STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF HUMAN TRANSCRIPTION-REGULATING KINASES

DISSERTATION

zur Erlangung des Doktorgrades (Dr. rer. nat.)

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

INES H. KALTHEUNER

aus Wermelskirchen

Bonn, 2023

Diese Dissertation wurde angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

1. Gutachter: PROF. DR. MATTHIAS GEYER

2. Gutachter: PROF. DR. OLIVER GRUSS

Tag der Promotion: 26.01.2024
Erscheinungsjahr: 2024

LIST OF PUBLICATIONS

Parts of this thesis are published in:

Kaltheuner IH, Devlin JR, Schmitz M, Anand K, Moecking J, Hageluken G, Johnstone RW, Geyer M. Novel transcriptional CDK/Cyclin pairs. *In preparation*.

Insco ML, Abraham BJ, Dubbury SJ, **Kaltheuner IH**, Dust S, Wu C, Chen KY, Liu D, Bellaousov S, Cox AM, Martin BJE, Zhang T, Ludwig CG, Fabo T, Modhurima R, Esgdaille DE, Henriques T, Brown KM, Chanock SJ, Geyer M, Adelman K, Sharp PA, Young RA, Boutz PL, Zon LI. Oncogenic CDK13 mutations impede nuclear RNA surveillance. *Science* 2023 Apr 21;380(6642):eabn7625. doi: 10.1126/science.abn7625.

Kaltheuner IH, Anand K, Moecking J, Düster R, Wang J, Gray NS, Geyer M. Abemaciclib is a potent inhibitor of DYRK1A and HIP kinases involved in transcriptional regulation. *Nat Commun*. 2021 Nov 16;12(1):6607. doi: 10.1038/s41467-021-26935-z.

Additional Publications:

Jenster LM, Lange KE, Normann S, Hemdt Av, Wuerth JD, Schiffelers LDJ, Tesfamariam YM, Gohr FN, Klein L, **Kaltheuner IH**, Ebner S, Lapp DJ, Mayer J, Moecking J, Ploegh HL, Latz E, Meissner F, Geyer M, Kümmerer BM, Schmidt FI. P38 kinases mediate NLRP1 inflammasome activation after ribotoxic stress response and virus infection. J Exp Med 2023 Jan 02; 220 (1); e20220837. doi: 10.1084/jem.20220837.

Baßler K, Fujii W, Kapellos TS, Dudkin E, Reusch N, Horne A, Reiz B, Luecken MD, Osei-Sarpong C, Warnat-Herresthal S, Bonaguro L, Schulte-Schrepping J, Wagner A, Günther P, Pizarro C, Schreiber T, Knoll R, Holsten L, Kröger C, De Domenico E, Becker M, Händler K, Wohnhaas CT, Baumgartner F, Köhler M, Theis H, Kraut M, Wadsworth MH 2nd, Hughes TK, Ferreira HJ, Hinkley E, **Kaltheuner IH**, Geyer M, Thiele C, Shalek AK, Feißt A, Thomas D, Dickten H, Beyer M, Baum P, Yosef N, Aschenbrenner AC, Ulas T, Hasenauer J, Theis FJ, Skowasch D, Schultze JL. Alveolar macrophages in early stage COPD show functional deviations with properties of impaired immune activation. Front Immunol. 2022 Jul 28;13:917232. doi: 10.3389/fimmu.2022.917232.

Ludwig KU, Schmithausen RM, Li D, Jacobs ML, Hollstein R, Blumenstock K, Liebing J, Słabicki M, Ben-Shmuel A, Israeli O, Weiss S, Ebert TS, Paran N, Rüdiger W, Wilbring G, Feldman D, Lippke B, Ishorst N, Hochfeld LM, Beins EC, **Kaltheuner IH**, Schmitz M, Wöhler A, Döhla M, Sib E, Jentzsch M, Borrajo JD, Strecker J, Reinhardt J, Cleary B, Geyer M, Hölzel M, Macrae R, Nöthen MM, Hoffmann P, Exner M, Regev A, Zhang F, Schmid-Burgk JL. LAMP-Seq enables sensitive, multiplexed COVID-19 diagnostics using molecular barcoding. Nat Biotechnol. 2021 Jun 29. doi: 10.1038/s41587-021-00966-9.

Jiang B*, Jiang J*, **Kaltheuner IH***, Iniguez AB, Anand K, Ferguson FM, Ficarro SB, Seong BKA, Greifenberg AK, Dust S, Kwiatkowski NP, Marto JA, Stegmaier K, Zhang T, Geyer M, Gray NS. Structure-activity relationship study of THZ531 derivatives enables the discovery of BSJ-01-175 as a dual CDK12/13 covalent inhibitor with efficacy in Ewing sarcoma. Eur J Med Chem. 2021 Oct 5;221:113481. doi: 10.1016/j.ejmech.2021.113481.
* equal author contribution

Jiang B, Gao Y, Che J, Lu W, **Kaltheuner IH**, Dries R, Kalocsay M, Berberich MJ, Jiang J, You I, Kwiatkowski N, Riching KM, Daniels DL, Sorger PK, Geyer M, Zhang T, Gray NS. Discovery and resistance mechanism of a selective CDK12 degrader. Nat Chem Biol. 2021 Jun;17(6):675-683. doi: 10.1038/s41589-021-00765-y.

Düster R*, **Kaltheuner IH***, Schmitz M, Geyer M. 1,6-Hexanediol, commonly used to dissolve liquid-liquid phase separated condensates, directly impairs kinase and phosphatase activities. J Biol Chem. 2021 Jan-Jun;296:100260. doi: 10.1016/j.jbc.2021.100260.
* equal author contribution

Mayor-Ruiz C, Bauer S, Brand M, Kozicka Z, Siklos M, Imrichova H, **Kaltheuner IH**, Hahn E, Seiler K, Koren A, Petzold G, Fellner M, Bock C, Müller AC, Zuber J, Geyer M, Thomä NH, Kubicek S, Winter GE. Rational discovery of molecular glue degraders via scalable chemical profiling. Nat Chem Biol. 2020 Nov;16(11):1199-1207. doi: 10.1038/s41589-020-0594-x.

Rimel JK, Poss ZC, Erickson B, Maas ZL, Ebmeier CC, Johnson JL, Decker TM, Yaron TM, Bradley MJ, Hamman KB, Hu S, Malojcic G, Marineau JJ, White PW, Brault M, Tao L, DeRoy P, Clavette C, Nayak S, Damon LJ, **Kaltheuner IH**, Bunch H, Cantley LC, Geyer M, Iwasa J, Dowell RD, Bentley DL, Old WM, Taatjes DJ. Selective inhibition of CDK7 reveals high-confidence targets and new models for TFIIH function in transcription. Genes Dev. 2020 Nov 1;34(21-22):1452-1473. doi: 10.1101/gad.341545.120.

LIST OF ABBREVIATIONS

Aβ	Amyloid β
AD	Alzheimer´s disease
ADP	Adenosine diphosphate
ADML	Adult myeloid leukaemia
AGC	Protein kinase A, G, and C
AMKL	Acute megakaryoblastic leukemia
AML	Acute myeloid leukemia
ALL	Acute lymphoblastic leukaemia
A-loop	Activation loop
APC/C	Anaphase-Promoting Complex
APP	Amyloid precursor protein
ATP	Adenosine triphosphate
bp	base pairs
BMP	Bone morphogenetic protein
Brd4	Bromodomain-containing protein 4
CAMK	Calcium/calmodulin-dependent kinases
CAK	CDK-activating kinase
CBP	CREB-binding protein
CDK	Cyclin-dependent kinase
CLK	CDK-like kinases
cCDK	cell-cycle regulating CDK
ChIP-seq	Chromatin immunoprecipitation sequencing
CKI	Casein kinase 1
C-lobe	C-terminal lobe
C-term.	C-terminal
CMGC	CDK, MAPK, GSK3, and CLK
cpm	counts per minute
cryo-EM	cryo-electron microscopy
CTD	Carboxy-terminal domain
Cyc	Cyclin
CV	Column volume
DDR	DNA damage-response
DH	DYRK homology box
DS	Down syndrome
DSIF	DRB sensitivity inducing factor
DYRK	Dual-specificity tyrosine kinase
E.coli	Escherichia coli
EGFR	Epidermal growth factor receptor
FCPI	TFIIF-associated component of the CTD phosphatase
FDA	U.S. Food and Drug Administration
FTD	Fronto-Temporal Degeneration
G-loop	Glycine-rich loop
GSK3β	Glycogen synthase kinase-3 beta
GST	Gluthation-S-transferase
HCMV	Human cytomegalovirus

HCV	Hepatitis C virus
HD	Huntington´s Disease
HER2	Human epidermal growth factor receptor 2
HEXIM	Hexamethylene bisacetamide inducible mRNAs
HID	Homeoprotein interaction domain
HIPK	Homeodomain interacting protein kinase
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HR	Homologous recombination
HSV	Herpes simplex virus
IP	Immunoprecipitation
IPTG	isopropyl β-D-thiogalactoside
MAPK	Mitogen-activated protein kinases
Mat1	Menage a trois 1
MBP	Maltose-binding protein
MePCE	Methyl phosphate capping enzyme
MLL	Mixed lineage leukaemia
mNET-seq	Mammalian native elongating transcript sequencing
MRD7	Mental Retardation Disease 7
mRNA	messenger RNA
MS	Mass spectrometry
NAPA	N-terminal auto-phosphorylation accessory domain
ncRNA	non-coding RNA
NELF	negative elongation factor
NIH	National Institute of Health
N-lobe	N-terminal lobe
N-term.	N-terminal
NLS	Nuclear localization signal
LARP7	La-related protein 7
LLPS	Liquid-liquid phase separation
LPS	lipopolysaccharide
OD_{600nm}	Optical density
PAF1	Polymerase associated factor
PARP	Poly(ADP-ribose) polymerase
PAS	Poly-adenylation site
PAXT	polyA tail eXosome Targeting
PCR	Polymerase Chain Reaction
PD	Parkinson´s Disease
PEST	Proline, glutamine, serine and threonine-rich region
PIC	Pre-initiation complex
PKA	Protein kinase A
PMA	phorbol-12-myristate-13-acetate
pol	RNA polymerase
PP	Protein phosphatase
ptRNA	prematurely terminated RNA
P-TEFb	Positive-transcription elongation factor beta
PTMs	Post-translational modifications
PROTAC	Proteolysis-targeting chimeras
PRP4	pre-mRNA-processing protein 4

RGC	Receptor guanylate cyclase
ROS	Reactive oxygen species
SCF	Skp1, Cullins, F-box proteins
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Super-elongation complex
Sf9	Spodoptera frugiperda
sgRNA	short guide RNA
SMKIs	small-molecule kinase inhibitors
snRNA	small nuclear RNA
SPR	Surface plasmon resonance
SQA	Serine, glutamine, and alanine-rich region
STE	Homologues of yeast sterile 7, 11, and 20
Strep	Streptavidin
Ssu72	Suppressor of sua7 gene 2
Tat	Transcriptional transactivator
TAR	Transactivation response RNA
tCDK	Transcription-regulating CDKs
TCERG1	Transcription elongation regulator 1 like
Tev	Tobacco etch virus
TGF	Transforming growth factor
TK	Tyrosine kinases
TKL	Tyrosine kinase-like
T-loop	Threonine-loop
TPD	Targeted protein degrader
TSS	Transcription start site
TTS	Transcription termination site
XPB	Xeroderma pigmentosum type B
XRN2	Terminal exoRNase
YAK1	Yet another kinase1

LIST OF FIGURES

ABSTRACT

Accurate regulation of RNA polymerase II by transcriptional kinases is essential for normal development and dysregulation can promote cancer onset and progression. Small molecules targeting transcriptional kinases, such as CDKs and DYRKs, hold promise for novel treatment options in oncology and beyond.

HIPKs are regulators of various cellular signalling pathways and belong to the DYRK kinase family. They are involved in the pathology of cancer, chronic fibrosis, diabetes, and multiple neurodegenerative diseases. In this study, characteristic structural features of the HIP kinase family were identified based on the HIPK3 crystal structure in the apo form at 2.5 Å resolution. Recombinant HIPKs and its close relative DYRK1A phosphorylate core components of the transcription machinery, such as the negative elongation factor SPT5, the transcription factor c-Myc, as well as pol II CTD, suggesting a direct role in transcriptional regulation. Abemaciclib, a Cdk4/Cdk6 inhibitor, which was approved by the FDA for the treatment of metastatic breast cancer, was identified in a database screen as a potent inhibitor of HIPK2, HIPK3, and DYRK1A. The crystal structures of HIPK3 and DYRK1A bound to abemaciclib were determined and demonstrated a binding mode to the kinase hinge region similar to Cdk6. Remarkably, abemaciclib inhibits DYRK1A to the same extent as Cdk4/Cdk6 in vitro, raising the question of whether transcriptional inhibition through DYRK1A contributes to the therapeutic activity of this breast cancer drug and whether abemaciclib may be a suitable treatment option for Alzheimer´s disease.

CDKs are dependent on a Cyclin subunit for activation and are central regulators of cell cycle and transcription. In contrast to cell cycle CDKs that can be activated promiscuously by multiple different Cyclins, transcriptional CDKs are thought to be activated by only one particular Cyclin partner. Although this long-established dogma exists, an unbiased approach to systematically test the formation of transcriptional CDK/Cyclin complexes revealed that both Cyclins, CycK and CycT1, are able to activate Cdk9, Cdk12 and Cdk13 in vitro. Cdk9/CycK has been a little appreciated phenomenon ever since the discovery of Cdk9 and we recently discovered a Cdk13/CycT1 complex mediating RNA surveillance in melanoma (Insco et al., *Science*, 2023). The crystal structure of the novel Cdk12/CycT1 complex at 2.0 Å resolution revealed a highly identical CDK/Cyclin interface compared to cognate Cdk12/CycK. Cdk13/CycT1 co-exists with Cdk13/CycK in HEK293T and A375 cells and ChIP-seq analysis confirmed co-localization of Cdk13 and CycT1 at transcription start sites of similar genes. Furthermore, Hex1 and Brd4 regulate Cdk9/CycT1, but no other transcriptional CDKs. Finally, small-molecular compounds inhibit CDKs independent of the bound Cyclin and Targeted protein degraders are highly specific for their targeted Cdk or Cyclin. The discovery of these novel CDK/Cyclin pairs challenges a long-standing dogma of CDK biology and may facilitate advances in the development of therapeutics targeting transcriptional CDKs.

TABLE OF CONTENT

1. INTRODUCTION

1.1 Eukaryotic transcription

Transcription is the first step of gene expression in which RNA is transcribed from a DNA template, which is catalyzed by a DNA-dependent RNA polymerase (pol). RNA polymerases are conserved in prokaryotes, archaea, and eukaryotes. While bacteria and archaea have only one RNA polymerase, eukaryotes possess three RNA polymerases, pol I, II and III[1]. This thesis focuses on the regulation of RNA pol II, a ~500 kDa multiprotein complex consisting of 12 subunits that transcribes all protein-coding genes as well as genes encoding noncoding (nc) RNAs. The human genome contains approx. 20,000 protein-coding genes and at least as many ncRNAs, highlighting the importance of a sophisticated regulatory network to control transcription.

1.1.1 Transcriptional control: From chromatin structure to gene level

Mammals have hundreds to thousands of different cell types[2]. A certain cell type is largely defined by the entirety of its gene products, thus cell types differ in their gene expression programs. The control of gene expression, its timing, location and the amount of gene product present in a cell have profound effects on the cell´s structure and function. A particular set of transcripts that establishes and maintains a cell state is under the control of specific DNA-binding transcription factors, their cofactors, chromatin modulators, the transcription machinery, and under the control of RNA pol II. Their precise crosstalk enables selective gene control, and the misregulation of gene expression can cause a variety of diseases[3].

Transcriptional control occurs at two interconnected levels: On the level of the 3D genome, the control of higher-order chromatin structure, and on the gene level, the control of the transcription machinery[4]. In eukaryotes, the genetic information is densely packed in form of higher-order chromatin (Fig.1): A nucleosome core complex is formed by two copies of the four different histone proteins H2A, H2B, H3 and H4, which are assembled into an octamer that has about 150 base pairs (bp) of DNA wrapped around it[5]. Such a nucleoprotein complex occurs in all eukaryotic genomes approx. every 200 bp, condensing the genetic information and thus storing DNA in a protected state from detrimental environmental factors. At the same time, nucleoprotein complexes control DNA accessibility along with other regulatory factors, such as general transcription factors, chromatin remodelers, and histone modifiers. A complex interplay of such regulators defines active or inactive genes, f.e. through acetylation and methylation, and genomic elements, f.e. through enhancers and promoters, which together regulate transcription at the chromatin level in a highly dynamic process, ultimately converging on the recruitment of the transcriptional machinery and cofactors to open chromatin regions[3,6,7].

Transcription of a single gene proceeds in an unidirectional multistep cycle, which is ordered in discrete phases, each controlled by checkpoints at which pol II irreversibly transitions to the next phase[8]. Each transcription cycle comprises the recruitment, initiation, pausing, elongation, termination and recycling phase and each phase is tightly regulated by the opposing activities of transcriptional kinases and antagonizing phosphatases[4]. Multiple transcriptional kinases, including transcriptional CDKs and DYRKs, mediate pol II through the

different phases of transcription by dynamic phosphorylation of multiple residues within the carboxy-terminal domain (CTD) of the Rpb1 subunit (Fig.1)[9].

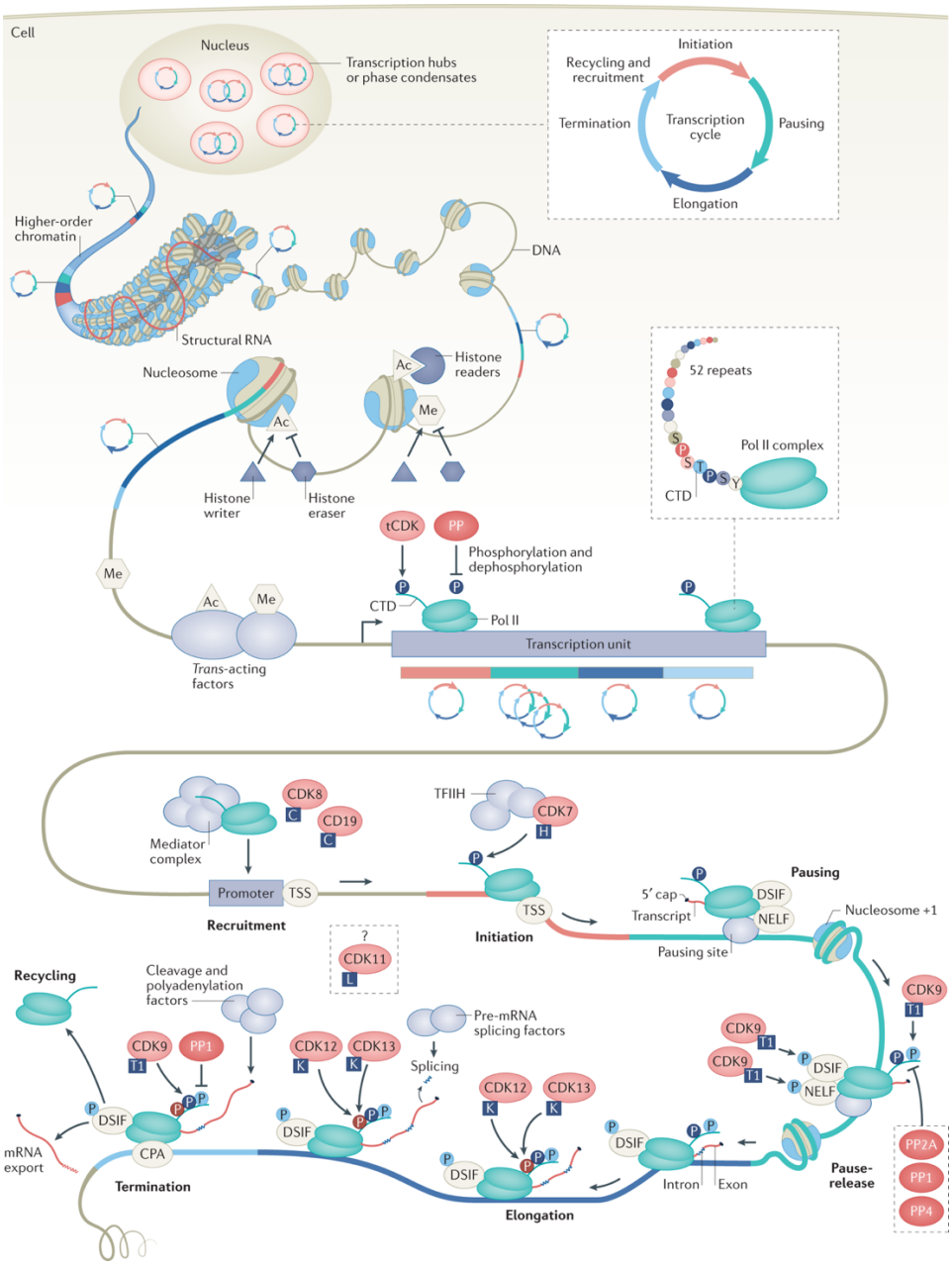

Fig.1 | Transcriptional control: From chromatin structure to gene level. DNA is densely packed as higher-order chromatin and wrapped around nucleosomes, which are formed of histones that can be modified by acetylation (Ac) and methylation (Me) influencing transcriptional states (f.e. active or inactive) and genomic elements (f.e. enhancers and promoters), which can recruit the core transcriptional machinery and cofactors. RNA pol II is governed through the different phases of the transcription cycle (Recruitment, Initiation, Pausing, Pause-release, Elongation, Termination, Recycling) by phosphorylation of its C-terminal domain (CTD) by transcriptional kinases (f.e. Cyclin-dependent kinases (CDKs) and their cognate Cyclins) and dephosphorylation by phosphatases (f.e. Protein phosphatase (PP) PP2A, PP1, PP4). DSIF: DRB sensitivity inducing factor; NELF: negative elongation factor; TSS: transcription start site; CPA: Cleavage and polyadenylation site (Vervoort et al. 2022)[4].

