C, Müller F, Janning P, Musacchio A. Reconstitution and use of highly active human CDK1:Cyclin-B:CKS1 complexes. *Protein Sci*. 2022;31(2):528-537. doi:10.1002/pro.4233
149. Ball KL. p21: structure and functions associated with cyclin-CDK binding. *Prog Cell Cycle Res*. 1997;3:125-134. doi:10.1007/978-1-4615-5371-7_10
150. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2
151. Terwilliger TC, Leibschnner DL, Croll T, et al. AlphaFold predictions are valuable

hypotheses, and accelerate but do not replace experimental structure determination. *bioRxiv*. Published online 2023:2022.11.21.517405. doi:10.1038/s41592-023-02087-4

152. Scarff CA, Fuller MJG, Thompson RF, Iadanza MG. Variations on negative stain electron microscopy methods: Tools for tackling challenging systems. *J Vis Exp*. 2018;2018(132):1-8. doi:10.3791/57199
153. Adrian M, Dubochet J, Lepault J, McDowell AW. Cryo-electron microscopy of viruses. *Nature*. 1984;308(5954):32-36. doi:10.1038/308032a0
154. Cheng Y. Single-particle cryo-EM—How did it get here and where will it go. *Science (80-)*. 2018;361(6405):876-880. doi:10.1126/science.aat4346
155. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol*. 1990;213(4):899-929. doi:10.1016/S0022-2836(05)80271-2
156. Frank J. Averaging of low exposure electron micrographs of non-periodic objects. *Ultramicroscopy*. 1975;1(2):159-162. doi:10.1016/S0304-3991(75)80020-9
157. Koning RI, Vader H, van Nugteren M, et al. Automated vitrification of cryo-EM samples with controllable sample thickness using suction and real-time optical inspection. *Nat Commun*. 2022;13(1):2985. doi:10.1038/s41467-022-30562-7
158. Vilas JL, Carazo JM, Sorzano COS. Emerging Themes in CryoEM—Single Particle Analysis Image Processing. *Chem Rev*. 2022;122(17):13915-13951. doi:10.1021/acs.chemrev.1c00850
159. Li X, Mooney P, Zheng S, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*. 2013;10(6):584-590. doi:10.1038/nmeth.2472
160. Wu M, Lander GC. How low can we go? Structure determination of small biological complexes using single-particle cryo-EM. *Curr Opin Struct Biol*. 2020;64:9-16. doi:10.1016/j.sbi.2020.05.007
161. Pantelic RS, Meyer JC, Kaiser U, Baumeister W, Plitzko JM. Graphene oxide: A substrate for optimizing preparations of frozen-hydrated samples. *J Struct Biol*. 2010;170(1):152-156. doi:10.1016/j.jsb.2009.12.020
162. Wang F, Yu Z, Betegon M, et al. Amino and PEG-amino graphene oxide grids enrich and protect samples for high-resolution single particle cryo-electron microscopy. *J Struct Biol*. 2020;209(2):107437. doi:10.1016/j.jsb.2019.107437
163. Cookis T, Sauer P, Poepsel S, et al. Streptavidin-Affinity Grid Fabrication for Cryo-