1.1.2 The C-terminal domain of RNA pol II

The transition of RNA pol II through the different phases of transcription, as well as co-transcriptional processes, such as splicing, capping and polyadenylation, are regulated by its largest subunit Rpb1 Carboxy-terminal domain (CTD), which is not present in RNA pol I and III. The pol II CTD is a repetitive and largely unstructured domain consisting of heptapeptide repeats with the consensus sequence $Tyr_1Ser_2Pro_3Thr_4Ser_5Pro_6Ser_7$ (YSPTSPS; Fig.2). The number of repeats varies between species with a tendency for more repeats in higher eukaryotes (52 in vertebrates) and less repeats in simpler eukaryotes (26 in *S. cerevisiae*), while some organisms lack a CTD entirely (f.e. *Trypanosoma brucei*)[10]. A full-length CTD is not required for cell viability or for sustaining wild-type growth rates in yeast, but growth defects were observed as soon as half or more repeats were removed in yeast or mammalian cells and a reduced number of repeats is unstable. Thus, it seems that a full-length CTD is preferred and actively maintained[9,10].

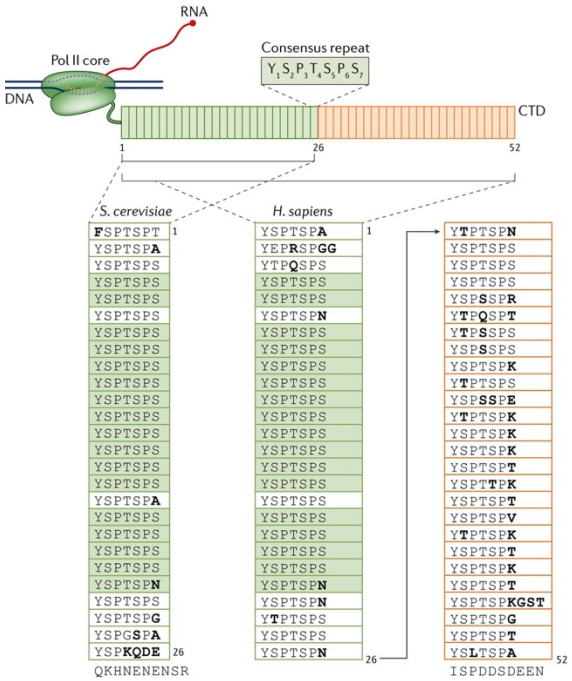

Fig.2 | The composition and conservation of pol II CTD. The amino acid sequence of the RNA pol II C-terminal domain (CTD) of yeast (*S. cerevisiae*) and human (*H. sapiens*) is depicted. Each rectangle represents one heptad repeat. The consensus sequence is shown above. The yeast CTD harbors 26 repeats of high conservation, whereas the human CTD is composed of 52 repeats. The first 26 repeats show a high degree of conservation (highlighted in green) between yeast and human while variations in the consensus sequence are more prominent in the distal part (Harlen & Churchman 2017)[10].

Of the 52 heptad repeats in human CTD, the N-terminal 26 repeats are highly conserved whereas the distal part shows frequent alterations from the consensus sequence, suggesting that the non-consensus repeats might have functions in the more sophisticated transcriptional regulation of higher eukaryotes (Fig.2). Most prominent is a change at position Ser7, as half of the serine residues in human are substituted by other amino acids, mostly lysine. The last repeat is extended to 17 amino acids and was shown to be important for CTD stability[11]. Studies in yeast showed that severe disruption but not deletion of single repeats

lead to growth defects. Furthermore, it has been shown that substitution of Ser2, Ser5, and Ser7 to glutamate is lethal, while the effects of Tyr1 substitution differ depending on the yeast strain[9,10].

The precise conformation of the CTD remains elusive, as it is not visible in crystal structures of yeast pol II due to its disordered nature. However, cryo-electron microscopy (EM) studies indicate that the CTD extends from the pol II core and forms a taillike structure, which functions as a binding platform for transcription and RNA processing factors[12,13]. The length of the CTD is dependent on the phospho-status. In yeast, the unphosphorylated CTD adopts a more compact state (~100 Å), while the phosphorylated CTD adopts an extended conformation of about 650 Å due to charge repulsion of the negative phosphate groups. Together with the linker, which can be up to 250 Å, the phosphorylated CTD would be five to six times longer than yeast RNA pol II (~150 Å)[9,14]. Thus, hypothetically, the CTD could reach basically anywhere on the pol II core complex, as well as toward DNA, pre-messenger (m) RNA, and other regulatory factors, allowing their spatiotemporal recruitment and crosstalk.

The CTD is not required for the catalytic function of pol II, but is central for governing pol II through the different phases of transcription, as well as mediating co-transcriptional processes, such as RNA processing and chromatin modification. The consensus sequence YSPTSPS consists exclusively of residues that can be subjected to post-translational modifications (PTMs). This includes phosphorylation (tyrosine, serine and threonine), isomerization (proline), glycosylation (serine and threonine), as well as methylation, acetylation and ubiquitylation of non-consensus lysine and arginine residues. These PTMs constitute the `CTD code´, controlled by the dynamic action of various enzymes. Hypothetically, originating from the unmodified CTD state, the possible number of combinations of PTMs together with the two proline configurations is 432 possibilities in a single hepta-consensus repeat[9]. Although the repetitive nature of the CTD increases the potential complexity of the code, recent work has demonstrated that many adjacent repeats are not phosphorylated. Thus, combinatorial phosphorylation implies that the number of actual CTD states is in the hundreds rather than astronomical[10].

By far the best-characterized CTD modification is phosphorylation. Each of the five hydroxy-side chain heptapeptide residues is dynamically phosphorylated and dephosphorylated depending on the phase of transcription, so that the relative abundance of these modifications changes during transcription along the gene[15]. These specific CTD phosphorylation patterns exist throughout the genome, are conserved among multiple species, and have been characterized extensively in yeast and human. Studies using Chromatin Immunoprecipitation (ChIP), mammalian native elongating transcript sequencing (mNET-seq), mass spectrometry (MS) analysis, and antibodies for specific CTD modifications in vitro, revealed that levels of pol II phosphorylated at Ser5 (pSer5) peak close to the transcription start site (TSS) of genes, as well as levels of pTyr1, pSer7, and acetylated and methylated Lys7, whereas levels of pSer2 and pThr4 peak at the 3´-end of genes (Fig.3)[10]. Overall, the patterns of modifications are similar in the consensus repeats in yeast and mammals, except for pTyr1, which is enriched along gene bodies in yeast and decreases at 5´- and 3´-ends of genes, as well as pThr4, which is more uniformly distributed along genes and peaks after polyadenylation (A) sites (PAS). Major changes in CTD phosphorylation occur at the transition between different transcription phases, providing a dynamic platform for the sequential recruitment of transcription and RNA-processing factors that control the transition to the next phase. The

CTD code follows a clear logic: factors that play a role early in transcription (f.e. capping enzymes) recognize modifications of the early phase (f.e. pSer5), while factors that have roles in later stages recognize late phospho-marks (f.e. pSer2).
Several transcription-regulating kinases have been shown to phosphorylate pol II CTD. Many studies suggest a certain redundancy among these CTD-phosphorylating kinases and in recent years also other cofactors of the transcription machinery have been identified as important substrates[16,17].

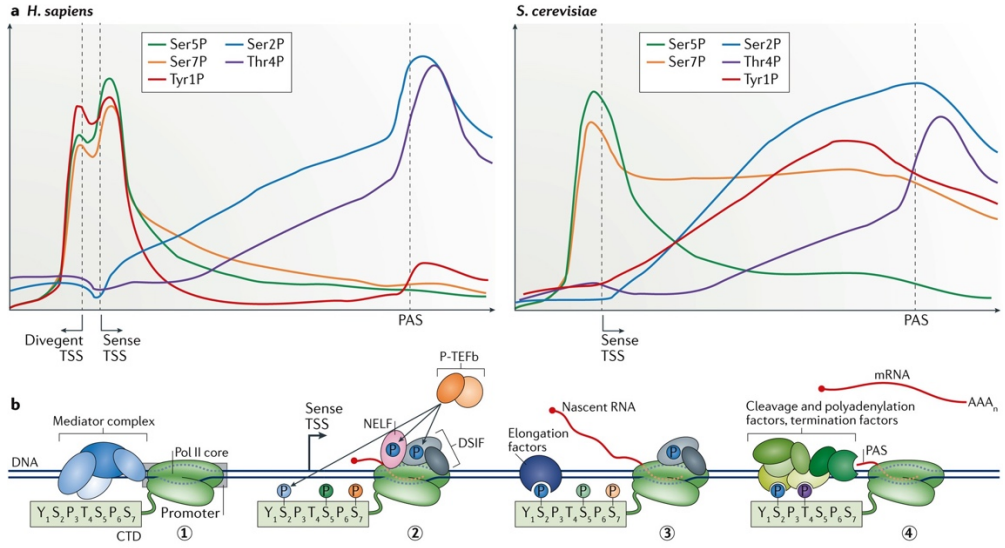

Fig.3 | The pol II CTD phosphorylation status along the gene. a Average chromatin immunoprecipitation (ChIP) profiles of phosphorylated residues of the RNA pol II CTD across protein-coding genes in human (*H. sapiens*) and yeast (*S. cerevisiae*). The profiles of yeast and human are highly similar, except for pTyr1. **b** The different phases of transcription are regulated by different complexes (f.e. TFIIH, Mediator, P-TEFb, NELF, and DSIF) containing transcriptional CDKs, which phosphorylate certain residues on the CTD corresponding for the particular phase of transcription. 1) Recruitment 2) Promoter-proximal pausing 3) Productive elongation 4) Termination; TSS: transcription start site; PAS: polyadenylation site (Harlen & Churchman 2017)[10].

1.1.3 The Transcription cycle

Transcription of a single gene proceeds in an unidirectional multistep transcription cycle (Fig.1). Transcription cycles can occur simultaneously as multiple pol II molecules can pass through a single gene at the same time. A transcription cycle can be divided into individual phases (Recruitment, Initiation, Pausing, Elongation, Termination and Recycling), and gene-intrinsic features determine the duration and speed of pol II in each phase. The distinct phases are defined by checkpoints that control pol II duration and irreversible transition to the next transcription phase[4,8].

Recruitment

During recruitment, chromatin is opened and a transcription bubble forms, which allows the assembly of the pre-initiation complex (PIC) at the promoter region, the site of transcription initiation[13]. The PIC is composed of pol II, general transcription factors, and the co-activator Mediator and its assembly has been studied extensively by structural and functional

approaches[18–21]. To enable gene activation, multiple transcription factors bind in a cooperative manner and make physical contacts between enhancer and promoter elements of nearby or distant genes through looping of DNA mediated by cohesin[2]. All genes require general transcription factors for recruitment of the PIC, while selective genes are additionally regulated by specific transcription factors and most protein-coding genes require Mediator[22]. The Mediator complex is composed of 26 subunits in mammals and integrates signals from diverse regulators, bridging interactions between transcription activators at enhancers and the core transcription machinery at promoters, and recruits cohesin and chromatin remodeling complexes to active genes[23]. Together, Mediator and general transcription factors mediate changes in the chromatin structure, unwind the DNA and recruit pol II through hydrophobic interactions between Mediator and unmodified pol II CTD[24]. These contacts are mediated by the Mediator subunits Med12 and Med13, which physically interact with the Mediator kinases Cyclin-dependent kinase (Cdk) 8 and its paralog Cdk19 in complex with Cyclin (Cyc) C[23]. Cdk8 and Cdk19 have been shown to phosphorylate pol II CTD at Ser2 and Ser5 residues in vitro, but the actual contribution to CTD phosphorylation and the resulting consequences in vivo are unclear[15]. However, a crucial role of both kinases resides in the phosphorylation of various transcription factors in the PIC, defining gene activation or repression[25]. Consequently, Cdk8 and Cdk19 are indirect but central regulators of the PIC, modulating Mediator complex formation and promoter recruitment. Finally, the Mediator complex recruits the TFIIH complex for the next step of transcription, the initiation of nascent RNA synthesis.

Initiation

Cdk7, CycH and other subunits such as the RING finger protein Menage a trois 1 (Mat1) and Xeroderma pigmentosum type B (XPB), are components of the general transcription factor TFIIH[26,27]. After the recruitment of TFIIH to the PIC by Mediator, Cdk7 phosphorylates pol II CTD at Ser5 and Ser7 residues[28]. Since the Mediator complex has a high affinity for unphosphorylated CTD, CTD Ser5 phosphorylation weakens pol II-Mediator association, thereby initiating pol II promoter escape into early transcription elongation[29]. Additionally, Cdk7 controls the recruitment of the mRNA capping machinery for the addition of the 5´-cap to mRNA, which protects it from exonuclease digestion[4].

Pausing

After early transcription elongation, the pol II complex almost immediately pauses 20-100 bp downstream of the TSS. This promoter-proximal pausing step is a key regulatory checkpoint and is one of the major rate-limiting steps of transcription[30,31]. Stalled pol II may transition to productive elongation through pause release, or it may ultimately terminate transcription with release of pre-mature RNA species[32,33]. Pol II pausing provides a window of opportunity for regulatory and co-transcriptional processes and is induced by pol II interaction with the DRB sensitivity-inducing factor (DSIF; a heterodimer comprised of SPT4 and SPT5) and the negative elongation factor (NELF; comprised of four subunits)[29,34].

To release pol II into productive elongation, Positive-transcription elongation factor beta (P-TEFb), composed of Cdk9 and CycT1, is recruited to pol II[35]. Cdk9 phosphorylates pol II CTD Ser2 and Ser5 residues, as well as a pol II CTD linker region, the C-terminal domain of the DSIF component SPT5, and NELF subunits[36,37]. Through phosphorylation of pol II CTD Ser2

residues, several elongation and chromatin-modifying factors are recruited to the paused pol II complex such as the polymerase associated factor (PAF1) complex, SPT6 and the human transcription elongation regulator 1 like (TCERG1)[10,28]. Upon phosphorylation, DSIF and NELF are released from pol II, which escapes into productive elongation[16]. Transcription factors such as c-Myc can stimulate increased elongation levels from entire gene expression programs by recruiting P-TEFb to stalled pol II complexes, accelerating pol II release into productive elongation[38].

Elongation

The molecular and catalytic mechanism of productive elongation have been studied extensively in a wide range of structural and functional studies[39–41]. During productive elongation, pol II processivity and elongation rate are controlled by the highly homologous kinases Cdk12 and Cdk13, which are both activated by CycK[4]. Cdk12 and Cdk13 phosphorylate pol II CTD at Ser2 in cells, and in vitro at Ser5 residues[42,43]. During productive elongation, levels of phosphorylated pol II CTD Ser5 and Ser7 decrease, while Ser2 phosphorylation increases[10]. This leads to the recruitment of diverse transcription elongation, chromatin-modifying and RNA-processing factors that regulate co-transcriptional processes that are facilitated by the positioning of the CTD proximal to the RNA exit channel. In addition, Cdk12 and Cdk13 regulate RNA-processing on multiple levels, as they control the synthesis of full-length mRNA transcripts by suppressing intronic and alternative polyadenylation. Also, Cdk11, which is less well understood, seems to be implicated in RNA-processing[4].

Termination

In contrast to pol I and III, pol II does not terminate transcription in a sequence-specific manner but rather at termination sites that are located several thousand bp past the PAS[44]. When pol II transcribes over a PAS, RNA is cleaved and a poly (A) tail is added to the 3´-end[45]. The pre-termination complex is postulated to be inherently labile, making structural studies difficult. Thus, it is yet unclear how transcription termination works mechanistically[4]. During transcription termination, phosphorylation levels of pol II CTD Ser2 and Thr4 residues peak, thereby promoting the recruitment of cleavage, polyadenylation, and termination factors that release pol II from DNA[10]. However, as pSer2 levels are high across gene bodies, but termination factors are concentrated only at the 3´end of genes, pTyr1 might play an additional role in preventing their recruitment along the gene[46]. The cross-talk between the three CTD phospho-marks, the PAS and termination factors enables the precise termination of transcription at a tight window at the 3´end of genes. Also, Cdk9 has been shown to play a direct role in transcription termination downstream of functional PAS and to phosphorylate the terminal exoRNase (Xrn2)[47], which is supported by the observation that Cdk9 inhibition causes termination defects[15].

Recycling

Phosphorylation of pol II CTD is a reversible process that allows its recycling and transition to a new transcription cycle. After termination, the loss of CTD modifications is essential and mediated by transcriptional phosphatases. In contrast to transcriptional kinases, the roles of protein phosphatases in transcription is yet poorly understood. Several phosphatases have been implicated in pol II CTD dephosphorylation, not only after transcription termination, but

rather consequently throughout the transcription cycle, thereby mediating the distinct phases by decreasing specific pol II CTD phospho-levels. Dephosphorylation of pol II at the recycling stage is controlled by phosphatases such as the TFIIF-associated component of the CTD phosphatase (FCP1), Suppressor of sua7 gene 2 (Ssu72) and Protein phosphatase 1 (PP1)[4,48].

1.2 Transcriptional kinases

1.2.1 The human kinome

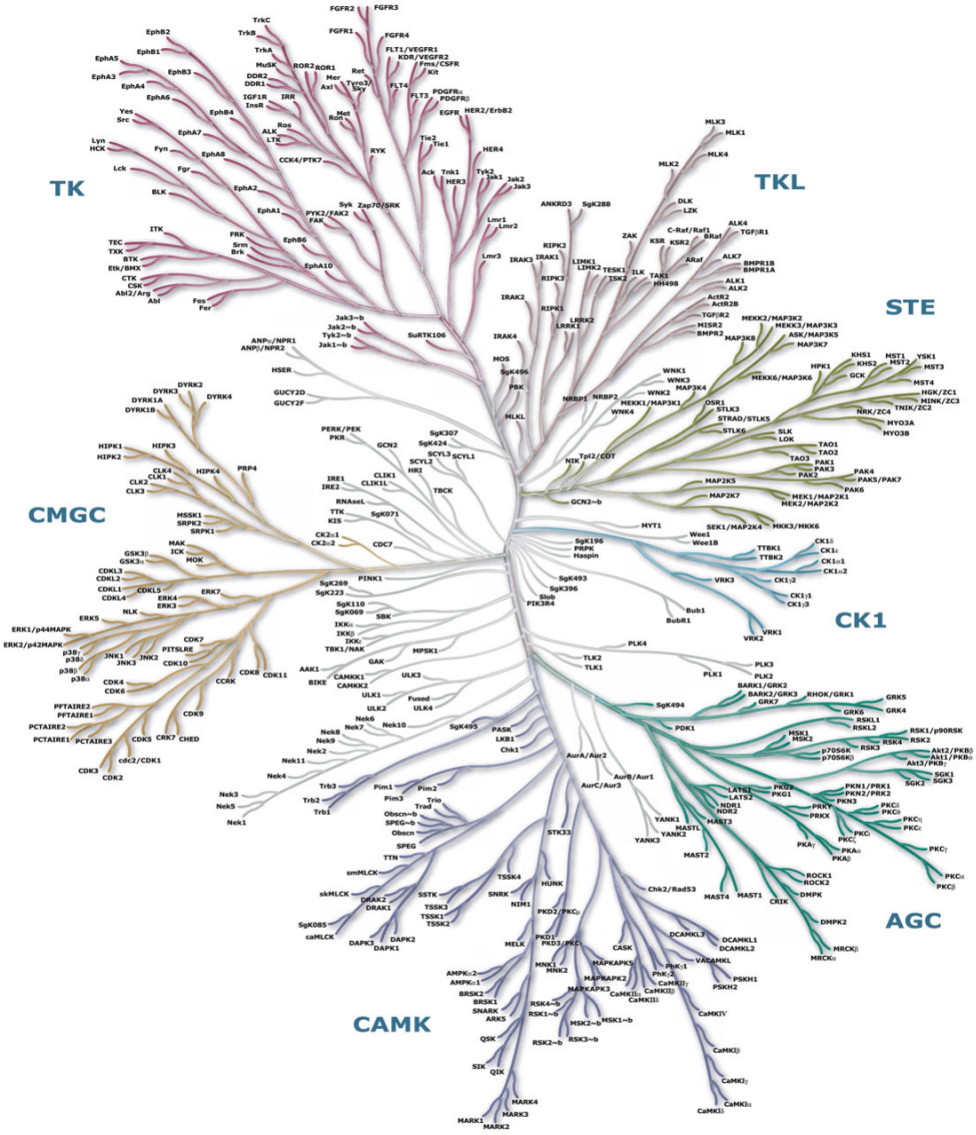

Fig.4 | The human kinome. Typical protein kinases are grouped based on sequence similarities. TK: Tyrosine kinases; RGC: Receptor guanylate cyclase (grey); CMGC: CDK, MAPK, GSK, CLK; CAMK: Calcium/Calmodulin dependent kinases; AGC: Protein kinase A, G, and C; CK1: Casein kinase 1; STE: Sterile of yeast homolog 7, 11 and 20; TKL: Tyrosine kinase like. Atypical protein kinases cannot be identified based on sequence homology (based on Manning et al.)[49].

Most cellular processes and signaling pathways in eukaryotes are regulated by reversible protein phosphorylation mediated by kinases and phosphatases. Protein kinases catalyze the

transfer of the terminal phosphate group from ATP to a hydroxyl-group of a protein substrate. The human kinome consists of 555 kinases, of which 497 are typical protein kinases that can be identified based on sequence similarity, and 58 atypical kinases, including lipid kinases, that lack sequence similarity to conventional kinases, but have been shown to exhibit kinase activity experimentally[50,51].

Typical protein kinases are further divided according to their residue specificity into serine/threonine- (388 kinases) or tyrosine-directed kinases (90 kinases). Furthermore, typical kinases can be subdivided into nine different groups (Fig.4): Tyrosine kinases (TK); Tyrosine kinase-like (TKL); Homologues of yeast sterile 7, 11, and 20 (STE); Receptor guanylate cyclase (RGC); CDK, MAPK, GSK3, and CLK (CMGC); Calcium/calmodulin-dependent kinases (CAMK); Protein kinase A, G, and C (AGC); Casein kinase 1 (CK1); and other[49].

1.2.2 The typical protein kinase fold

All typical protein kinases share a common kinase fold within a catalytic domain of approx. 300 residues, which contains highly conserved structural elements and sequence motifs that allow the coordination of ATP and a substrate in a precise spatial configuration that is critical for the catalytic phospho-transfer. The typical kinase domain consists of an N-terminal and a C-terminal lobe, which are connected by a flexible hinge region (Fig.5a). Sandwiched in between both lobes resides the active center of the kinase[52]. While the smaller N-lobe contains five antiparallel β-strands (β1-β5) and a regulatory αC-helix, the larger C-lobe is predominantly α-helical. In the N-lobe resides the glycine-rich loop (G-loop) between β1 and β2, which is a highly flexible region with the consensus sequence motif GXGXϕG (ϕ represents a hydrophobic residue) and a second key sequence motif (V)AxK is located in β3[51]. The C-lobe harbors the catalytic loop with the conserved Y/HRD motif, as well as the activation segment, which is a mobile element of variable length that begins with the DFG motif, extends over the activation loop, and ends with the helical APE motif[50] (Fig.5b).

Catalysis is mediated by the precise coordination of the ATP γ-phosphate and the substrate hydroxy group, resulting in a rigid active conformation. Hence, all active protein kinases share a common overall conformation, whereas inactive kinases adopt a more flexible and structurally diverse state as no spatial requirements constrain the fold. Nevertheless, inactive kinases share some common structural themes, in which the activation segment is partially disordered because the DFG phenylalanine is rotated out of the catalytic cleft (‘DFG-out’), similar to the αC-helix (‘αC-out’) and the catalytical glutamate[52].

Kinase activation is a highly complex process. Most signaling kinases are expressed in an inactive state that is activated by different control mechanisms, such as homo- and heterodimerization through the association of activation domains or subunits, removal of inhibitory segments or subunits, as well as phosphorylation of residues in the activation segment or at other sites. Finally, all these mechanisms promote closure of the two kinase lobes and repositioning of the αC-helix and the activation segment, which allows the precise coordination of all catalytic residues and eventually space for substrate binding. The activation segment forms a crucial part of the substrate-binding site, which orientates the substrate in a way that the hydroxyl group is directed toward a catalytic aspartate (D166 in protein kinase A (PKA)) in close proximity to the ATP γ-phosphate, which points out of the adenosine pocket (Fig.5a and d). In serine/threonine kinases, a lysine residue (K168 in PKA) two residues away from the catalytic aspartate contacts the γ-phosphate and stabilizes the

developing local negative charge during catalysis, whereas in tyrosine kinases, the stabilizing residue is an arginine, which is four residues away to allow for the larger tyrosine residue[52]. ATP is coordinated by residues of both lobes, forming hydrogen bonds between the adenine ring, the peptide backbone of the hinge region and the G-loop. Nonpolar aliphatic groups provide van der Waals contacts to the purine moiety. A conserved salt bridge forms in all active kinases between the lysine residue (K72 in PKA) in the (V)AxK motif and a conserved glutamate (E91 in PKA) located in the αC-helix, resulting in the `αC-in´ conformation and positioning the α- and β-phosphate groups. The C-terminus of the αC-helix is stabilized by interactions with the conserved DFG motif, resulting in the active `DFG-in´ conformation, and the DFG aspartate (D184 in PKA) pointing towards the phosphate moieties of ATP. Additional interactions between the negatively charged phosphate groups are mediated by two Mg^{2+} ions. One Mg^{2+} ion is chelated by the aspartate of the DFG motif (D184 in PKA) coordinating the β- and γ-phosphate group, while a second Mg^{2+} ion is coordinated by a asparagine residue (N171 in PKA) located on the C-lobe, which further contributes to the stabilization of ATP.

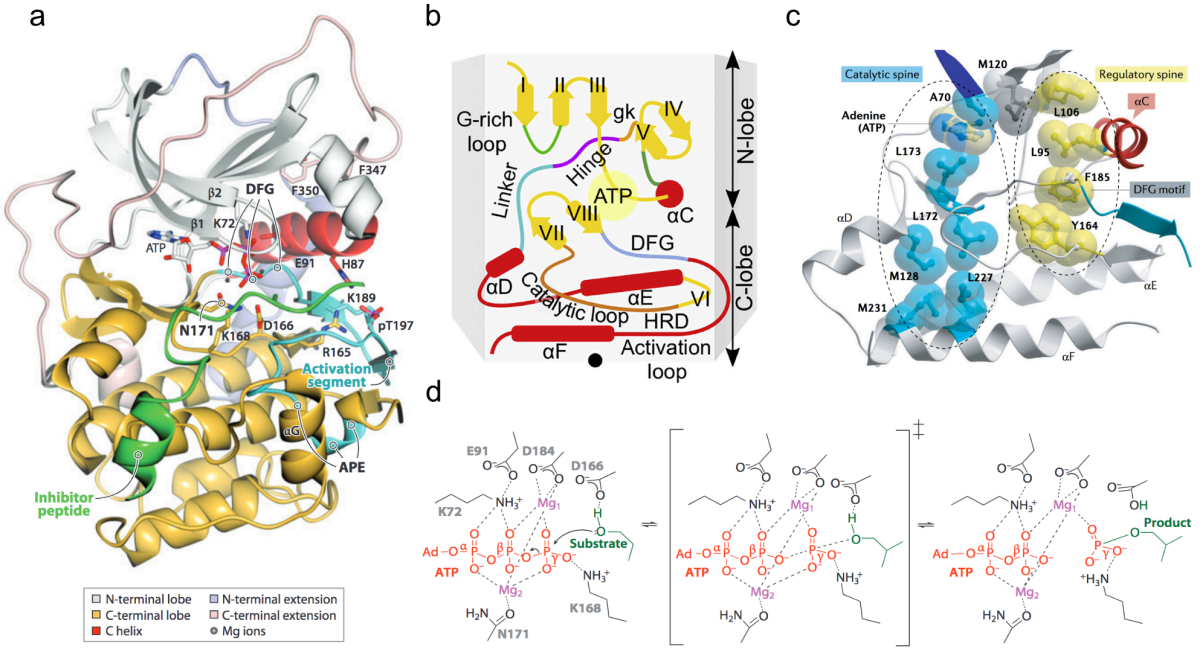

Fig.5 | The structural basis for kinase catalysis. a The typical active kinase fold with conserved structural elements and sequence motifs of the catalytic site of protein kinase A (PKA; PDB: 1ATP) with ATP and an inhibitor peptide (Endicott et al., 2012)[52]. **b** Schematic representation of conserved structural elements and sequence motifs of the typical kinase domain (Kanev et al., 2019)[50]. **c** Aligned residues of the catalytic and regulatory spine in the PKA active conformation. The gatekeeper residue (M120 in PKA) bridges the two spines (Attwood et al., 2021)[51]. **d** Schematic diagram of the protein kinase catalytic mechanism. The phospho-transfer proceeds from the kinase/substrate complex (left) through a transition state (center) to the kinase/product complex (right). Residue numbers for PKA are annotated (Endicott et al., 2012)[52].

In addition to the catalytically important residues that directly interact with the substrate and ATP, a secondary network of conserved hydrophobic residues form a catalytic and a regulatory spine (`C- and R-spine´) (Fig.5c). The hydrophobic contacts of both spines

assemble the catalytic motifs within both kinase lobes and simultaneously sense changes in the catalytic cleft, and transmit these changes to the rest of the domain[51]. The hydrophobic interactions that generate the C-spine result from residues of the N- and C-lobe, contact the ATP adenine ring and connect structural elements at the hinge region, whereas the R-spine interactions are promoted by the state of the activation segment in response to ATP binding and sense the proper alignment of the β4 sheet, the αC-helix, and the DFG motif[51]. A so-called gatekeeper residue (M120 in PKA) is located at the N-terminus of the hinge region and forms essential contacts between the C- and the R-spine.
The catalytic mechanism of the phospho-transfer is chemically simple, beginning with the nucleophilic attach of the substrate OH group in the kinase/substrate complex (Fig.5d). The attacking group (the substrate OH group) is aligned to the leaving group (phosphate ester oxygen) so that the oxygen electron pair is directed toward the ATP βγ-phosphate bridging[53]. The transition state involves a metaphosphate intermediate in which the bond breaking of the βγ-bridging oxygen of ATP is well advanced, while the incoming nucleophile bond making to the phosphorus is only just beginning. The two Mg^{2+} ions and nearby lysine residue compensate the developing negative charge that develops on the bridging oxygen as the reaction proceeds, finally allowing the departure of the leaving ADP. As the reaction proceeds, the acidity of the substrate hydroxyl group increases, and its pKa will become lower than the pKa of the nearby catalytic aspartate, thus allowing transfer of a proton from the hydroxyl to the aspartate. Finally, this proton is transferred to the phosphate dianion of the product restoring the catalytic site aspartate to the carboxylate state in the kinase/product complex[52]. While the kinase catalytic mechanism is chemically simple, the structural diversity of protein kinase activation and repression offers an extraordinary level of complexity to the enzyme family, highlighting the importance of structural and functional studies on kinase regulation.

1.2.3 The Cyclin-dependent Kinase Family

Cyclin-dependent kinases (CDKs) belong to the CMGC subfamily along with mitogen-activated protein kinases (MAPKs), glycogen synthase kinase-3 beta (GSK3β), dual-specificity tyrosine-regulated kinases (DYRK), and CDK-like kinases (CLK). CDKs are characterized by their dependency on a Cyclin (Cyc) subunit for their activation and were originally discovered for their role in cell cycle regulation in yeast and frogs. Cyclins that activate cell cycle-regulating CDKs (cCDKs) oscillate in their protein levels as they are subjected to degradation depending on the cell division phase. Since these pioneer studies conducted in the 1980s, the importance of CDKs acting not only in cell cycle regulation but also in transcriptional control of RNA pol II has been clearly established. Accordingly, CDKs can be grouped into cCDKs (Cdk1-6 and 14-18), regulating cell cycle progression by the promiscuous interaction with multiple different Cyclins, and into transcriptional CDKs (tCDKs; Cdk7-13, and 19-20) that regulate RNA pol II progression through the different stages of the transcription cycle (Fig.6). In contrast to cCyclins, tCyclins are continuously expressed throughout the cell cycle and thus activate only one particular Cdk[54]. CDKs have undergone an exceptional degree of evolutionary divergence and specialization demonstrated by increased numbers of CDKs during evolution (6 to 8 CDKs and 9 to 15 Cyclins in Fungi, 11 CDKs and 14 Cyclins in flies and echinodermata, and 20 CDKs and 29 Cyclins in humans)[54,55]. Presumably, tCDKs originated after cCDKs and became more diverse as the complexity of transcription increased throughout evolution[56]. Furthermore, CDKs range in size from approx.

300 amino acids, encompassing only the kinase domain, to approx.1,500 amino acids, with huge N- and C-terminal extensions (Fig.6a).
CDKs are serine/threonine kinases that show a preference for proline in the +1 position for their substrates as a consequence of the presence of a hydrophobic pocket near the catalytic site[57]. The empirically determined consensus sequence is (S/T)Px(R/K). However, recent findings show that tCDKs do not strictly depend on this sequence and that cCDKs and tCDKs have slightly different phospho-motifs correlating with different substrate preference in cell cycle and transcription[17,47,58].

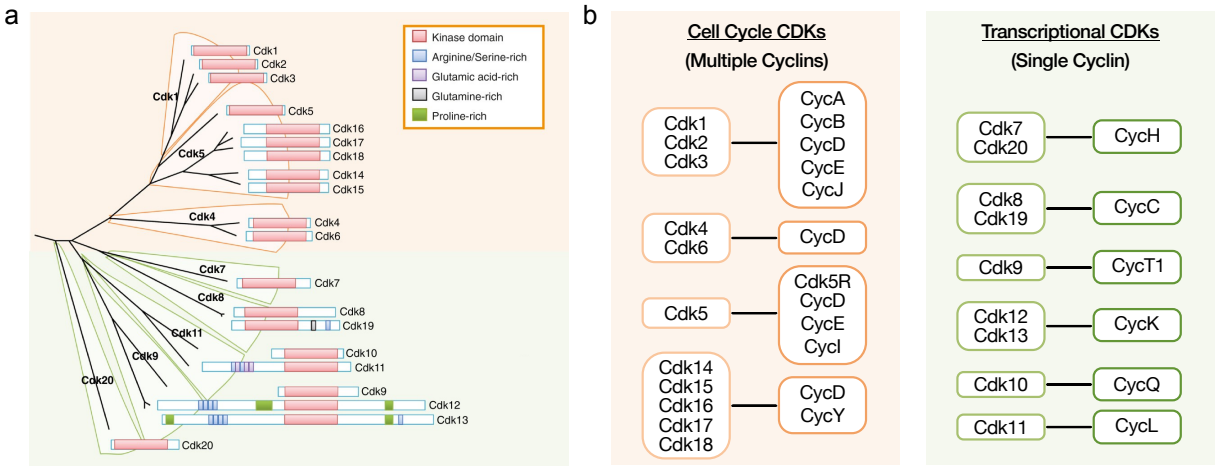

Fig.6 | Cyclin-dependent kinases. a Cyclin-dependent kinases (CDKs) were grouped based on sequence homology and divided into cell cycle (Cdk1-6, 14-18) and transcriptional CDKs (Cdk7-13, 19-20). Two genes exist for Cdk11 which encode two nearly identical kinases Cdk11A and Cdk11B respectively. **b** Cell cycle CDKs can be activated by multiple different Cyclins (Cyc), whereas transcriptional CDKs can be activated by only one single Cyclin partner (modified after Malumbres 2014)[54].

1.2.3.1 Structural features of CDKs

The first protein kinase structure was determined in 1991 by Knighton et al., who reported the structure of protein kinase A (PKA) in the active conformation[59], followed by the Cdk2 structure in the inactive conformation in 1993 by De Bondt et al.[60]. Overall, the Cdk2 fold is similar to other protein kinases but harbors some features that are exclusive to the CDK family. Like all protein kinases, CDKs comprise a two-lobed architecture with an αC-helix, a glycine-rich loop (G-loop) and an activation segment including the activation loop (A-loop). However, all CDKs rely on association with a Cyclin subunit for kinase activation and therefore, harbor a region critical for Cyclin binding, a PSTAIRE sequence (in Cdk2; with slight sequence deviations in other CDKs) in the αC-helix. Another important feature of many CDKs is a conserved threonine in the so called T-loop (T160 in Cdk2) in the activation segment, which is phosphorylated by a CDK-activating kinase (CAK) for kinase activation, which in human can be Cdk7[28].
In the Cyclin-free monomeric form, CDKs adopt an auto-inhibited conformation due to closure of the catalytic cleft by the T-loop (‘T-loop in’), which prevents ATP and substrate binding (Fig.7a). The activation loop, which is important for binding of the substrate phospho-acceptor Ser/Thr region, is partially disordered (‘A-loop out’) and the αC-helix swings out of the active

site (ˋαC-outˊ), breaking a catalytically important salt bridge of the catalytic glutamate (E51 in Cdk2, Fig7a)[61].

Upon Cyclin binding to the PSTAIRE region, the CDK αC-helix rotates into the active site (αC-in), the A-loop refolds into an extended conformation that opens up the catalytic cleft (ˋA-loop outˊ) and the N- and C-lobes shift in their relative orientation resulting in a conformation which allows correct binding of the substrate and ATP[54] (Fig.7b).

Full activity of many but not all CDK/Cyclin complexes is acquired by phosphorylation of a conserved threonine in the T-loop (T160 in Cdk2; Fig.7b), which results in further changes of the activation segment[62]. Upon phosphorylation, the conserved threonine coordinates three arginine residues (R50, R126, and R150 in Cdk2), ultimately rigidifying and stabilizing the entire active kinase conformation. However, not all CDKs harbor a conserved threonine residue in their T-loop, but a serine residue instead (Cdk14-Cdk18), a phospho-mimetic aspartate (Cdk8 and Cdk19), or circumvent phosphorylation in the T-loop through interaction with other regulatory factors (Cdk5)[63]. In Cdk2, the conserved threonine (T160) is first phosphorylated by CAK and subsequently participates in the interaction with the Cyclin subunit. Cyclin binding buries the phosphorylated T160 residue, finally blocking dephosphorylation and inactivation of the kinase[64]. On the other hand, some CDKs bind their Cyclin first, which rotates the activation segment out of the catalytic cleft so that the T-loop becomes accessible for activating phosphorylation by CAK. However, how Cyclin binding and T-loop phosphorylation are ordered and whether there are any differences, enzymatically, between the same CDK phosphorylated prior to or after Cyclin binding is still under investigation and can differ between CDKs[28].

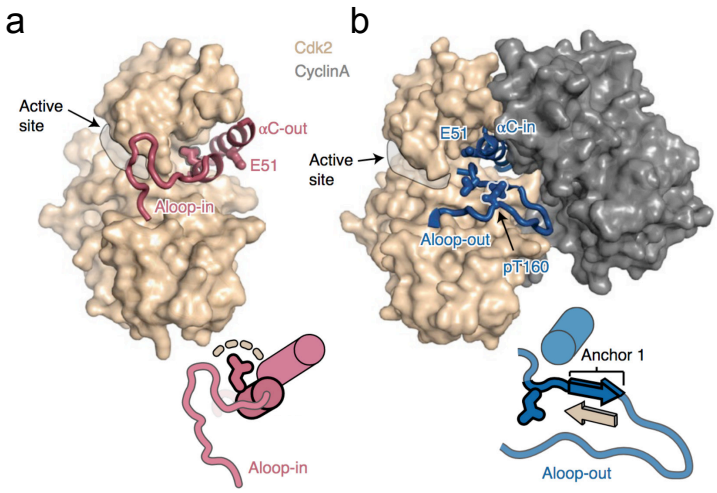

Fig.7 | Structural features of CDK activation. a Crystal structure of inactive monomeric Cdk2 (PDB: 1HCL) in the activation loop (Aloop)-in and αC-out conformation with an inaccessible catalytic pocket. **b** Crystal structure of the activated Cdk2/CycA complex (PDB: 1JST) with the phosphorylated threonine (pT160), the Aloop-out and αC-in conformation in which the C-helix and the activation segment are pulled apart, making the enzymatic pocket accessibly for ATP and substrate binding; E51: catalytic glutamate (Majumdar et al., 2021)[61].