- Electron Microscopy Sample Preparation. *J Vis Exp.* 2023;(202). doi:10.3791/66197
164. Kim Y-M, Kang J-S, Kim J-S, Jeung J-M, Lee J-Y, Kim Y-J. Ultrathin Carbon Support Films for High-Resolution Electron Microscopy of Nanoparticles. *Microsc Microanal.* 2007;13(4):285-290. doi:10.1017/S1431927607070250
165. Bepler T, Morin A, Rapp M, et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods.* 2019;16(11):1153-1160. doi:10.1038/s41592-019-0575-8
166. Rowland RJ, Korolchuk S, Salamina M, et al. Cryo-EM structure of the CDK2-cyclin A-CDC25A Complex. *bioRxiv.* Published online 2023:1-36.
167. Greber BJ, Perez-Bertoldi JM, Lim K, Iavarone AT, Toso DB, Nogales E. The cryoelectron microscopy structure of the human CDK-activating kinase. *Proc Natl Acad Sci.* 2020;117(37):22849-22857. doi:10.1073/pnas.2009627117
168. Russo CJ, Passmore LA. Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. *J Struct Biol.* 2016;193(1):33-44. doi:10.1016/j.jsb.2015.11.006
169. Naydenova K, Jia P, Russo CJ. Cryo-EM with sub-1 Å specimen movement. *Science (80-).* 2020;370(6513):223-226. doi:10.1126/science.abb7927
170. Britt HM, Cragolini T, Thalassinou K. Integration of Mass Spectrometry Data for Structural Biology. *Chem Rev.* 2022;122(8):7952-7986. doi:10.1021/acs.chemrev.1c00356
171. Engen JR, Komives EA. Complementarity of Hydrogen/Deuterium Exchange Mass Spectrometry and Cryo-Electron Microscopy. *Trends Biochem Sci.* 2020;45(10):906-918. doi:10.1016/j.tibs.2020.05.005
172. Goda T, Funakoshi M, Suhara H, Nishimoto T, Kobayashi H. The N-terminal Helix of Xenopus Cyclins A and B Contributes to Binding Specificity of the Cyclin-CDK Complex. *J Biol Chem.* 2001;276(18):15415-15422. doi:10.1074/jbc.M011101200
173. Brown NR, Korolchuk S, Martin MP, et al. CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nat Commun.* 2015;6(1):6769. doi:10.1038/ncomms7769
174. Goda T, Ishii T, Nakajo N, Sagata N, Kobayashi H. The RRASK motif in Xenopus cyclin B2 is required for the substrate recognition of Cdc25C by the cyclin B-Cdc2 complex. *J Biol Chem.* 2003;278(21):19032-19037. doi:10.1074/jbc.M300210200
175. Song H, Hanlon N, Brown NR, Noble MEM, Johnson LN, Barford D. Phosphoprotein-protein interactions revealed by the crystal structure of kinase-associated

- phosphatase in complex with phosphoCDK2. *Mol Cell*. 2001;7(3):615-626.
doi:10.1016/S1097-2765(01)00208-8
176. Gould KL, Nurse P. Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. *Nature*. 1989;342(6245):39-45.
doi:10.1038/342039a0
 177. Norbury C, Blow J, Nurse P. Regulatory phosphorylation of the p34cdc2 protein kinase in vertebrates. *EMBO J*. 1991;10(11):3321-3329. doi:10.1002/j.1460-2075.1991.tb04896.x
 178. Contreras A, Hale TK, Stenoien DL, Rosen JM, Mancini MA, Herrera RE. The Dynamic Mobility of Histone H1 Is Regulated by Cyclin/CDK Phosphorylation. *Mol Cell Biol*. 2003;23(23):8626-8636. doi:10.1128/MCB.23.23.8626-8636.2003
 179. al-Rawi A, Kaye E, Korolchuk S, Endicott JA, Ly T. Cyclin A and Cks1 promote kinase consensus switching to non-proline-directed CDK1 phosphorylation. *Cell Rep*. 2023;42(3):112139. doi:10.1016/j.celrep.2023.112139
 180. Kimanius D, Jamali K, Wilkinson ME, et al. Data-driven regularisation lowers the size barrier of cryo-EM structure determination. *bioRxiv*. Published online 2023:2023.10.23.563586.
<https://www.biorxiv.org/content/10.1101/2023.10.23.563586v1%0Ahttps://www.biorxiv.org/content/10.1101/2023.10.23.563586v1.abstract>
 181. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci*. 2008;33(1):2-8.
doi:10.1016/j.tibs.2007.10.003
 182. Banerjee PR, Holehouse AS, Kriwacki R, et al. Dissecting the biophysics and biology of intrinsically disordered proteins. *Trends Biochem Sci*. Published online November 2023. doi:10.1016/j.tibs.2023.10.002
 183. Ray A, James MK, Larochelle S, Fisher RP, Blain SW. p27 Kip1 Inhibits Cyclin D-Cyclin-Dependent Kinase 4 by Two Independent Modes. *Mol Cell Biol*. 2009;29(4):986-999. doi:10.1128/MCB.00898-08
 184. Grimm M, Wang Y, Mund T, et al. Cdk-Inhibitory Activity and Stability of p27 Are Directly Regulated by Oncogenic Tyrosine Kinases. *Cell*. 2007;128(2):269-280.
doi:10.1016/j.cell.2006.11.047
 185. Huang Y, Yoon M-K, Otieno S, Lelli M, Kriwacki RW. The Activity and Stability of the Intrinsically Disordered Cip/Kip Protein Family Are Regulated by Non-Receptor Tyrosine Kinases. *J Mol Biol*. 2015;427(2):371-386. doi:10.1016/j.jmb.2014.11.011