Cyclins are classified by a Cyclin box, which spans over approx. 100 amino acids and consist of a five α-helix bundle, mediating the interaction with the CDK PSTAIRE motif through key contacts of the H5 and H3-helix. These interactions represent the conserved heart of the

Cdk/Cyclin interface[65]. Most Cyclins contain two Cyclin boxes that are tandemly arranged and separated by a linker of variable length and can harbor N- and C-terminal extensions of different length adjacent to the Cyclin box region that can have regulatory functions. Both Cyclin boxes share a high degree of structural similarity but little sequence identity. Cyclins, however, undergo no conformational changes between their free, bound and kinase active state[66].

1.2.3.2 Regulation of Cdk9, Cdk12 and Cdk13

Transcriptional CDKs have specific functions during the different phases of the transcription cycle, often in association with larger regulatory complexes. Cdk9 together with CycT1 is also known as Positive Elongation Factor b (P-TEFb), which is further regulated by the 7SK small nuclear (sn) RNP complex or the Super Elongation Complex (SEC) for pol II pause-release. The importance of P-TEFb in transcriptional regulation became apparent when CycT1 was discovered as the cellular target of the HIV-1 transcriptional transactivator (Tat)/transactivation response RNA (TAR) ribonucleoprotein complex, which increases replication by stimulating the transcription of viral genes in the elongation phase. P-TEFb regulates many genes that are required for cell growth and proliferation. Therefore, a precise equilibrium of P-TEFb levels is required for proper homeostasis. The 7SK snRNP complex has long been known to regulate Cdk9 activity and consists of a 7SK RNA, which is stabilized by methyl phosphate capping enzyme (MePCE) and La-related protein 7 (LARP7). Hexamethylene bisacetamide inducible mRNAs (Hexim) 1 or 2 assemble with Cdk9 and CycT1, thereby sequestering Cdk9 in an inactive state and functioning as an adapter to 7SK RNA[66,67]. In proliferating cells exists an equilibrium between the active, free and inactive, sequestered P-TEFb (approx. 50-90% inactive)[68]. The release of P-TEFb from the 7SK complex occurs when transcription is inhibited, f.e. after treatment with UV light, actinomycin D, or P-TEFb inhibitors[64]. However, details of the release mechanism are not known.
The free portion of P-TEFb can assemble with Bromodomain-containing protein 4 (Brd4) or components of the Super Elongation Complex (SEC), which recruit P-TEFb to chromatin and facilitate pol II release. SEC consists of one of four AFF scaffold proteins (AFF1-4), one of three ELL proteins (ELL1-3), and ENL or its analogue AF9. P-TEFb recruited by Brd4 has been shown to phosphorylate DSIF, while SEC-recruited P-TEFb phosphorylates NELF and pol II CTD Ser2 residue[68].
The transcriptional elongation and RNA processing CDKs, Cdk12 and Cdk13, are less well studied and have not yet been shown to participate in larger complex formation or the assembly with regulators.

1.2.4 The Dual-specificity tyrosine-regulated Kinase Family

The family of dual-specific tyrosine-regulated kinases (DYRKs) belongs to the CMGC group and is highly conserved from yeast to human. The DYRK family can be further divided into five subfamilies: Class I DYRKs, Class II DYRKs, Yet another kinase1 (YAK1), Homeodomain-interacting protein kinases (HIPK), and pre-mRNA-processing protein 4 (PRP4)[69]. Members of the DYRK family auto-phosphorylate a conserved tyrosine residue in their activation loop, but phosphorylate substrates at serine and threonine residues and are therefore named for their dual-specific character[70]. For the class II DYRK ortholog from *D. melanogaster*, it has been shown that the critical tyrosine is *cis*-auto-phosphorylated by the nascent kinase in a transient

intermediate state during its maturation at the ribosome[71]. Phosphorylation of the critical tyrosine was later confirmed for members of all DYRK subfamilies by biochemical and structural analyses[72,73].
The DYRK substrate binding pocket has the typical appearance of serine/threonine kinases. Initially, DYRKs were expected to be specific for substrates with a proline at P+1 and an arginine in the P-3 position[74–76]. However, cellular substrates of DYRKs exhibit a wide variety of phosphorylation motifs and functional and structural analysis revealed that DYRK1A also recognizes substrates with a small hydrophobic residue such as valine or alanine in the P+1 position, demonstrating that DYRKs are not stringently proline directed[58,72].

1.2.4.1 DYRK1A

The Class I and Class II DYRK families consist of DYRK1A and DYRK1B, and DYRK2-4, respectively[77]. Class I and Class II members share a conserved central kinase domain and an adjacent N-terminal DH-box (DYRK homology box) but have divergent extended N- and C-terminal regions[72] (Fig.8a). The N-terminal region of all DYRKs except for DYRK3 contains a nuclear localization signal (NLS)[78] and DYRK2-4 harbor a conserved N-terminal auto-phosphorylation accessory (NAPA) domain crucial for tyrosine auto-phosphorylation[79]. The C-terminal region of DYRK1A and DYRK1B contains a proline, glutamine, serine and threonine-rich (PEST) region, which is important for proteasomal degradation[80] and DYRK1A additionally possesses a histidine-rich region in its C-terminus that mediates localization to nuclear speckles.
The best-studied member of the DYRK family is DYRK1A, owing to its role in the pathology of Down syndrome and neurodegenerative diseases[81-84]. In mice, the absence of DYRK1A is lethal at the embryonic stage. DYRK1A is ubiquitously expressed in all human tissues from early embryonic development to adulthood[85,86] with particularly high expression levels in several areas of the adult brain, especially in the cerebellum[87]. DYRK1A has diverse cellular functions in chromatin modification, transcriptional regulation, splicing, DNA damage repair, apoptosis, cell cycle, neuronal development, and synaptic plasticity[88]. Regarding the control of gene expression, DYRK1A phosphorylates and thereby activates several transcription factors important for muscle development, neurogenesis, and immune responses, as well as chromatin-modifying enzymes. Furthermore, DYRK1A has been shown to directly interact with the general transcription factor TFII-I and the transcriptional co-activator CREB-binding protein (CBP) and to phosphorylate RNA pol II on Ser2 and Ser5 residues, suggesting that it may play a role in the regulation of transcriptional elongation[89].
Due to its low tissue specificity and involvement in a variety of signaling pathways, alterations in DYRK1A protein levels result in a wide number of human diseases (Fig.8b). Initially, DYRK1A has been discovered due to its localization in the Trisomy 21 critical region on chromosome 21.3, resulting in 1.5-fold increased levels of DYRK1A in Down syndrome (DS), which are responsible for multiple symptoms observed in DS including cognitive defects[87]. Interestingly, cognitive functions are restored by genetic normalization of DYRK1A expression or pharmacological inhibition of its catalytic activity[90]. Additionally, protein levels of DYRK1A are elevated in a variety of neurological diseases with cognitive impairment and have been shown to be critical for onset of Alzheimer´s Disease (AD), Parkinson´s Disease (PD), Huntington´s Disease (HD) and dementia as it phosphorylates key substrates, such as tau, parkin, septin 4, α-synuclein, amyloid precursor protein (APP), and β-tubulin[91,92]. Consistent

with this, DYRK1A inhibition in AD transgenic mouse models decreased APP and amyloid-β accumulation[93–95] and DYRK1A inhibition has been shown to inhibit tau oligomerization and aggregation[96,97] and to restore cognitive defects[98]. On the other hand, in Mental Retardation Disease 7 (MRD7), an autism spectrum disorder characterized by microcephaly, intellectual disability, and speech impairment, which results from haploinsufficiency of the DYRK1A gene, DYRK1A levels are diminished[87]. Both, DS and MRD7 affect similar brain areas and functions but both diseases demonstrate the critical dosage sensitivity of DYRK1A as either too much or too little kinase activity has negative health consequences, probably resulting from the critical role of DYRK1A in neuronal development and timing of neural progenitor cell differentiation.

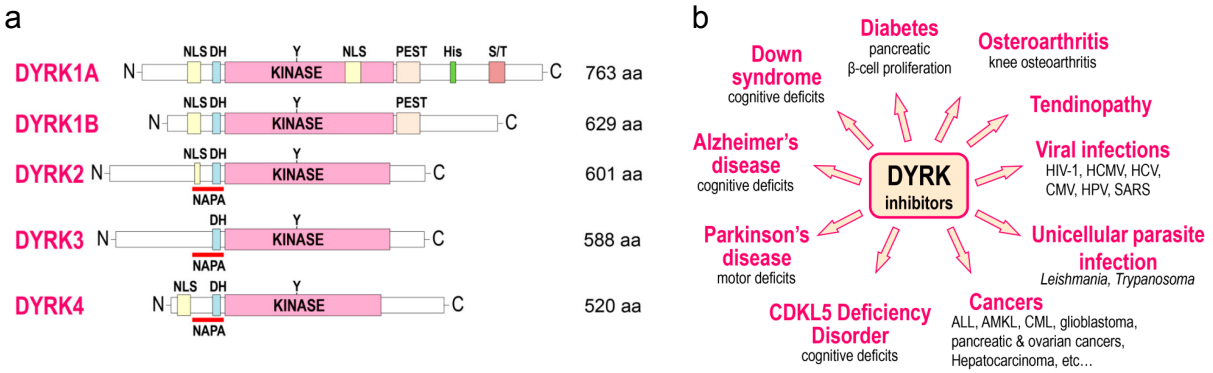

Fig.8 | The Class I and II DYRK families. a Domain architecture of human DYRK1-4. DH: DYRK homology box; His: Histidine-rich domain; Kinase: Kinase domain; NAPA: N-terminal autophosphorylation accessory domain; NLS: nuclear localization signal; PEST: proline, glutamine, serine, and threonine-rich domain; S/T: serine/threonine-rich region; Y: conserved tyrosine residue. **b** Schematic representation of diseases that could benefit from DYRK inhibitors (Lindberg & Meijer, 2021)[81].

There is growing evidence that DYRK1A inhibition induces the proliferation of insulin-producing pancreatic β-cells both in cells and animal models, bringing DYRK1A inhibitors in focus as a potential therapeutic target for type 1 and type 2 diabetes by restoring endogenous insulin production[99,100]. Moreover, DYRK1A is involved in the progression of viral infections, as DYRK1A inhibition has been shown to prevent replication of various viruses, including hepatitis C virus (HCV), human cytomegalovirus (HCMV), human immunodeficiency virus type 1 (HIV-1), and herpes simplex virus 1 (HSV-1)[101-102]. Furthermore, DYRK1A interacts with oncoproteins from adenovirus and human papillomavirus (HPV), an oncogenic virus that can cause the development of cervical cancer.
Overall, DYRK1A functions are complex and it seems that both, increased and decreased DYRK1A levels are critical for a variety of diseases[84]. Therefore, the therapeutic potential of selective DYRK1A inhibitors has attracted increasing attention in recent years, as they may provide new therapeutic opportunities for some, so far untreatable diseases such as AD and DS.

1.2.4.2 The HIPK family

Homeodomain-interacting protein kinases (HIPKs) are an evolutionarily conserved kinase family consisting of the members HIPK1, HIPK2, and HIPK3 in vertebrates and additionally HIPK4 in mammals[103,104]. HIPK1 and HIPK4 are mainly located in the cytoplasm, whereas HIPK3 is found in the nucleus and HIPK2 undergoes dynamic shuttling between different cellular compartments[105,106].
The close relationship between HIPKs and Class I DYRKs is highlighted by the fact that DYRK1A and HIPK3 share 39% sequence identity within their catalytic domain (Fig.9a). Among the HIPK family, HIPK1 and HIPK2 are the most closely related members, showing approx. 93% sequence identity in their kinase domains, whereas HIPK3 is slightly less conserved with 87% identity. HIPK4 is the most divergent member of the family, sharing only 50% sequence identity with the catalytic domains of the other three HIP kinases.
The smaller and more unique HIPK4 comprises only the kinase domain[105,107], while HIPK1-3 have a complex domain architecture consisting of multiple domains involved in protein-protein interaction (Fig.9b). In HIPK1-3 the highly conserved kinase domain is located at the N-terminus followed by the name-giving homeoprotein interaction domain (HID), which mediates interactions with homeodomain transcription factors[103]. A proline, glutamate, serine, and threonine (PEST)-rich domain follows the HID, mediating proteasomal degradation of these kinases. The C-terminus of HIPK1-3 harbors a serine, glutamine, and alanine (SQA)-rich region, which is involved in the interaction with various co-factors[108].

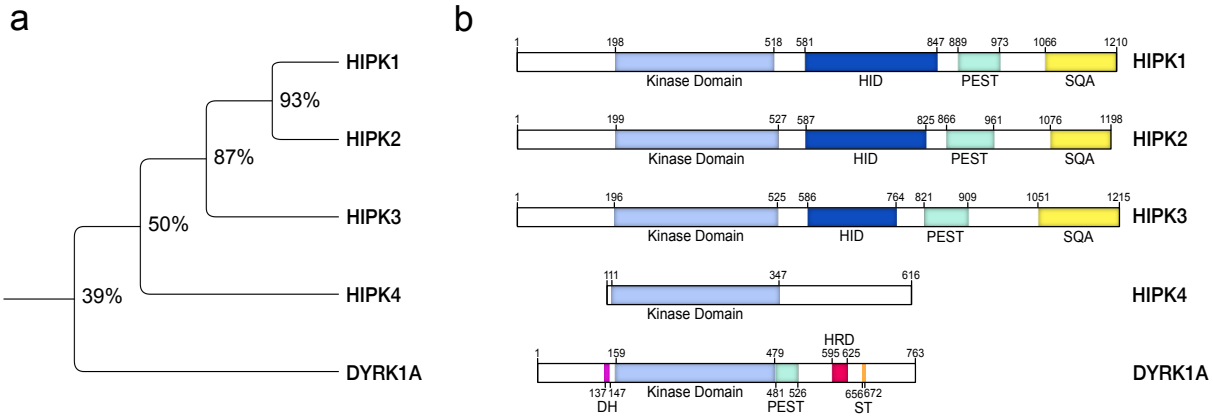

Fig.9 | DYRK1A and the HIPK family. a Phylogenetic tree of human HIPK1-4 and DYRK1A illustrating the identity of the kinase domains. **b** Domain architecture of human HIPK1-4 and DYRK1A. HID: homeoprotein-interaction domain; PEST: proline, glutamate, serine, threonine-rich region; SQA: serine, glutamine, alanine-rich region; DH: DYRK homology box; HRD: histidine-rich domain; ST: serine/threonine-rich region.

Due to autophosphorylation of their activation loop, DYRKs and HIPKs are in a constitutive, at least partially active state after maturation at the ribosome. Therefore, HIPKs undergo additional regulation through extensive post-translation modifications (PTMs) such as phosphorylation, acetylation, ubiquitination, SUMOylation, and caspase cleavage[109]. Collectively, these PTMs have been shown to alter protein stability and subcellular localization, ultimately controlling kinase activity.

Several studies have shown that HIPKs phosphorylate various transcriptional regulators and chromatin modifiers in a variety of signaling pathways such as nutritional stress, DNA damage, reactive oxygen species (ROS), virus infection, and hypoxia stress signaling[104]. Thus, the main function of HIPKs appears to be their ability to promote or repress gene expression. Importantly, HIPKs have not been identified as key components of directed downstream signaling pathways. Instead, HIPKs seem to play a more subtle role as "fine-tuners" that integrate multiple incoming signals into downstream effector pathways such as Salvador-Warts-Hippo, bone morphogenetic protein (BMP), Wnt/Wingless, DNA damage response, Notch, MAPK, p53, and transforming growth factor (TGF) β[104]. Consequently, dysregulation of HIPKs links them to the pathology of cancer, diabetes, chronic fibrosis, and certain neurodegenerative diseases such as Alzheimer's, Parkinson´s and Huntington's disease[110,111]. Therefore, similar to DYRK1A, the specific inhibition of HIPKs may provide new therapeutic opportunities for some of these diseases.

1.3 Targeting Transcriptional Kinases in Cancer

1.3.1 Therapeutic Kinase Inhibition in Cancer

Kinases are attractive targets in oncology and beyond due to their central roles in tumour growth, survival and resistance. The U.S. Food and Drug Administration (FDA)-approval of imatinib in 2001 was a breakthrough in targeted cancer therapy and marked the beginning of the kinase inhibitor era, which has been rising ever since. Twenty years later 71 small-molecule kinase inhibitors (SMKIs) were approved by the FDA predominantly in the oncological field and a few other areas, targeting approx. 20% of the human kinome[51]. The success rates of kinase inhibitors in clinical trials are relatively high compared to other drug targets, as the probabilities for phase transitions are 72% from phase I to phase II, 66% from phase II to phase III and 86% from phase III to approval and marketing[51]. The growing understanding of catalytic structures (covering approx. 300 kinases from 104 families) has significantly advanced selective kinase inhibitor design. Nearly all FDA-approved SMKIs target the ATP-binding pocket with only few exceptions. Due to high conservation of the ATP-binding site between all kinases, cross-reactivity of SMKIs is a common phenomenon resulting in promiscuous inhibitory profiles. Such pan-selective inhibitors can exhibit high toxicology, or on the other hand synergistic effects by blocking multiple signalling pathways and thus can offer potential clinical benefits[112].

Strategies for rational kinase inhibitor design

The overwhelming majority of kinase inhibitors compete with ATP for interactions with the adenine pocket close to the hinge region and often interact with spine residues. For the design of ATP mimetic inhibitors the hydrogen bonds between ATP and the hinge region represent important anchor points and define the basic kinase inhibitor scaffold. SMKIs target kinases in different modes of action: Type 1 inhibitors target the kinase active site (Fig.10a), in which the compound´s central purine moiety mimics hydrogen bonds of ATP binding to the hinge-region, and both αC and the DFG motif are in their active ‘in’ position. Type 2 inhibitors bind to an inactive kinase state that is generally highly diverse between different kinases and is characterized by the αC-in and DFG-out conformation, which creates a large hydrophobic pocket that is targeted by the compound. Allosteric kinase inhibitors often mimic natural

regulatory mechanisms in a highly specific manner and differ in their location of binding. Type 3 inhibitors bind adjacent to the ATP-binding site in a pocket created by an αC-out conformation and in some cases by an inactive conformation of the A-loop, whereas Type 4 inhibitors bind remote sites from the ATP-binding pocket. Compared to Type 1 and 2 inhibitors, non-canonical binding modes represent attractive design strategies because these compounds are often highly selective and have prolonged target residence times. Moreover, additional approaches can provide higher selectivity with enormous therapeutic potential and the kinase field has only started to explore these new strategies in clinical settings[51] (Fig.10b). Covalent inhibitors interact with kinases by forming a covalent bond with nucleophilic residues, which can be either reversible or irreversible. Bivalent inhibitors target kinases simultaneously at two different binding sites or domains, f.e. the ATP-binding site and a remote secondary site. Eventhough the increase in selectivity may provide interesting chemical tools for functional studies and target validation, their large molecular weight challenges their clinical development. Targeted protein degraders (TPDs) target kinases via recruitment of components of the ubiquitin degradative pathways. Different classes of TPDs exist: Proteolysis-targeting chimeras (PROTACs) comprise bivalent inhibitors, which link a conventional ATP mimetic inhibitor via a chemical linker to a ligand that binds to an E3 ligase, f.e. thalidomide or its derivative lenalidomide. Since protein kinases are usually highly amenable to degradation, PROTACs have received a lot of attention owing to increased isoform and mutant selectivity[113-116] and first PROTACs have now entered clinical testing[113]. Molecular glues are variants of PROTACs that bind to the interface of two proteins without a linker[117]. However, strategies for the rational design of molecular glues are currently lacking, as most molecular glues have been discovered serendipitously or have been developed based on natural products. However, owing to their low molecular weight, molecular glues can be attractive lead compounds for drug development with very specific modes of action and can allow targeting of so far `undruggable targets´[118].

The most important challenges that remain in kinase drug discovery comprise validation of novel kinase targets, utilizing kinase inhibitors in non-oncological therapeutic areas, developing efficient screening and profiling technologies, increasing target selectivity to reduce off-target toxicity and overcoming drug resistance[112]. Identification of the molecular characteristics of tumours enabled the development of first-generation SMKIs. However, long-lasting therapeutic effects are frequently counteracted by pre-existing or de novo evolving resistance mechanisms resulting from therapy, which lead to the development of second-, third- and next-generation inhibitors[50]. The most common mutations responsible for drug resistance in kinases are mutations at position 4 in the G-loop and the gatekeeper residue, f.e. T790M in the epidermal growth factor receptor (EGFR; in up to 50% of patients treated with EGFR inhibitors). Kinases with gatekeeper mutations exhibit bulkier residues in this position that sterically exclude inhibitor binding and significantly decrease the efficacy of first- and second-generation kinase inhibitors. Attempts to prevent escape mechanisms comprise the identification of synergistic drug combinations, the rational design of inhibitors with a specific poly-pharmacological profile that matches the genetic fingerprint of the patient and the search for alternative targets[50]. Shifting the focus from inhibition of oncogenes to basal cellular processes which tumors disproportionally rely on has yielded some promising results, as demonstrated for kinases that regulate the transcriptional machinery[4].

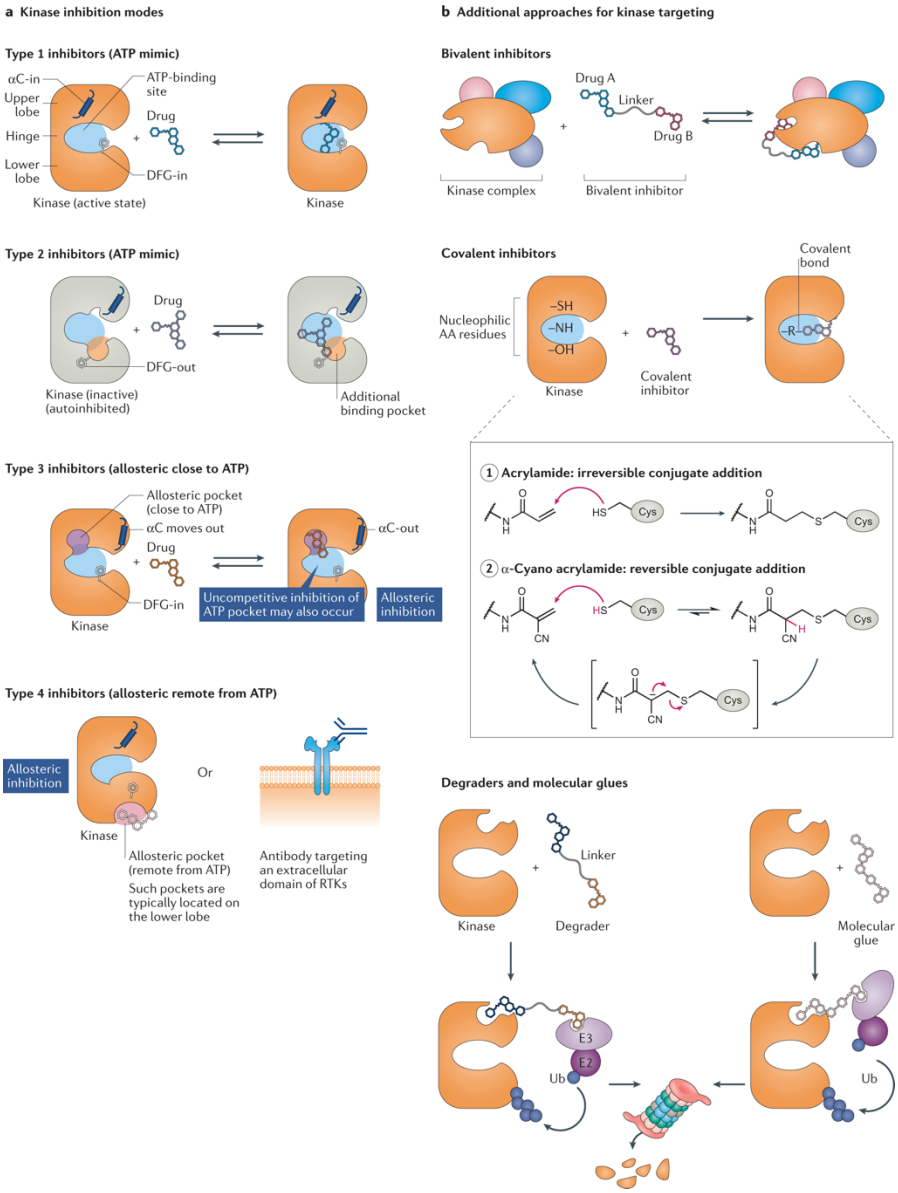

Fig.10 | Kinase Inhibitor binding modes. a Primary modes of kinase inhibition. Type 1 inhibitors interact with kinases at the ATP-binding site in the active, DFG-out and αC-in conformation. Type 2 inhibitors interact with kinases at the ATP-binding site in the inactive, DFG-out and αC-in conformation. Type 3 inhibitors interact with kinases at an allosteric site close to the ATP-binding site, resulting in an αC-out conformation. Type 4 inhibitors interact with kinases at an allosteric site remote from the ATP-binding site, f.e. targeting of extracellular domains by antibodies. **b** Additional approaches for kinase targeting. Bivalent inhibitors target kinases simultaneously at two different binding sites or domains. Covalent inhibitors interact with kinases by forming a covalent bond with nucleophilic residues, which can be either reversible or irreversible. Degraders and molecular glues target kinases via recruitment of components of the ubiquitin (Ub) degradative pathways. E3: E3 ubiquitin ligase; E2: E2 ubiquitin ligase (Attwood et al., 2021)[51].

1.3.2 Transcriptional Addiction in Cancer

Cancer combines a group of diseases of which more than 100 types are known in humans. Cancer cells are characterized by abnormal proliferation and the potential to invade and spread to other parts of the body. Overall, cells are defined by their gene expression programs

and the oncogenic transformation from a healthy cell to a cancer cell is caused by dysregulation of such finely tuned programs due to genetic and epigenetic alterations. Alterations in the genome can develop as a consequence from exposure to environmental pollutants, UV and ionizing radiation, infections with oncogenic viruses, or inherited gene defects. The rapid growth and proliferation of cancer cells leads to the dependency on high transcriptional turn-over of oncogenic gene expression programs, rendering cancer cells transcriptionally addicted[3].

The molecular events that drive aberrant transcription in cancer are diverse. Reports about oncogenic alterations within genes encoding core components of the pol II transcription machinery are usually specific to individual cancer types[4]. In contrast, key oncogenic drivers of a majority of cancers encompass transcription factors such as the tumor suppressor p53 and oncogene c-Myc, which show altered expression levels in 50-70% of all tumors and are associated with tumor aggression and poor clinical outcome[119]. Because such key transcription factors occupy the majority of open promoters that define cell identity, dysregulation results in genome-wide alterations of gene expression programs, but at the same time render cancer cells particularly sensitive to transcriptional disruption. The rationale for targeting components of the core transcription machinery has been proposed to be less prone to acquire therapeutic resistance[120]. Therefore, SMKIs and TPDs targeting transcriptional key factors such as transcriptional kinases hold promise as an attractive new class in cancer drug development.

1.3.3 Targeting Transcriptional Kinases in Cancer

Although the field of transcriptional regulation opened up by the discovery of RNA pol II more than 50 years ago and since then transcriptional kinases have been studied intensively, the role of many transcriptional kinases remains controversial. A substantial fraction of transcriptional kinases are still understudied as recently highlighted by the National Institute of Health (NIH) Understudied Kinome Consortium and the NIH´s Illuminating the Druggable Genome Program[121,122]. Transcriptional kinases are promising druggable targets of the core transcription machinery and CDKs have been therapeutic targets for 20 years. First-generation CDK inhibitors, such as DRB, roscovitine, and flavopiridol, have been used to study tCDK functions in preclinical cancer models but act promiscuously on many CDKs with a narrow therapeutic window. These inhibitors have shown inadequate efficacy and considerable toxicity, which is related to the universal functions of tCDKs for transcriptional regulation in healthy cells resulting in broad off-target effects[123]. The development of next-generation CDK inhibitors led to the successful development of the selective Cdk4/6 inhibitors palbociclib, abemaciclib, ribociclib and trilaciclib, which are now the standard of care for advanced and metastatic breast cancer and small cell lung cancer[124]. CDK inhibitors continue to be intensively investigated in phase I and II clinical trials, with five novel SMKIs targeting the Cdk4 family and additional 12 SMKIs targeting Cdk7, Cdk8/19, and Cdk2 and Cdk9, either selectively or in combinations with other agents[51]. Combinational therapy may also allow the use of lower concentrations than single agents, potentially reducing toxicity and resistance mechanisms.

Cdk7

Cdk7 activity is critical for transcription initiation as well as for the activation of CDKs as a CDK-activating kinase (CAK). Loss-of-function mutations of Cdk7 have not been widely reported. However, a wide variety of cancers, including T cell acute lymphoblastic leukaemia (ALL), neuroblastoma and triple-negative breast cancer, show elevated Cdk7 occupancy and activity at tumour-specific super-enhancer elements[112]. Cdk7 inhibition by the covalent inhibitor THZ1 has demonstrated the selective dependency of these cancers on Cdk7 activity[125], but potent Cdk12/13 off-target activity obscured Cdk7 contribution to this phenotype. A derivative of THZ1 with higher selectivity is YKL-5-124 that induces a strong cell-cycle arrest and a surprisingly weak effect on pol II phosphorylation[126]. Phase I clinical trials of the Cdk7 inhibitor SY-1365 are ongoing[112].

Cdk8

Mutations in Cdk8 and other components of Mediator such as Med12 are frequently observed at super-enhancer elements in various tumours such as Uterine leiomyomas, colon cancer, melanoma, and prostate cancer[4]. Cdk8 has been shown to possess both, tumour-suppressive and oncogenic functions in different cellular contexts. Overall, altered Cdk8 levels or activity have minimal impact on global gene expression but rather selectively impact the transcription of genes controlled by specific transcription factors such as E2F1 and STAT1. Selective Cdk8 inhibition by the natural product cortistatin A leads to upregulation of SE-associated genes and apoptosis in acute myeloid leukemia (AML)[127]. Conversely, the small-molecule inhibitor CCT251545 induces downregulation of WNT pathway-regulated genes with efficacies in WNT-driven breast and colorectal cancer models. An optimized derivative is currently in preclinical trials[112].

Cdk9

The pause-release kinase Cdk9 is globally essential for pol II progression through the transcription cycle so that no Cdk9 loss-of-function mutations are known in cancer. On the contrary, many cancer cells show aberrant recruitment of P-TEFb to oncogenic key loci and super-enhancer regions and are thus disproportionally sensitive to Cdk9 inhibition. The oncogenic transcription factor c-Myc occupies the majority of open promoters and physically interacts with P-TEFb, recruits it to TSS proximal regions and thus promotes the release of stalled pol II from essentially the entire active gene expression program in diverse cell types. Therefore, c-Myc globally amplifies the output of existing gene expression programs[38,119].
In addition to activating mutations, chromosomal rearrangements are another cause of elevated kinase activity as these can lead to oncogenic fusion chimeras, f.e. BCR-ABL[128]. Another example is the genetic rearrangement of the mixed lineage leukaemia (MLL) gene that results in fusion chimaeras important in infant B cell ALL, AML, and adult myeloid leukaemia (ADML). In MLL fusion chimaeras the chromatin-binding domain of MLL gets fused to components of the super-elongation complex (SEC) such as AF4/9/10, and ENL, aberrantly recruiting Cdk9 and Brd4 to MLL target genes, ultimately promoting elevated elongation rates through excessive pol II pause-release. MLL target genes encode critical regulators of haematopoietic progenitor cell programs such as c-Myc, BCL2, MCL1, and CDK6, which in turn contribute to aggressive acute leukemia[4].

Cdk9 inhibitors described to date, f.e. flavopiridol, show typically high levels of poly-pharmacology and suffer from dose-limiting toxicity in the bone marrow and the gastrointestinal tract in clinical trials[112]. NVP2 and iCDK9 are more selective Cdk9 inhibitors[129], which cause genome-wide pol II pausing in cells, whereas the only selective Cdk9 inhibitor in clinical trials is BAY1251152.

Cdk12 and Cdk13

Many cancer types have been shown to contain loss-of-function mutations of Cdk12, including high-grade serous ovarian cancers, triple negative breast cancer, gastric, intestinal, and lung cancer and metastatic castration-resistant prostate cancer[130–132]. Cdk12 mutated cancers typically exhibit `BRCAness´ phenotypes characterized by decreased transcription of longer genes, such as genes involved in DNA damage-response (DDR) and homologous recombination (HR)[17,133,134], which leads to the elevated expression of short transcripts, sensitization of cancer cells to poly(ADP-ribose) polymerase (PARP) inhibitors[135], as well as genomic instability characterized by genome-wide focal tandem duplications. Nonetheless, loss of Cdk12 activity does not results in global elongation defects, which might results from the redundancy between Cdk12 and Cdk13[136]. Mutations of Cdk13 are less well studied but have been associated with poor prognosis in hepatocellular carcinoma[137].
The covalent and selective Cdk12 and Cdk13 inhibitor THZ531 downregulates DDR genes already at low doses and super-enhancer associated driver genes in T-ALL cells[138]. In high grade serous ovarian cancer and triple-negative breast cancer, THZ531 effectively synergizes with PARP inhibition. Dinaciclib, a pan-CDK inhibitor, disrupts transcription of HR genes and reverses PARP inhibitor resistance in triple-negative breast cancer models via Cdk12 inhibition[139]. Thus, combined inhibition of PARP and Cdk12 could expand current limitations of both single inhibitors.

DYRK1A and HIPKs

Due to its roles in cell cycle and DNA damage response, DYRK1A has oncogenic and tumor-suppressor functions depending on the cellular context. A growing body of literature links DYRK1A to glioblastoma, head and neck squamous cell carcinoma, pancreatic ductal adenocarcinoma, hepatocellular carcinoma, ovarian cancer, acute megakaryoblastic leukemia (AMKL), and ALL[140]. Similarly, HIPKs dysregulation is involved in the progression of AML, prostate carcinoma, colorectal cancer, skin carcinogenesis and thymic lymphoma[141]. Selective DYRK1A or HIPKs inhibitors are still under development.

Although there have been substantial advances in kinase drug discovery, there are still many challenges. At least 70% of the kinome (approx. 400 kinases) is still unexplored in drug development in the sense that no compounds targeting these kinases have entered clinical trials[51]. Additionally, the emergence of resistance is a major barrier to achieving long-term remission in cancer. Further studies on transcriptional kinases will help to overcome these current issues.

1.4 Research objectives

Eukaryotic transcription is tightly controlled by transcriptional kinases phosphorylating the C-terminal domain of RNA pol II and other key regulators throughout the distinct phases of the transcription cycle. Transcriptional addiction of a broad range of cancer types highlights the importance of a better understanding of transcriptional kinases and will aid the development of novel kinase inhibitory drugs.

The first part of this thesis addresses functional and structural issues of the poorly studied HIP kinase family. A strategy for expression and purification of all four HIPKs in their recombinant active form will be elaborated. Crystallization and structure determination of HIPK3 would provide a better understanding of the fundamental structural features of the HIPK family. Strong similarities with DYRK1A suggest a role of HIP kinases in the direct phosphorylation of general components of the transcription machinery, which will be investigated by in vitro kinase assays. Based on a database screen, putative HIPK3 inhibitors will be identified, followed by further characterization of the best hits and co-crystallization, enabling the structure-guided optimization for selective HIP kinase inhibitors.

The second part of this thesis focuses on tCDKs and their activation by cognate and novel Cyclin partners. Analytical size exclusion chromatography will be used as an unbiased approach to systematically test complex formation between tCDKs and Cyclins. Kinase activities, substrate specificity and interaction with known regulators, such as Hex1 and Brd4 will be analyzed in radioactive kinase assays. At the gene level novel CDK/Cyclin pairs will be characterized by ChIP-seq, as well as by co-immunoprecipitation and Knock-out mutations in cells. In addition, crystallization studies will help to understand similarities and differences in Cyclin interaction. Given the increasing interest in tCDK inhibitor development for cancer therapy, the inhibitory potential of novel small-molecular compounds will be characterized in terms of binding affinities, inhibitory efficacy and by co-crystallization to determine differences in their target profiles for cognate and novel CDK/Cyclin pairs.

The results of this thesis will contribute to a better understanding of these understudied transcriptional kinases and pave the way for the development of more selective targeted therapies in the field of oncology and beyond.

2. MATERIAL & METHODS

2.1 Material

2.1.1 Devices

Device	Type	Manufacturer
Agarose Gel Chamber	Sub® Cell GT (Mini, Midi, Maxi)	Biorad, CA, USA
Autoclave	5075 EL	Systec, Wettenberg, Germany
Balances	CPA 324S	Sartorius, Göttingen, Germany
	BP110S	Sartorius, Göttingen, Germany
	PCD	Kern, Balingen-Frommern, Germany
Blotting Device	Semi-dry Blotter V20 SDB	SciePlas, Cambridge, UK
Cell Culture Bench	HeraSafe	Heraeus Instruments, Hanau, Germany
Cell Counter	Countess II FL automated cell counter	Thermo Fisher Scientific, MA, USA
Centrifuges	Avanti J-26S XP	Beckman Coulter, CA, USA
	5804	Eppendorf, Hamburg, Germany
	table top 3424	Eppendorf, Hamburg, Germany
Chromatography Systems (FPLC)	Äkta Start	Cytiva, Opfikon, Switzerland
	Äkta Prime Plus	Cytiva, Opfikon, Switzerland
	Äkta Pure	Cytiva, Opfikon, Switzerland
	Äkta Avant	Cytiva, Opfikon, Switzerland
Cooling Cabinett	Unichromat 1500	UniEquip, Martinsried, Germany
Crystal Equipment	Rock Imager 1000	Formulatrix, Dubai, United Arab Emirates
	Griffon	Art Robbins Instruments, CA, USA
	epMotion 5070	Eppendorf, Hamburg, Germany
Fluorometer	Qubit 4 Fluorometer	Thermo Fisher Scientific, MA, USA
Gel Documentation System	ChemiDocXRS+	Biorad, CA, USA
Heat Steriliser	Typ ST6120	Heraeus Instruments, Hanau, Germany
HPLC	Agilent 1260 infinity 1260 bio-inert HPLC System	Agilent Technologies Inc., CA, USA

Incubator	Multitron	Infors HT, Bottmingen, Switzerland
Liquid Scintillation Counter	LS 6500	Beckman Coulter, CA, Usa
Magnetic Rack	DynaMag-2 magnet	Thermo Fisher Scientific, MA, USA
Magnetic Stirrer	MR 3002	Heidolph, Schwabach, Germany
Microscope	Eclipse TS100	Nikon, Tokio, Japan
Microwave	Privileg 8018	Privileg, Stuttgart, Germany
Nanodrop	Nanodrop 2000c	Thermo Fisher Scientific, MA, USA
Nano DSF	Prometheus NT.48	NanoTemper Technologies GmbH, Munich, Germany
Nucleofector	4D-Nucleofector X Unit	Lonza, Basel, Switzerland
PAGE Electrophoresis Chamber	MiniProtean III	Biorad, CA, USA
PAGE Pouring Station	MiniProtean	Biorad, CA, USA
PCR Cycler	Mastercycler Nexus SX1	Eppendorf, Hamburg, Germany
pH Meter	Lab 850	Xylem Analytics, Mainz, Germany
Power Supply	Basic	Biorad, CA, USA
Shaker, rolling	RM5-30V CAT	NeoLab, Heidelberg, Germany
Sonifier	Sonics	Vibra Cell, CT, USA
	ME220	Covaris, LLC., MA, USA
Sonifier Waterbath	Bandelin Sonorex	Sonorex Digitec, Berlin, Germany
Surface Plasmon Resonance	Biacore 8K	Cytiva, Opfikon, Switzerland
Thermo Mixer	Thermomixer comfort	Eppendorf, Hamburg, Germany
UV-Spectrophotometer	Nanodrop 2000C	Thermo Fisher Scientific, MA, USA
UV-table	Benchtop 2UV Transilluminator	UVP, Cambridge, UK
Vortex Mixer	Vortex Genie 2	Bender&Hobein, Bruchsal, Germany
Waterbath	Julabo 5	Julabo, Seelbach, Germany
White Light Table	White light transilluminator	UVP, Cambridge, UK

2.1.2 Chemicals & Reagents

All chemicals were of analytical grade and purchased from the following suppliers: AppliChem (Darmstadt, Germany), Carl Roth GmbH (Karlsruhe, Germany), Cytiva (Uppsala, Sweden),

Merck (Darmstadt, Germany), Sigma Aldrich (München, Germany) and Thermo Fisher Scientific Inc. (MA, USA).