186. Sivakolundu SG, Bashford D, Kriwacki RW. Disordered p27Kip1 Exhibits Intrinsic Structure Resembling the Cdk2/Cyclin A-bound Conformation. *J Mol Biol.* 2005;353(5):1118-1128. doi:10.1016/j.jmb.2005.08.074
187. Basu S, Roberts EL, Jones AW, Swaffer MP, Snijders AP, Nurse P. The Hydrophobic Patch Directs Cyclin B to Centrosomes to Promote Global CDK Phosphorylation at Mitosis. *Curr Biol.* 2020;30(5):883-892.e4. doi:10.1016/j.cub.2019.12.053
188. Brown NR, Noble MEM, Endicott JA, Johnson LN. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol.* 1999;1(7):438-443. doi:10.1038/15674
189. Ohta T, Okamoto K, Isohashi F, et al. T-loop deletion of CDC2 from breast cancer tissues eliminates binding to cyclin B1 and cyclin-dependent kinase inhibitor p21. *Cancer Res.* 1998;58(6):1095-1098.
190. Schulze-Gahmen U, Kim SH. Structural basis for CDK6 activation by a virus-encoded cyclin. *Nat Struct Biol.* 2002;9(3):177-181. doi:10.1038/nsb756
191. Lu Z, Hunter T. Ubiquitylation and proteasomal degradation of the p21 Cip1 , p27 Kip1 and p57 Kip2 CDK inhibitors. *Cell Cycle.* 2010;9(12):2342-2352. doi:10.4161/cc.9.12.11988
192. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022;19(6):679-682. doi:10.1038/s41592-022-01488-1
193. Pettersen EF, Goddard TD, Huang CC, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 2021;30(1):70-82. doi:10.1002/pro.3943
194. Wang H, Lo W-T, Vujičić Žagar A, et al. Autoregulation of Class II Alpha PI3K Activity by Its Lipid-Binding PX-C2 Domain Module. *Mol Cell.* 2018;71(2):343-351.e4. doi:10.1016/j.molcel.2018.06.042
195. Shevchenko A, Wilm M, Vorm O, Mann M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal Chem.* 1996;68(5):850-858. doi:10.1021/ac950914h
196. Schilling B, Rardin MJ, MacLean BX, et al. Platform-independent and Label-free Quantitation of Proteomic Data Using MS1 Extracted Ion Chromatograms in Skyline. *Mol Cell Proteomics.* 2012;11(5):202-214. doi:10.1074/mcp.M112.017707
197. Spataro S, Guerra C, Cavalli A, et al. CEMIP(HYBID, KIAA1199): structure, function and expression in health and disease. *FEBS J.* 2023;290(16):3946-3962.