2.1.3 Consumables

Consumable	Type	Manufacturer
Bottle Top Filter	Whatman Membrane Filters (0.2 µm, 0.4 µm)	Thermo Fisher Scientific, MA, USA
Cell Culture Supply	Cell culture plates (10 cm, 15 cm) and flasks (T75, T175)	Sarstedt AG & Co.KG, Nümbrecht, Germany
Chromatography Supply	Protino columns (12 ml, 35 ml)	Macherey-Nagel, Düren, Germany
	HisTrap FF	Cytiva, Opfikon, Switzerland
	GSTrap FF	Cytiva, Opfikon, Switzerland
	MBPTrap HP	Cytiva, Opfikon, Switzerland
	StrepTrap HP	Cytiva, Opfikon, Switzerland
	SP IEX	Cytiva, Opfikon, Switzerland
	Q IEX	Cytiva, Opfikon, Switzerland
	HiLoad 16/60 Superdex 75 pg	Cytiva, Opfikon, Switzerland
	HiLoad 16/60 Superdex 200 pg	Cytiva, Opfikon, Switzerland
	Superdex75 increase 10/300 GL	Cytiva, Opfikon, Switzerland
	Superdex200 increase 10/300 GL	Cytiva, Opfikon, Switzerland
	Superose6 increase 10/300 GL	Cytiva, Opfikon, Switzerland
	SuperdexS200 increase 3.2/300 GL	Cytiva, Opfikon, Switzerland
	Amylose Resin Hi Flow	New England Biolabs, MA, USA
	Pierce Glutathion Agarose	Thermo Fisher Scientific, MA, USA
	His Pur Nickel-NTA Resin	Thermo Fisher Scientific, MA, USA
	Dynabeads Protein A or G	Thermo Fisher Scientific, MA, USA
Concentrators	Amicon Ultra Centrifugal Devices (MWCO: 3K, 10K, 30K, 50K, 100K)	Merck Millipore, MA, USA
	Vivaspin 500 (MWCO: 3K, 10K, 30K, 50K, 100K)	Sartorius, Göttingen, Germany
Crystallization Consumables	Nylon loops	Hampton Research, CA, USA
	JBScreen JCSG++	Jena Bioscience, Jena, Germany

	MemGold2	Molecular Dimension, OH, USA
	ProPlex	Molecular Dimension, OH, USA
	The Ligand Friendly Screen	Molecular Dimension, OH, USA
Desalting Columns	PD-10	Cytiva, Opfikon, Switzerland
Dialysis Tubing	SpectraPor (MWCO: 6-8K)	Spectrum Laboratories, CA, USA
Membranes	Amersham Protran	Cytiva, Opfikon, Switzerland
	Optitran Nitrocellulose	
	Immobilon-FL polyvinylidene fluoride membranes	Merck Millipore, MA, USA
Pipette Tips	10 µl, 20 µl, 200 µl, 1000 µl	Sarstedt AG & Co.KG, Nümbrecht, Germany
Reaction Tubes	Eppis (0.5 ml, 1 ml, 2 ml, 5 ml)	Eppendorf, Hamburg, Germany
	Falcons (15 ml, 50 ml)	Greiner bio-one, Frickenhausen, Germany
SPR Consumables	Sensor chip (CM5)	Cytiva, Opfikon, Switzerland
	Amine coupling kit	Cytiva, Opfikon, Switzerland
Stripettes	5 ml, 10 ml, 25 ml, 50 ml	Sarstedt AG & Co.KG, Nümbrecht, Germany
Syringe Filter	Filtropur S (0.2 µm, 0.4 µm, 0.8 µm)	Sarstedt AG & Co.KG, Nümbrecht, Germany

2.1.4 Kits & Standards

Kits & Standards	Type	Manufacturer
ChIP Kit	ChIP DNA Clean & Concentrator kit	Zymo research, CA, USA
DNA Ladder	(100 bp, 1 kbp)	Roth, Karlsruhe, Germany
DNA Prep Kit	ExtractMe (Mini, Midi, Maxi Prep)	Blirt, Gdańsk, Poland
DNA Purification Kit	ExtractMe	Blirt, Gdańsk, Poland
DNA/RNA Dye	PeqGreen	PeqLab Biotechnologies, Erlangen, Germany
	SYBR Safe	Thermo Fisher Scientific, MA, USA
ECL Solution	ECL Western Blotting Detection Reagent	Cytiva, Opfikon, Switzerland
	SuperSignal West Pico Plus Chemiluminescent Substrate	Thermo Fisher Scientific, MA, USA
Mycoplasma detection Kit	Plasmo Test	Invivogen, CA, USA
Nucleofection Kit	SF cell line 4D-nucleofector X kit	Lonza, Basel, Switzerland

Protein Quantification Kit	Pierce BCA protein assay kit	Thermo Fisher Scientific, MA, USA
Protein Marker	PageRuler Plus Prestained Protein Ladder	Thermo Fisher Scientific, MA, USA
Qubit Kit	Qubit dsDNA HS assay kit	Thermo Fisher Scientific, MA, USA
RNAeasy Micro Kit	Qiagen	Qiagen, Hilden, Germany

2.1.5 Enzymes

Enzyme	Manufacturer
Alt-R SpCas9 nuclease V3	Integrated DNA Technologies, CA, USA
Benzonase Nuclease	Merck Millipore, MA, USA
Cre Recombinase	New England Biolabs, MA, USA
DNaseI	AppliChem, Darmstadt, Germany
Lysozyme	Roth, Karlsruhe, Germany
Proteinase K	Promega, WI, USA
ProtoScript First Strand cDNA Synthesis kit	New England Biolabs, MA, USA
Q5 DNA polyermase	New England Biolabs, MA, USA
Restriction Endonucleases Type II	New England Biolabs, MA USA
Shrimp Phosphatase	New England Biolabs, MA, USA
T4 DNA Ligase	New England Biolabs, MA, USA
TEV Protease	Homemade (Geyer Lab)

2.1.6 Antibodies

Antibodies	Host	Dilution	Manufacturer
Primary Antibodies			
Actin (PA5-16914)	rabbit	1:10,000 (blot)	Thermo Fisher Scientific, MA, USA
Actin (A2228)	mouse	1:10,000 (blot)	Merck Millipore, MA, USA
Cdk9 (2316S)	rabbit	5 ug (ChIP) 1:1,000 (blot)	Cell Signaling Technology, MA, USA
Cdk9 (sc-13130)	mouse	200 ug/ml (blot)	Santa Cruz, TX, USA
Cdk12 (11973S)	rabbit	5 ug (ChIP) 1:1,000 (blot)	Cell Signaling Technology, MA, USA
Cdk13 (A301-458A)	rabbit	5 ug (ChIP) 1:5,000 (blot)	Thermo Fisher Scientific, MA, USA
CycK (A301-939A)	rabbit	5 ug (ChIP) 1:5,000 (blot)	Thermo Fisher Scientific, MA, USA

CycK (ab85854)	rabbit	1:5,000 (blot)	Abcam, Cambridge, UK
CycT1 (81464S)	rabbit	5 ug (ChIP) 1:1,000 (blot)	Cell Signaling Technology, MA, USA
pol II (CTD4H8)	mouse	5 ug (ChIP)	Merck Millipore, MA, USA
pol II CTD pSer2 (3E8)	rat	1:100 (blot)	Prof. Dirk Eick, Helmholtz Center Munich, Germany
pol II CTD pSer2 (ab5095)	rabbit	1:1,000 (blot)	Abcam, Cambridge, UK
pol II CTD pSer5 (3E10)	rat	1:100 (blot)	Prof. Dirk Eick, Helmholtz Center Munich, Germany
pol II CTD pSer5 (ab5131)	rat	1:5,000 (blot)	Abcam, Cambridge, UK
pol II CTD pSer7 (4E12)	rat	1:100 (blot)	Prof. Dirk Eick, Helmholtz Center Munich, Germany
pol II CTD pThr4 (6D7)	rat	1:100 (blot)	Prof. Dirk Eick, Helmholtz Center Munich, Germany
pol II CTD pTyr1 (3D12)	rat	1:100 (blot)	Prof. Dirk Eick, Helmholtz Center Munich, Germany
Thiophosphate ester (ab133473)	rabbit	1:5,000 (blot)	Abcam, Cambridge, UK
Secondary Antibodies			
Rat IgG, HRP-coupled	chicken	1:10,000	Santa Cruz, TX, USA
Rabbit IgG, HRP-coupled	goat	1:10,000	Thermo Fisher Scientific, MA, USA
Mouse IgG, HRP-coupled	goat	1:10,000	Thermo Fisher Scientific, MA, USA
IRDye 800 CW anti-rabbit IgG	goat	1:20,000	Li-COR Inc., NE, USA
IRDye 680 CW anti-mouse IgG	goat	1:20,000	Li-COR Inc., NE, USA
IgG controls			
IgG Isotype Control (31235)	rabbit		Thermo Fisher Scientific, MA, USA
IgG Isotype Control (DA1E)	rabbit		Cell Signaling Technology, MA, USA

2.1.7 Nucleotides

Nucleotide	Manufacturer
ATP	Jena Bioscience, Jena, Germany
ADP	Jena Bioscience, Jena, Germany
ATP-γ-S	Jena Bioscience, Jena, Germany
[^{32}P]-γ-ATP	Perkin Elmer, MA, USA

dNTPs	Roche, Mannheim, Germany
Oligonucleotides	Metabion International AG, Planegg/Steinkirch, Germany
sgRNAs	Synthego Corporation, CA, USA

2.1.8 Vectors

Vector backbones were modified with an N- or C-terminal His-, MBP-, GST-, or Strep-tag, respectively, followed by a Tobacco Etch Virus (TEV)-protease cleavage site.

Vector backbone	Organism	Manufacturer
pET28a	*E.coli*	GE Healthcare, Munich, Germany
pGEX	*E.coli*	GE Healthcare, Munich, Germany
pProExHtb	*E.coli*	GE Healthcare, Munich, Germany
pFastBac	Insect cells	Thermo Fisher Scientific, MA, USA
pACEBac1	Insect cells	ATG Biosynthetics, Merzhausen, Germany
pIDK	Insect cells	ATG Biosynthetics, Merzhausen, Germany
pIDC	Insect cells	ATG Biosynthetics, Merzhausen, Germany

2.1.9 Synthetic peptides

Peptide	Sequence	Manufacturer
Cons. $CTD_{[3]}$	YSPTSPS YSPTSPS YSPTSPS-[PEG2]-RR-amid	Biosyntan, Berlin, Germany
pY1 $CTD_{[3]}$	PS pYSPTSPS pYSPTSPS pYSPTSPS-[PEG2]-RR-amid	Biosyntan, Berlin, Germany
pS2 $CTD_{[3]}$	S YpSPTSPS YpSPTSPS YpSPTSPS-[PEG2]-RR-amid	Biosyntan, Berlin, Germany
pT4 $CTD_{[3]}$	S YSPpTSPS YSPpTSPS YSPpTSPS Y-[PEG2]-RR-amid	Biosyntan, Berlin, Germany
pS5 $CTD_{[3]}$	YSPTpSPS YSPTpSPS YSPTpSPS-[PEG2]-RR-amid	Biosyntan, Berlin, Germany
pS7 $CTD_{[3]}$	ac-YSPTSPpS YSPTSPpS YSPTSPpS Y-[PEG2]-RR-amid	Biosyntan, Berlin, Germany
K7 $CTD_{[3]}$	YSPTSPK YSPTSPK YSPTSPK-[PEG2]-RR-amid	Biosyntan, Berlin, Germany

2.1.10 Buffers

Buffers	Components
PBS	20 mM Na_2HPO_4 pH 7.5, 4.6 mM NaH_2PO_4, 150 mM NaCl
PBS-T	PBS with 0.05% (v/v) Tween20

TBS	50 mM Tris-HCl pH 8.0, 150 mM NaCl
TBST-T	TBS with 0.05% (v/v) Tween20
TAE	40 mM Tris-HCl pH 8.5, 20 mM acetic acid, 1 mM EDTA
SDS-PAGE running buffer	25 mM Tris (pH not adjusted), 194 mM Glycine, 0.1% (w/v) SDS
SDS-PAGE sample buffer (6x)	62.5 mM Tris-HCl pH 6.8, 25% (v/v) Glycerol, 20% (v/v) β-mercaptoethanol, 1.6% (w/v) SDS, 0.04% (w/v) Bromophenol blue
SDS-PAGE Coomassie staining solution	40% (v/v) EtOH, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Brilliant Blue R-250
SDS-PAGE Coomassie destaining solution	10% (v/v) EtOH, 5% (v/v) acetic acid
SDS-PAGE separating gel buffer	1.5 M Tris-HCl pH 8.8, 0.4% (w/v) SDS
SDS-PAGE stacking gel buffer	0.5 M Tris-HCl pH 6.8, 0.4% (w/v) SDS
Separating gel (12%)	6.8 ml ddH_2O, 8 ml 30% (v/v) Acrylamide/Bis, 5 ml Separating gel buffer, 100 µl 10% (v/v) APS 10 µl TEMED
Stacking gel	2.7 ml ddH_2O, 1 ml 30% (v/v) Acrylamide/Bis, 1.25 ml Stacking gel buffer, 25 µl 10% (v/v) APS, 2.5 µl TEMED
Semi-dry blotting buffer	48 mM Tris (pH not adjusted), 39 mM glycine, 0.04% (v/v) SDS, 20% (v/v) methanol

2.1.11 Cell lines

Cell Line	Genotype/Origin
***E.coli* Strains**	
Beta10	Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- ϕ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (Str^R) rph spoT1 Δ(mrr-hsdRMS-mcrBC)
DH5alpha	F^- ϕ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17(r_K^-, m_K^+) phoA supE44 λ^-thi-1 gyrA96 relA1
PIR	F^-Δlac169 rpoS(am) robA1 creC510 hsdR514 endA recA1 uidA(ΔMluI)::pir-116
DH10Bac	F^-mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara-leu)7697 galU galK λ^-rpsL nupG / bMON14272 / pMON7124
BL21 (DE3) pLysS	F^-ompT $hsdS_B$ (r_B^-, m_B^-) gal dcm (DE3) pLysS(Cam^R)
Rosetta™ (DE3)	F^- ompT $hsdS_B$(r_B^- m_B^-) gal dcm (DE3) pRARE (Cam^R)
Insect cell lines	
Spodoptera frugiperda 9 (*Sf9*)	(Vaughn et al., 1977)[142], Thermo Fisher Scientific, MA, USA
Human cell lines	

A375	University Clinics Bonn, Germany
HEK293T	University Clinics Bonn, Germany
THPI	Peter MacCallum Cancer Centre, Melbourne, Australia
THPI FUCCI	Peter MacCallum Cancer Centre, Melbourne, Australia

2.1.12 Media & Antibiotics

All *E.coli* media were autoclaved before usage.

E.coli Media	**Components**
LB	10 g/l Bacto™ tryptone 5 g/l Bacto™ yeast extract 10 g/l NaCl
TB	12 g/l Bacto™ tryptone 24 g/l Bacto™ yeast extract 4 ml Glycerin 2.31 g KH_2PO_4 12.54 g K_2PO_4
LB agar	1.5% (w/v) Agar in LB

Insect cell Media	**Manufacturer**
Sf-900 III SFM	Thermo Fisher Scientific, MA, USA

Human tissue-culture Media	**Manufacturer**
Dulbecco's Modified Eagle Medium (DMEM)	Thermo Fisher Scientific, MA, USA
Fetal bovine serum (FBS)	Thermo Fisher Scientific, MA, USA
Penicillin-Streptomycin	Thermo Fisher Scientific, MA, USA
Roswell Park Memorial Institute (RPMI) 1640 Medium	Thermo Fisher Scientific, MA, USA
TrypLE Select Enzyme	Thermo Fisher Scientific, MA, USA
Trypsin-EDTA, phenol-red	Thermo Fisher Scientific, MA, USA

Antibiotics	**Final Concentration**
Ampicillin	100 µg/ml (diluted in H_2O)
Chloramphenicol	10 µg/ml (diluted in EtOH)
Gentamycin	10 µg/ml (diluted in H_2O)
Kanamycin	50 µg/ml (diluted in H_2O)
Streptomycin	10 µg/ml (diluted in H_2O)
Spectinomycin	50 µg/ml (diluted in H_2O)
Tetracyclin	10 µg/ml (diluted in Isopropanol)
X-Gal, BCIG (5-bromo-4-Chloro-3-indolyl-B-D-galactopyranoside)	100 µg/ml (diluted in EtOH)

2.1.13 Kinase inhibitors

Small-molecular compounds and Targeted protein degraders were dissolved in DMSO and stored as aliquots at -20°C to avoid freeze-thaw cycles. The small-molecular compounds flavopiridol, (R)-roscovitine, staurosporine, SCH727965 (dinaciclib), PD0332991 (palbociclib), LY2835219 (abemaciclib), GSK1059615, AS601245, and NVP2 were purchased from MedChemExpress. Compounds XMD8-70, XMD8-62-i, JWD-065, HTH-01-091, CVM-05-145-3, and CVM-06-033-2 were developed and provided by the Gray laboratory, Dana-Farber Cancer Institute, MA, USA. The targeted protein degraders CR8 and THAL-SNS-032 were purchased from MedChemExpress, and dCEMM2 was provided by the Winter laboratory, CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna Austria.

2.1.14 Software

Software	Version	Manufacturer
Agilent OpenLab CDS Software	22.2.1	Agilent Technologies Inc., CA, USA
Biacore Insight Evaluation Software	3.0.12	Cytiva, Opfikon, Switzerland
COOT	1.0.0	XQuartz 2.8.0_beta3, X.org Foundation, Inc.
Geneious	8.1.9	Biomatters, Auckland, New Zealand
GraphPad Prism	9.4.1 v.7	Software MacKiev, GraphPad Software, LLC.
ImageJ	1.52q	Wayne Rasband, NIH, USA
Image Lab	1.6.0	BioRad Laboratories, Inc., CA, USA
Integrative Genomics Viewer	2.14.1	Broad Institute and the Regents of the University of California, CA, USA
PDBePISA	1.52	EMBL-EBI, Cambridgeshire, UK
PR.ThermControl	2.1.6	NanoTemper Technologies GmbH, Munich, Germany
PyMOL Molecular Graphics System	2.2.0	Schrödinger, LLC
Rock Imager	3.7.1.18	Formulatrix, Dubai, United Arab Emirates
Rock Maker	3.13.3.1	Formulatrix, Dubai, United Arab Emirates
Scaffold	5.1.1	Thermo Fisher Scientific, MA, USA

2.2 Methods

2.2.1 Molecular biological methods

2.2.1.1 cDNA Synthesis

HEK293T or THP1 cDNA was used as a template for cloning of human genes. For cDNA synthesis, total RNA was isolated from HEK293T or THP1 cells using an RNAeasy Micro kit (Qiagen) according to the manufacturer´s instructions, followed by reverse transcription to cDNA using a ProtoScript First Strand cDNA Synthesis kit (New England Biolabs).

2.2.1.2 Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) was performed to amplify or modify DNA fragments for cloning. Primers were designed for melting temperatures of 54-56°C as calculated by Geneious (Biomatter) and synthesized by metabion (Metabion International AG). PCR reactions were set up as shown in Tab.1 and incubated following the PCR cycles shown in Tab.2 using a Mastercycler nexus (Eppendorf). PCR reaction mixtures were subjected to agarose gel electrophoresis to control for successful amplification and purified from gels using the ExtractMe DNA Purification Kit (Blirt) according to the manufacturer´s instructions.

Tab.1 | PCR reaction mixture

Component	final Volume
5x Q5 Reaction Buffer	10 µl
Q5 GC Enhancer	2 µl
dNTPs (10 mM)	4 µl
Primer forward (10 µM)	0.5 µl
Primer reverse (10 µM)	0.5 µl
Template DNA	30 ng
Q5 polymerase	1 µl
ddH_2O	ad 50 µl

Tab.2 | PCR program

Phase	Temperature	Duration
Initial Denaturation:	98°C	30 sec
Amplification:		
5x repeat	98°C	5 sec
	54°C	30 sec
	72°C	30 sec /1kb
25x repeat	98°C	5 sec
	56°C	30 sec
	72°C	30 sec /1kb
Final Extension:	72°C	2 min

2.2.1.3 Site-directed mutagenesis

Point mutations were introduced into genes by site-directed mutagenesis[143] in a PCR reaction followed by digestion of the non-mutated parental plasmid using the methylation-sensitive restriction endonuclease *DpnI* (New England Biolabs), which specifically cuts DNA at GATC sites containing methylated adenine nucleotides. DNA methylation is generated naturally by amplification of the plasmid in *E.coli* cells, while PCR-generated, mutated DNA lacks methylation and is therefore not digested. Primers for site directed mutagenesis consisted of an overlapping and an extended non-overlapping sequence. Melting temperatures for overlapping sequences were designed for 55°C according to Geneious (Biomatter) and of

non-overlapping regions for 45°C. Mutation sites can be placed either in the overlapping or non-overlapping sequence.
PCR reactions were set up as shown in Tab.3 and incubated following the PCR cycles shown in Tab.4 using a Mastercycler nexus (Eppendorf). After PCR, 5 µl of the reaction mixture were subjected to agarose gel electrophoresis to control for successful amplification. The residual sample was digested by addition of 1 µl *DpnI* in Cutsmart Buffer, followed by incubation for 2-4 h at 37°C. Digestion mixtures were purified using the ExtractMe DNA Purification Kit (Blirt) according to the manufacturer´s instructions.

Tab.3 | PCR reaction mixture

Component	final Volume
5x Q5 Reaction Buffer	10 µl
Q5 GC Enhancer	2 µl
dNTPs (10 mM)	4 µl
Primer forward (10 µM)	2.5 µl
Primer reverse (10 µM)	2.5 µl
Template DNA	30 ng
Q5 polymerase	1 µl
ddH_2O	ad 50 µl

Tab.4 | PCR program

Phase	Temperature	Duration
Initial Denaturation:	95°C	5 min
Amplification:		
15x repeat	95°C	1 min
	50°C	1 min
	72°C	10 min
	40°C	1 min
Final Extension:	72°C	30 min

2.2.1.4 Agarose gel electrophoresis

DNA was separated and analyzed by agarose gel electrophoresis. Depending on the size of the DNA fragments, 1-3% (w/v) agarose was heated in TAE-buffer until agarose was melted completely. PeqGreen (PeqLab) or SYBRSafe (Thermo Fisher Scientific) was added to the gel in a 1: 20,000 dilution to visualize DNA bands under UV light. Gel loading dye (New England Biolabs) was added to DNA samples before loading onto the gel. For size determination, 100 bp or 1 kbp DNA ladder (Roth) was used. Agarose gel electrophoresis was performed in TAE buffer at 120 V for 30-45 min using a Mini-, Midi- or Maxi-Sub® Cell GT Cell (BioRad). Gels were visualized and documented under UV light using a ChemDocXRS+ (BioRad). For preparative agarose gels, DNA bands of the respective size were excised from the gel with a

clean scalpel and extracted using the ExtractMe DNA Purification kit (Blirt) according to the manufacturer´s instructions.

2.2.1.5 Restriction digest

For cloning or analysis of plasmid DNA, DNA was digested using specific endonucleases (New England Biolabs) according to the manufacturer´s instructions. For restriction digest, restriction enzymes and Cutsmart Buffer (New England Biolabs) were added to the DNA samples and incubated for 30-120 min. Afterwards, DNA was purified using the ExtractMe DNA Purification kit (Blirt) according to the manufacturer´s instructions.

2.2.1.6 Ligation

For Ligation of DNA, a fivefold molar excess of insert was added to 25 ng of vector DNA and incubated with T4 DNA Ligase (New England Biolabs) in T4 Ligase buffer (New England Biolabs) for 1 hour to overnight at room temperature.

2.2.1.7 Generation of chemically competent *E.coli* cells

To generate chemically competent *E.coli* cells, a pre-culture was inoculated from an agar plate and grown overnight at 37°C at 150 rpm. The next day, the pre-culture was diluted to an optical density (OD_{600}) of 0.25 in 1 L of LB medium supplemented with the respective antibiotics of the *E.coli* strain. Cells were grown to an OD_{600} of 0.6 and harvested by centrifugation at 4,000 g for 10 min at 4°C. The medium was discarded and cell pellets were thoroughly resuspended in 3 ml of ice-cold 0.1 M $CaCl_2$ followed by 20 min incubation on ice. Cells were pelleted by centrifugation at 4,000 g for 10 min at 4°C. The supernatant was discarded and pellets were resuspended in 3 ml of ice-cold 0.1 M $CaCl_2$ and 1.5 ml ice-cold glycerol. Competent cells were aliquoted, directly snap-frozen in liquid nitrogen and stored at -80°C.

2.2.1.8 Transformation of *E.coli* cells

For vector amplification or protein expression, competent *E.coli* strains were transformed with plasmid DNA by heat shock. After thawing competent cells on ice, 1 µl of plasmid DNA was added and cells were subjected to heat shock for 42 sec at 42°C, followed by addition of 800 µl LB medium. For recovery, cells were incubated at 37°C for 30-120 min at 800 rpm in a ThermoMixer C (Eppendorf). Afterwards, bacteria were plated on agar plates containing antibiotics for the selection of transformed cells and incubated overnight at 37°C.

2.2.1.9 Plasmid preparation from *E.coli* cells

Competent beta10 or DH5alpha *E.coli* strains were transformed and grown on agar plates. For small scale DNA amplification, single clones were picked and grown in 2-5 ml LB medium containing the respective antibiotics overnight at 37°C with shaking. Plasmid DNA was purified using the ExtractMe Plasmid Mini Prep kit (Blirt) according to the manufacturer´s instruction.

2.2.1.10 DNA concentration determination

Concentration of DNA was determined by absorbance measurement at 260 nm using a Nanodrop 2000c (Thermo Fisher Scientific).

2.2.1.11 DNA sequencing

The correct nucleotide sequence of all cloned constructs was verified by Sanger sequencing at GATC Biotech (Eurofins Scientific, Ebersberg, Germany) or Microsynth AG (Balgach, Switzerland).

2.2.1.12 Generation of multigene expression vectors

In *Sf9* insect cells, protein expression was performed using the MultiBacTurbo System[144]. The system offers the possibility to generate multi-gene expression vectors by Cre-LoxP fusion of plasmids. This is achieved by a set of donor or acceptor vectors, in which genes of interest are inserted separately. Donor (pIDK, pIDC) and acceptor (pACEBac1) vectors are fused using a Cre-recombinase recognizing LoxP sites to form a multigene fusion vector. Vector fusion was performed according to the manufacturer´s protocol. Vectors were mixed in equal concentrations and Cre recombination was performed at 37°C for 2 h. After *E.coli* transformation and DNA preparation, correct assembly of fusion vectors was analyzed by restriction digest and DNA sequencing.

2.2.2 Cell biological methods

2.2.2.1 Maintenance of *Sf9* insect cell culture

Many protein kinases require post-translational modifications or have specific folding requirements that are not met by expression in bacteria. For most kinases in this study, *Sf9* insect cells from lepidoptera were used as an eukaryotic expression system[142]. *Sf9* insect cells were grown as suspension culture in SF900 III serum-free Medium (Thermo Fisher Scientific) at 27°C at 80 rpm. Cells were kept at 0.5 x 10^6 to 4 x 10^6 cells/ml and cultures were passaged two to three times per week. *Sf9* cells were cultured in sterilized glassware of appropriate size to allow good aeration.

2.2.2.2 Transfection of *Sf9* insect cells with recombinant baculoviruses

Bacmid preparation

Bacmid preparation was preformed using the MultiBac system[144]. Competent DH10 MultiBac^{Turbo} cells were transformed with 100 ng plasmid DNA by heat shock followed by incubation in 1 ml LB medium for 2-4 hours at 37°C at 800 rpm. Cells were selected by blue/white screening on agar plates containing kanamycin (50 µg/ml), gentamycin (10 µg/ml), ampicillin (100 µg/ml), tetracyclin (10 µg/ml), X-Gal (100 µg/ml), and IPTG (40 µg/ml) at 37°C for 1-2 days. White colonies were selected and grown overnight at 37°C in 4 ml LB medium containing kanamycin (50 µg/ml), gentamycin (10 µg/ml), ampicillin (100 µg/ml), and tetracyclin (10 µg/ml). Cells were harvested by centrifugation, resuspended, and lysed using ExtractMe DNA Clean-up kit (Blirt) buffers P1 and P2 according to the manufacturer´s protocol. After neutralization with P3 buffer, lysates were cleared by centrifugation at 20,000 g for 10 min. Supernatant was transferred to a new tube, centrifuged again, transferred to a new tube filled with 800 µl isopropanol and stored at -20 °C.

Virus propagation and transfection of Sf9 cells

Sf9 cells were incubated in 6-well plates at 0.3 x 10^6 cells/ml for at least 15 min prior to transfection to allow cells to become adherent. Bacmid DNA was collected by centrifugation

at 20,000 g for 30 min at 4°C. Isopropanol was removed, precipitated DNA was washed twice with 70% (w/v) ice-cold ethanol and pelleted for 10 min at 20,000 g at 4°C. Ethanol was removed under a sterile cell culture hood. Pellets were air-dried for approximately 10 min followed by resuspension in 100 µl SF900 III serum-free medium per tube. Then, 100 µl medium with 5 µl of Cellfectin (Thermo Fisher Scientific) was added. After incubation for 20 min, the mixture was added to the cells and incubated at 27°C. After 3 days, the virus containing supernatant was collected (V_0) and sterile filtered. V_0 was used for further virus amplification of 20-100 ml of 0.5 x 10^6 cells/ml. Successful infection of *Sf9* cells was monitored by a stop of cell division of transfected cells due to viral infection. Cells were incubated for 3 days at 27°C at 80 rpm and pelleted by centrifugation at 400 g for 10 min. The supernatant was collected (V_1). To obtain higher virus titer, 5% (v/v) V_1 was used to infect 1 x 10^6 cells/ml *Sf9* cells. Cells were incubated for 3 days and pelleted by 400 g for 10 min. Supernatant was collected (V_2). All viral stocks were sterile filtered and stored at 4 °C.

2.2.2.3 Maintenance of human cell culture

Human cell cultures were routinely checked for mycoplasma contamination using Plasmo Test Mycoplasma detection kit (Invivogen). Experiments with human cells were performed in part in the laboratory of Prof. Ricky Johnstone, Department of Gene Regulation at the Peter MacCallum Cancer Centre, Melbourne, Australia. HEK293T and A375 cells were cultured in Dulbecco´s modified Eagle´s medium (DMEM) (Thermo Fisher Scientific) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (Thermo Fisher Scientific), 100 U/ml penicillin, 100 µg/ml streptomycin at 37°C and 10% CO_2. THP-1 AML and THP-1 FUCCI cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 (Thermo Fisher Scientific) supplemented with 10% (v/v) FBS, 100 U/ml penicillin, 100 µg/ml streptomycin at 37°C and 5% CO_2. Human cultures were passaged two to three times per week. For cell passage of adherent cell lines, culture medium was removed, cells were washed once with PBS and incubated with either 0.05% Trypsin/EDTA (Thermo Fisher Scientific) or TrypLE Express Enzyme (Thermo Fisher Scientific) for 5 min, followed by resuspension in an appropriate amount of culture medium. Cell counting was performed with a Countess II FL automated cell counter (Thermo Fisher Scientific).

2.2.2.4 Generation of Knock-out cell lines using CRISPR/Cas9

Short guide (sg) RNAs were designed and synthesized by Synthego. For nucleotransfection, 200 nM of each sgRNA were incubated with 100 pmol of recombinant Alt-R SpCas9 nuclease V3 (Integrated DNA Technologies) for 20 min at room temperature. HEK293T cells (0.5 x 10^6 cells per transfection) were harvested as described (see chapter 2.2.2.3) and pelleted by centrifugation for 4 min at 400 g at room temperature. Cell pellets were washed twice with PBS and resuspended in 20 µl nucleofector solution (16.4 µl SF nucleofector solution + 3.6 µl supplement 1) from the SF cell line 4D-nucleofector X kit (Lonza). The Cas9/sgRNA mix was added to the 20 µl cell suspension and transferred to a 16-well Nucleocuvette strip. After electroporation using the 4D-Nucleofector X Unit (program CM-130, Lonza), cells were transferred to a 96-well plate and incubated in 100 µl prewarmed culture medium for 10 min at 37°C and 10% CO_2. Cells were then transferred to a 6-well plate containing 3 ml of cell culture medium and were incubated overnight under the same culturing conditions. After

expansion, cell lysates were stained by immunoblot with the respective primary antibody to confirm knock-out of the protein of interest.

2.2.3 Protein biochemical methods

A list summarizing the cloning, expression, and purification conditions for all recombinant proteins generated for experiments in this thesis is included in Appendix B.

2.2.3.1 Protein Expression

Protein Expression in bacteria

For large-scale protein expression in bacteria, *E.coli* BL21 (DE3) pLysS were transfected with the respective plasmid of interest by heat-shock (see chapter 2.2.1.8) and grown as a pre-culture in 50-100 ml LB medium supplemented with the respective antibiotics at 37°C at 150 rpm overnight. For induction of expression by isopropyl β-D-thiogalactoside (IPTG), precultures were diluted to an OD_{600} of 0.1 in LB or TB medium supplemented with antibiotics. Cultures were grown at 37°C at 150 rpm to an OD_{600} of 0.6-1.0. Protein expression was induced by addition of 0.4 mM IPTG while the temperature was lowered to 20-30°C for 4 hours to overnight growth. For protein expression by autoinduction, precultures were diluted to an OD_{600} of 0.5 in TB medium supplemented with 17 mM KH_2PO_4, 72 mM K_2HPO_4, 1.5% (w/v) lactose, 0.05% (w/v) glucose, 2 mM $MgSO_4$ and antibiotics. Bacteria were grown at 37°C at 150 rpm to an OD_{600} of 1.0-2.0 and then grown overnight at 20°C. Cells were harvested by centrifugation with a JLA-8.1 rotor in an Avanti J-26S XP centrifuge (Beckman Coulter) at 4,000 g for 20 min. Pellets were washed with PBS, transferred to Falcon tubes followed by centrifugation in an Eppendorf 5804 centrifuge with an S-4-72 rotor (Eppendorf) at 4,000 g for 20 min. Pellets were subjected to cell lysis or snap frozen in liquid nitrogen for storage at -20°C.

Protein Expression in Sf9 insect cells

For protein expression in insect cells, *Sf9* cultures were infected by addition of 2% (v/v) of baculovirus stock V_2 at a cell density of 1.5-2.0 x 10^6 cells/ml. Cells were harvested 3-4 days after infection by centrifugation with a JLA-8.1 rotor in an Avanti J-26S XP centrifuge (Beckman Coulter) at 800 g for 20 min. Pellets were washed with PBS, transferred to Falcon tubes, and then centrifuged in an Eppendorf 5804 centrifuge with an S-4-72 rotor (Eppendorf) at 800 g for 20 min. Pellets were subjected to cell lysis or snap frozen in liquid nitrogen for storage at -20°C.

2.2.3.2 Cell lysis

Cell pellets were resuspended in lysis buffer (50 mM Hepes pH7.6, 500 mM NaCl, 5 mM b-mercaptoethanol; approximately 40 ml lysis buffer per pellet from a 500 ml cell culture) on ice with magnetic stirring followed by sonication (5 sec on/off interval) at 40% amplitude with a standard probe (Vibra Cell) until cells were properly lysed. Cell debris was pelleted by centrifugation at 4°C for 45 min at 50,000 g in a JA-25.50 rotor in an Avanti J-26S XP centrifuge (Beckman Coulter). The supernatant was filtered through a 0.45 µm syringe filter.

2.2.3.3 Protein Purification

Proteins were purified from filtered cell lysate using chromatographic procedures either at room temperature using gravity columns (Marcherey-Nagel) filled with different affinity resins (Cytiva or NEB) or at 4°C using the FPLC systems Äkta Start, Äkta Prime Plus, Äkta Pure or Äkta Avant (Cytiva). After each chromatographic procedure, protein-containing fractions were analyzed by SDS-PAGE (see chapter 2.2.3.9).

Inclusion body purification

After cell lysis in IB resuspension buffer (50 mM Hepes pH 7.6, 100 mM NaCl, 5 mM b-mercaptoethanol) and centrifugation (see 2.2.3.2) supernatant was discarded and cell pellets were resuspended in IB wash buffer (50 mM Hepes pH7.6, 100 mM NaCl, 1 M Urea, 0.5% (v/v) Triton X-100, 5 mM b-mercaptoethanol). All following steps were performed at room temperature. After centrifugation for 30 min at 50,000 g, resuspension and centrifugation were repeated two times. Next, supernatant was discarded and pellets were resuspended in IB resuspension buffer, followed by centrifugation. Pellets were resuspended in IB extraction buffer (50 mM Hepes pH7.6, 100 mM NaCl, 8 M Urea, 5 mM b-mercaptoethanol) under stirring for 1 hour, followed by centrifugation. Supernatant was dialyzed overnight in IB refolding buffer 1 (50 mM Hepes pH7.6, 100 mM NaCl, 4 M Urea, 5 mM b-mercaptoethanol), followed by dialysis on the next day in IB refolding buffer 2 (50 mM Hepes pH7.6, 100 mM NaCl, 10 mM Imidazol, 5 mM b-mercaptoethanol) for 6 hours at 4°C. The sample was filtered through a 0.45 µm syringe filter and applied to affinity chromatography.

Affinity chromatography

For the separation of the protein of interest from other cellular components, affinity chromatography was used to immobilize the affinity-tagged protein on a stationary phase, wash and elute by the addition of a competitive ligand of higher binding efficiency. Affinity tags were introduced into the protein sequence via the plasmid vector. For affinity chromatography of Gluthation-S-transferase (GST)-, Maltose-binding-protein (MBP)-, hexahistidine (His)-, or Streptavidin (Strep)-tagged proteins, GSTrap columns (Cytiva) or Pierce Glutathion Agarose (Thermo Fisher Scientific), MBPTrap columns (Cytiva) or Amylose Resin Hi Flow (New England Biolabs), HisTrap columns (Cytiva) or His Pur Nickel-NTA Resin (Thermo Fisher Scientific), and StrepTrap columns (Cytiva) were used, respectively.
For purification, affinity resins or columns were equilibrated with approx. 5 column volumes (CV) of lysis buffer (50 mM Hepes pH7.6, 500 mM NaCl, 5 mM b-mercaptoethanol) followed by immobilization of the filtered cell lysate to the matrix. After extensive washes with 10 CV of wash buffer (lysis buffer containing 1 M NaCl) to remove unspecific binders, proteins were eluted with 2 to 5 CV elution buffer containing 20 mM reduced glutathione (GSH) for GST-tagged proteins, 20 mM maltose for MBP-tagged proteins, 250 mM imidazole for His-tagged proteins or 20 mM desthiobiotin for Strep-tagged proteins. Fractions containing the protein of interest in high amount and purity were pooled and further subjected to protein concentration, TEV protease digest, and size exclusion chromatography. In case of strong impurity, additional ion exchange chromatography was performed prior to size exclusion chromatography.

Ion exchange chromatography

To remove residual impurities from the protein of interest, anion or cation exchange was performed via Q or SP chromatography columns (Cytiva). Ion exchange (IEX) chromatography is based on the binding of a charged protein sample to an oppositely charged matrix. This binding is electrostatic and reversible. The pH value at which a protein of interest carries no net charge is called isoelectric point (pI) and is depending on the amino acid composition of a protein, so that the pI of each protein is different. When exposed to a pH below its pI, the protein of interest will carry a positive net charge and will bind to a cation exchanger (SP column). At a pH above its pI the protein will carry a negative net charge and will bind to an anion exchanger (Q column). To ensure uniform charge, proteins were dialyzed against a low salt loading buffer (20 mM Hepes pH7.6, 100 mM NaCl, 1 mM TCEP) with a pH below or above the protein´s pI. After loading the protein sample to the matrix, elution was achieved using a gradient over 200 ml with a high salt elution buffer (20 mM Hepes pH7.6, 1 M NaCl, 1 mM TCEP).

Ammonium sulfate precipitation

For highly impure proteins with good yields, ammonium sulfate precipitation can be performed as an additional purification step after cell lysis. The solubility of proteins is dependent on their amino acid composition and thus differs between proteins. Therefore, varying concentrations of ammonium sulfate can be used to alter solvent conditions and to separate cell lysates into fractions of different solubility, leading to increased purity of the protein of interest. For ammonium sulfate precipitation, cell pellets were resuspended in resuspension buffer (50 mM Tris pH 8.5, 150 mM KCl, 0.02% NP-40, 5 mM $MgCl_2$, 1 mM EDTA, 10% glycerol), disrupted by sonication, followed by centrifugation at 50,000 g for 30 min at 10°C. The supernatant was subjected to step-wise addition of 10%, 30% or 50% ammonium sulfate, followed by centrifugation at 50,000 xg for 15 min at 10°C after each step. After each centrifugation step, pellets were resuspended in resuspension buffer without salt, which was adjusted to a concentration of 150 mM NaCl after complete resuspension. The single fractions were analyzed for the protein of interest by SDS-PAGE and further subjected to affinity chromatography and size exclusion chromatography.