doi:10.1111/febs.16600

198. Yoshida H, Nagaoka A, Kusaka-Kikushima A, et al. KIAA1199, a deafness gene of unknown function, is a new hyaluronan binding protein involved in hyaluronan depolymerization. *Proc Natl Acad Sci*. 2013;110(14):5612-5617. doi:10.1073/pnas.1215432110
199. Bayer IS. Hyaluronic Acid and Controlled Release: A Review. *Molecules*. 2020;25(11):2649. doi:10.3390/molecules25112649
200. Vigetti D, Viola M, Karousou E, De Luca G, Passi A. Metabolic control of hyaluronan synthases. *Matrix Biol*. 2014;35:8-13. doi:10.1016/j.matbio.2013.10.002
201. Fallacara A, Baldini E, Manfredini S, Vertuani S. Hyaluronic Acid in the Third Millennium. *Polymers (Basel)*. 2018;10(7):701. doi:10.3390/polym10070701
202. Stern R, Jedrzejewski MJ. Hyaluronidases: Their Genomics, Structures, and Mechanisms of Action. *Chem Rev*. 2006;106(3):818-839. doi:10.1021/cr050247k
203. Zheng X, Wang B, Tang X, et al. Absorption, metabolism, and functions of hyaluronic acid and its therapeutic prospects in combination with microorganisms: A review. *Carbohydr Polym*. 2023;299:120153. doi:10.1016/j.carbpol.2022.120153
204. Laurent TC, Fraser JRE. Hyaluronan. *FASEB J*. 1992;6(7):2397-2404. doi:10.1096/fasebj.6.7.1563592
205. Yoshida H, Nagaoka A, Komiya A, et al. Reduction of hyaluronan and increased expression of HYBID (alias CEMIP and KIAA1199) correlate with clinical symptoms in photoaged skin. *Br J Dermatol*. 2018;179(1):136-144. doi:10.1111/bjd.16335
206. Deroyer C, Charlier E, Neuville S, et al. CEMIP (KIAA1199) induces a fibrosis-like process in osteoarthritic chondrocytes. *Cell Death Dis*. 2019;10(2):103. doi:10.1038/s41419-019-1377-8
207. Shimoda M, Yoshida H, Mizuno S, et al. Hyaluronan-Binding Protein Involved in Hyaluronan Depolymerization Controls Endochondral Ossification through Hyaluronan Metabolism. *Am J Pathol*. 2017;187(5):1162-1176. doi:10.1016/j.ajpath.2017.01.005
208. Evensen NA, Kuscu C, Nguyen H-L, et al. Unraveling the Role of KIAA1199, a Novel Endoplasmic Reticulum Protein, in Cancer Cell Migration. *JNCI J Natl Cancer Inst*. 2013;105(18):1402-1416. doi:10.1093/jnci/djt224
209. Yoshida H, Nagaoka A, Nakamura S, Tobiishi M, Sugiyama Y, Inoue S. N-terminal signal sequence is required for cellular trafficking and hyaluronan-depolymerization of KIAA1199. *FEBS Lett*. 2014;588(1):111-116. doi:10.1016/j.febslet.2013.11.017

210. Guo J, Cheng H, Zhao S, Yu L. GG: A domain involved in phage LTF apparatus and implicated in human MEB and non-syndromic hearing loss diseases. *FEBS Lett.* 2006;580(2):581-584. doi:10.1016/j.febslet.2005.12.076
211. Abe S, Usami S, Nakamura Y. Mutations in the gene encoding KIAA1199 protein, an inner-ear protein expressed in Deiters' cells and the fibrocytes, as the cause of nonsyndromic hearing loss. *J Hum Genet.* 2003;48(11):564-570. doi:10.1007/s10038-003-0079-2
212. Jenkins J, Pickersgill R. The architecture of parallel β -helices and related folds. *Prog Biophys Mol Biol.* 2001;77(2):111-175. doi:10.1016/S0079-6107(01)00013-X
213. Dong A, Xu X, Edwards AM. In situ proteolysis for protein crystallization and structure determination. *Nat Methods.* 2007;4(12):1019-1021. doi:10.1038/nmeth1118
214. Yamamoto H, Tobisawa Y, Inubushi T, Irie F, Ohyama C, Yamaguchi Y. A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. *J Biol Chem.* 2017;292(18):7304-7313. doi:10.1074/jbc.M116.770149
215. Niu M, McGrath M, Sammon D, et al. Structure of the transmembrane protein 2 (TMEM2) ectodomain and its lack of hyaluronidase activity. *Wellcome Open Res.* 2023;8:76. doi:10.12688/wellcomeopenres.18937.1
216. Schmaus A, Rothley M, Schreiber C, et al. Sulfated hyaluronic acid inhibits the hyaluronidase CEMIP and regulates the HA metabolism, proliferation and differentiation of fibroblasts. *Matrix Biol.* 2022;109:173-191. doi:10.1016/j.matbio.2022.04.001
217. Shostak K, Zhang X, Hubert P, et al. NF- κ B-induced KIAA1199 promotes survival through EGFR signalling. *Nat Commun.* 2014;5(1):5232. doi:10.1038/ncomms6232
218. Wu S, Avila-Sakar A, Kim J, et al. Fabs Enable Single Particle cryoEM Studies of Small Proteins. *Structure.* 2012;20(4):582-592. doi:10.1016/j.str.2012.02.017