Size exclusion chromatography

Size exclusion chromatography was used as a final purification step to separate proteins from residual substances. For size exclusion chromatography, proteins were loaded onto a gel matrix that separates molecules according to their molecular weight, so that molecules of higher molecular mass pass the matrix faster than particles with lower molecular weight. The chromatography columns used were Superdex 75pg or 200pg increase 10/300 or 16/600, or Superose6 10/300 or 16/600 (Cytiva). Columns were equilibrated with 1 CV size exclusion chromatography buffer (20 mM Hepes pH7.6, 150 mM NaCl, 1 mM TCEP) followed by injection of the concentrated protein sample via an injection loop. Fractions containing the protein of interest were pooled and concentrated. Proteins were aliquoted, snap frozen in liquid nitrogen, and stored at -80°C.

2.2.3.4 Buffer exchange of protein solutions

Buffer exchange of protein solutions was performed with dialysis tubings (Spectrum Laboratories) or PD-10 desalting columns (Cytiva) according to the manufacturers´ instructions.

2.2.3.5 Concentration of Protein Samples

Protein solutions were concentrated by ultrafiltration using Amicon Ultra Centrifugal Filter Devices (Merck Millipore) or Vivaspin 500 (Sartorius) with a molecular weight cut-off of 3, 10, 30, 50 or 100 kDa according to the manufacturers´ instructions. In this application, protein solutions are filtered through a cellulose membrane. Dependent on the pore size, molecules of a certain sizes are retained, while smaller molecules pass through the filter membrane.

2.2.3.6 Determination of Protein Concentration

The concentration of protein samples was determined according to Beer´s law by measuring the absorbance at 280 nm (A_{280}) with a Nanodrop 2000c (Thermo Fisher Scientific). Contaminations with DNA or RNA were monitored by the A260/280 ratio, with an ideal ratio of approx. 0.6 for pure protein samples. Exact protein concentrations were calculated from the absorbance using a theoretical absorbance coefficient and the molecular mass of the protein. The molar extinction coefficient was calculated from the specific protein sequence using Geneious.

2.2.3.7 Protease digestion

For affinity tag removal, recombinant proteins carried a specific recognition site for tobacco etch virus (Tev) protease, which was introduced into the protein sequence via the plasmid vector. The Tev protease is a 27 kDa catalytic domain of the Nuclear Inclusion a (Nia) protein encoded by the tobacco etch virus (Tev), which specifically cleaves after the recognition motif ENLYFQ. For Tev-protease digestion, homemade His-tagged TEV-protease was added to the protein sample in a 1:50 molar ratio and incubated overnight at 4°C under rotation. The protease was removed afterwards by size exclusion chromatography or reverse His-affinity chromatography.

2.2.3.8 Protein storage

For storage of recombinant proteins, concentrated samples were aliquoted, snap-frozen in liquid nitrogen to avoid the formation of ice crystals and stored at -80°C. Freeze-thaw cycles were generally avoided.

2.2.3.9 SDS-PAGE

Protein samples were analyzed by discontinuous Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with a Tris-HCl/Tris-Glycine buffer system. Homemade SDS-polyacrylamide gels were casted according to the molecular weight of the protein of interest with acrylamide concentrations between 4-20% using the MiniProtean III system (BioRad). Samples were mixed with 6x SDS sample buffer (100 mM Tris-HCl pH 6.8, 50% (v/v) glycerol, 4% SDS (v/v), 5% (v/v) b-mercaptoethanol, 0.01% (w/v) bromophenol blue) and incubated at 95°C for 5 min before loading the gel. Electrophoresis was performed in SDS running buffer (25 mM Tris (pH not adjusted), 194 mM Glycine, 0.1% (w/v) SDS) at 30 mA per gel for approx.

40 min. PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific) was used as a molecular weight standard.

Coomassie staining

After gel electrophoresis, gels were shortly boiled in Coomassie brilliant blue staining solution (40% (v/v) EtOH, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Brilliant Blue R-250) and subsequently incubated for 5 min followed by incubation in Coomassie destaining solution (10% (v/v) EtOH, 5% (v/v) acetic acid) for 10 min.

2.2.3.10 Immunoblotting

After SDS-PAGE, proteins were transferred from an unstained polyacrylamide gel onto Amersham Protran nitrocellulose membrane (Cytiva) or Immobilon-FL polyvinylidene fluoride membranes (Merck). Prior to blotting, membranes and filter paper were activated with methanol and equilibrated with semi-dry blotting buffer (48 mM Tris (pH not adjusted), 39 mM glycine, 0.04% (v/v) SDS, 20% (v/v) methanol). Blotting was performed at 120 mA for 45 min using a semi-dry Blotting chamber V20 SDB (SciPlas), followed by a blocking step in 5% (w/v) bovine serum albumin (BSA) in TBST or in Intercept TBS blocking buffer (Li-COR) for 1 h at room temperature. Primary antibodies were diluted in TBST with 5% (w/v) BSA or Intercept TBS blocking buffer and incubated overnight at 4°C. After three washing steps for 5 min each with TBST, the secondary antibody coupled to horseradish peroxidase (HRD) or IRDye was diluted in TBST with 5% (w/v) BSA or Intercept TBS blocking buffer and applied to the membrane, followed by incubated for 1 h at room temperature. Membranes were washed three times for 5 min each and then visualized using enhanced chemiluminescence (ECL) with Amersham ECL Western Blotting Detection Reagents (Cytiva) or SuperSignal West Pico Plus Chemiluminescent Substrate (Thermo Fisher Scientific) according to the manufacturers´ instructions with a CCD camera ChemiDoc XRS+ system (BioRad) or a ChemiDoc MP Imaging system (BioRad). Analysis of immunoblots was performed using ImageJ (v.1.52q, NIH) and Image Lab (v.6.1.0, BioRad).

2.2.3.11 Kinase assays

In vitro kinase assays were performed to analyze the activity of kinases towards different protein substrates, their regulation by co-factors or their inhibition by small-molecular compounds. For all kinase assays, 0.2-0.5 µM kinase was incubated with 10 µM substrate and ATP in kinase buffer (150 mM Hepes pH 7.6, 34 mM KCl, 7 mM $MgCl_2$, 2.5 mM dithiothreitol, 5 mM β-glycerol phosphate) for 15-60 min at 30°C at 300 rpm. GraphPad Prism (v.7) was used for data analysis and representation.

ATP kinase assay

Kinases were incubated with 200 µM ATP. Assays were stopped by the addition of 6x SDS sample buffer and samples were analyzed by SDS-PAGE or further by immunoblots using phospho-specific antibodies.

[^{32}P]-γ-ATP kinase assay

For radioactive kinase assays, kinases were incubated with 200 µM 0.45 mCi [^{32}P]-γ-ATP/ml (Perkin Elmer). Assays were stopped by the addition of EDTA to a final concentration of 50 mM. Samples were spotted onto Amersham Protran nitrocellulose membrane (Cytiva)

followed by three washing steps for 5 min each with PBS. Counts per minute (cpm) were determined in a Liquid Scintillation Counter LS6500 (Beckman Coulter). For inhibition by small-molecular compounds or targeted protein degraders, compounds were pre-incubated with the kinase for 5 min before kinase activity assays were started by addition of substrate at indicated concentrations.

ATP-γ-S kinase assay

For ATP-γ-S kinase assays, kinases were incubated with 2 mM ATP-γ-S leading to thio-phosphorylation of the substrate, and stopped by the addition of EDTA to a final concentration of 50 mM followed by an alkylation reaction with 5 mM p-Nitrobenzylmesylate (PMBM) for 2 h at room temperature. Samples were analyzed by immunoblotting, using an antibody that detects an alkylated thiophosphosphate on the substrate[145].

2.2.3.12 Co-Immunoprecipitation

For each sample, 45 µl of dynabeads protein A (Thermo Fisher Scientific) were pre-blocked with 1 ml 0.5% (w/v) BSA in PBS for 2 h at room temperature under rotation. The beads were then washed three times with IP resuspension buffer (50 mM Hepes pH 7.6, 150 mM NaCl, 2.5 mM $MgCl_2$, 10% (v/v) glycerol, 0.5% (v/v) NP-40) followed by incubation with primary antibodies or a rabbit IgG control in 600 µl of IP resuspension buffer for 2 h at 4°C under rotation (for primary antibodies concentrations see chapter 2.1.6).
For co-immunoprecipitation (IP), a 70-80% confluent 15 cm dish of HEK293T or A375 cells was washed twice with ice-cold PBS. Cells were scrapped in ice-cold PBS and harvested by centrifugation at 400 g for 5 min at 4°C. Cell pellets were resuspended in 800 µl of IP lysis buffer (50 mM Hepes pH 7.6, 150 mM NaCl, 2.5 mM $MgCl_2$, 10% (v/v) glycerol, 1% (v/v) NP-40, 1x Halt Protease and Phosphatase Inhibitor-Cocktail (Thermo Fisher Scientific) and 25-29 units of Benzonase (Merck)) and incubated at 4°C for 30 min under rotation. After centrifugation at 13,000 g for 15 min at 4°C, 800 µl of IP dilution buffer (50 mM Hepes pH 7.6, 150 mM NaCl, 2.5 mM $MgCl_2$, 10% (v/v) glycerol) was added to the supernatant. Total protein concentrations were determined using a Pierce BCA protein assay kit (Thermo Fisher Scientific). The beads conjugated with antibodies were washed three times with IP resuspension buffer followed by incubation with the supernatant containing 1 mg total protein at 4°C overnight under rotation. The next day, beads were washed three times with 800 µl IP resuspension buffer and three times with IP dilution buffer. Proteins were eluted with 45 µl 1x NuPAGE LDS sample buffer (Thermo Fisher Scientific) supplemented with 100 mM DTT at 70°C for 15 min. Immunoblots were performed with either 4-20% Bis-Tris Mini-Protean TGX Precast Protein Gels (BioRad) or 4-12% NuPAGE Bis-Tris gels (Thermo Fisher Scientific) together with NuPAGE MOPS SDS Running Buffer (Thermo Fisher Scientific) depending on the molecular weight of the protein of interest.

2.2.3.13 ChIP-seq

For each sample, 25 µl dynabeads protein A and dynabeads protein G (Thermo Fisher Scientific) were mixed and washed twice with 1 ml ChIP blocking buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% (v/v) Triton-X 100, 0.15% (v/v) SDS, 0.1% (w/v) BSA, 1x Halt Protease and Phosphatase Inhibitor-Cocktail) followed by incubation in ChIP blocking buffer for approx. 6 h at 4°C under rotation.

For Chromatin immunoprecipitation (ChIP), two 80% confluent 10 cm dishes per sample were washed once with ice-cold PBS. For cross-linking, 9 ml PBS was added to each dish followed by the addition of 1 ml formaldehyde solution (50 mM Hepes-NaOH pH 7.6, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 11% (v/v) formaldehyde) and incubation for 30 min at room temperature with shaking. Residual formaldehyde was quenched by the addition of 125 mM glycine and incubation for 5 min at room temperature with shaking. Cells were washed once with 10 ml ice-cold PBS, scraped in 10 ml ice-cold nuclear extraction buffer (20 mM Tris-HCl pH 8.0, 10 mM NaCl, 2 mM EDTA, 0.5% (v/v) NP-40, 1x Halt Protease and Phosphatase Inhibitor-Cocktail), dishes from same conditions were pooled together and pelleted by centrifugation at 250 xg for 5 min at 4°C. Cells were washed with nuclear extraction buffer three times total and were then resuspended in 8 ml ice-cold sonication buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% (v/v) NP-40, 0.3% (v/v) SDS, 1x Halt Protease and Phosphatase Inhibitor-Cocktail). Sonification was performed using an ME220 Focused-Ultra sonicator (Covaris) at 4°C for 15 min per 1 ml sample volume (Peak power: 75.0, duty factor 15.0, cycles/burst: 1.000, average power: 11.3). To ensure proper sonication, 20 µl lysate was collected before and after sonication, digested with 0.5 µg Proteinase K (Promega) for 1h at 55°C and analyzed on a 1% (w/v) agarose gel.
For immunoprecipitation, samples were cleared by centrifugation for 20 min at 20,000 g at 4°C. Supernatants were transferred to a new tube and 1 volume of ChIP dilution buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% (v/v) Triton-X 100, 1x Halt Protease and Phosphatase Inhibitor-Cocktail) was added. Blocked dynabeads were resuspended in 500 µl of ChIP buffer per sample (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% (v/v) Triton-X 100, 0.15% (v/v) SDS, 1x Halt Protease and Phosphatase Inhibitor-Cocktail). Beads, primary antibodies and the cleared supernatant were mixed and incubated with 0.3% (w/v) BSA overnight at 4°C with shaking (For antibody concentrations see chapter 2.1.6). On the next day, beads were washed with 1 ml ChIP buffer followed by a second washing step with 1 ml wash buffer 1 (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% (v/v) Triton-X 100, 0.1% (v/v) SDS, 1x Halt Protease and Phosphatase Inhibitor-Cocktail), and a third washing step with wash buffer 2 (20 mM Tris-HCl pH 8.0, 250 mM LiCl, 2 mM EDTA, 0.5% (v/v) NP-40, 0.5% (v/v) deoxycholate, 1x Halt Protease and Phosphatase Inhibitor-Cocktail), each for 5 min at 4°C under rotation followed by two washing steps with TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) for 1 min at 4°C under rotation. For decrosslinking, 150 µl of reverse crosslinking buffer (200 mM NaCl, 100 mM $NaHCO_3$, 1% (v/v) SDS, 300 ug Proteinase K) was added to each sample and incubated for 4 h at 55°C at 600 rpm. Supernatant was transferred to a new tube and incubated at 65°C overnight at 600 rpm. DNA was isolated using a ChIP DNA Clean & Concentrator kit (Zymo research) according to the manufacturer´s instructions. Samples were eluted in 12 µl of elution buffer. DNA concentrations were determined a Qubit dsDNA HS assay kit (Thermo Fisher Scientific) according to the instructions of the manufacturer with 1 µl of each sample. ChIP sample library preparation and sequencing were performed at the Genomics Core Facility at the Peter MacCallum Cancer Centre, Melbourne, Australia and visualized using Integrative Genomics Viewer (v.2.14.1). Data analysis was performed by Jennifer R. Devlin, PhD, Peter MacCallum Cancer Center, Melbourne, Australia.

2.2.3.14 Analytical size exclusion chromatography

Analytical size exclusion chromatography was performed to analyze protein-protein interactions with a Superdex 200 Increase 3.2/300 column using an Agilent 1260 infinity 1260 bio-inert HPLC System (Agilent Technologies) in HPLC buffer (20 mM Hepes pH 7.6, 300 mM NaCl).

2.2.4 Biophysical methods

2.2.4.1 Mass spectrometry

For peptide mass fingerprint analysis, protein samples were separated by SDS-PAGE and stained with Coomassie brilliant blue. Protein bands of interest were excised from the gel and analyzed by mass spectrometry at the Proteomics Facility, Max Planck Institute for Biophysical Chemistry in Göttingen, Germany. Mass spectrometry data was analyzes using Scaffold 4 (v.4.8.7).

2.2.4.2 Protein thermal stability analysis

Nano-differential scanning fluorimetry was used for the determination of thermal protein stability using a Prometheus NT.48 (NanoTemper) devise operating at the two wavelengths 330 and 350 nm. Proteins were diluted to a final concentration of 5 µM in kinase buffer and incubated with 1, 10, or 100 µM ADP, ATP, or compound for 10 min prior to measurements. The thermal stability was monitored from 20 to 70 °C at a heating rate of 1.5°C/min using the PR.ThermControl software. Measurements were performed in duplicates.

2.2.4.3 Surface plasmon resonance spectroscopy

Surface plasmon resonance (SPR) experiments were performed using a Biacore 8K instrument (Cytiva). All steps were performed at 25°C in SPR buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1 mM $MgCl_2$, 0.05% (v/v) Tween20). Kinases were immobilized by amine coupling: The flow cell 1 and 2 surfaces of a CM5 sensor chip (Cytiva) were both activated with 50 mM NaOH for 15 sec at a flow rate of 30 µl/min followed by activation with a 1:1 mixture of 0.1 M N-hydroxysuccinimide (NHS) and 0.1 M 3-(N,N-dimethylamino)propyl-N-ethylcarbodimide (EDC) for 7 min at a flow rate of 10 µl/min. Subsequently, the flow system was washed with 1 M ethanolamine pH 8.0. Optimal kinase concentration varied and were therefore optimized for the respective batch before each measurement. Kinase immobilization was carried out in acetate pH 5.0-5.5 or MES pH 6.5 on the surface of flow cell 1 for 160 sec at a flow rate of 10 µl/min. Surfaces were blocked with 1 M ethanolamine pH 8.0 for 7 min at a flow rate of 10 µl/min. Data was collected at a rate of 10 Hz and double referenced by blank cycles and reference flow cell subtraction. Processed data was fitted using a 1:1 interaction model with the Biacore Insight Evaluation Software (Cytiva).

Binding affinity measurements of protein-protein interactions

Binding affinity measurements of protein-protein interactions were performed in SPR buffer. Kinetic binding of ligands was measures as single-cycle kinetics. Ligands were injected at a flow rate of 30 µl/min (association: 120 sec, dissociation: 300 sec) over both flow cells at increasing concentrations.

Binding affinity measurements of protein-compound interactions

Binding affinity measurements of protein-compound interactions were performed in SPR buffer containing 2% (v/v) DMSO. Kinetic binding of compounds was measured as single-cycle kinetics. Compounds were injected at a flow rate of 30 µl/min (association: 120 sec, dissociation: 600-900 sec) over both flow cells at increasing concentrations. To compensate any potential effects of DMSO, the data was solvent corrected.
For SPR measurements with abemaciclib, kinases were additionally stabilized by pre-incubation with a two-fold excess of compound for 5 min at room temperature to increase the amount of active immobilized kinase on the sensor chip surface, prior to a 1:5 dilution in acetate buffer. Abemaciclib was washed out for at least 60 min at a flow rate of 20 µl/min before conducting kinetic measurements.

2.2.4.4 Protein Crystallization

Crystallization of apo-HIPK3

Initial crystallization screenings of HIPK3 were conducted using a homemade PegMix screen (0.1 M Hepes pH 7.0, 30% (w/v) medium weight PEG-mixture). Purified HIPK3 protein was concentrated to 10.5 mg/ml, mixed with 1 mM ADP/Mg^{2+} and crystallized by the hanging drop vapor diffusion method at 15°C. Optimized hexagonal-shaped crystals were grown at a 1:1 ratio of protein and reservoir solution containing 0.1 M Hepes pH 7.5, 0.2 M $MgCl_2$, and 15-17% (w/v) medium weight PEG mix. Crystals were cryoprotected with 15-20% (v/v) ethylene glycol in mother-liquor and flash frozen in liquid nitrogen. Crystallization of apo-HIPK3 was performed by Dr. Kanchan Anand.

Crystallization of HIPK3 with abemaciclib

Crystals of HIPK3 in complex with abemaciclib were obtained by mixing 10.5 mg/ml HIPK3 with a fivefold molar excess of abemaciclib followed by incubation on ice for 30 min prior to crystallization. Each hanging drop was set using a 1:1 ratio of protein-ligand mixture and mother-liquor. Initial crystals were further optimized by micro-seeding. Optimized crystals appeared in about 4 to 5 days at 15°C in drops containing 0.1 M Tris/HCl pH 8.0, 0.2 M $MgCl_2$, and 8% (w/v) PEG 8000 solution. Crystals were cryoprotected with 15-20% (v/v) ethylene glycol in mother-liquor and flash frozen in liquid nitrogen. Crystallization of HIPK3 with abemaciclib was performed by Dr. Kanchan Anand.

Crystallization of DYRK1A with abemaciclib

DYRK1A in complex with abemaciclib was crystallized at a protein concentration of 14.4 mg/ml with 14-16% (w/v) PEG 4000, 0.1 M sodium citrate pH 6.5 and 0.1 M ammonium sulfate at a fivefold molar excess of the inhibitor. Crystals optimized for diffraction data collection appeared in about a week at 15°C using the hanging drop vapor diffusion method. Crystals were cryoprotected with 15-20% (v/v) ethylene glycol in mother-liquor and flash-frozen in liquid nitrogen.

Crystallization of Cdk12/CycK with NVP2

For initial crystallization screenings, the sitting drop vapour diffusion technique in 96-well plate format was performed with 12 mg/ml Cdk12/CycK mixed in a 1:10 molar ratio with NVP2 using a mosquito crystallization robot (TTP Labtech). Crystal optimization was performed in

24-well-plates (Hampton Research) at 15 °C by the hanging drop vapour diffusion technique, mixing equal volumes (0.6 µl/ 0.6 µl) of protein and crystallization buffer (MES pH 6.0, 0.2M $MgCl_2$, 30% mixture of medium weight polyethylene glycols). Rod-shaped small crystals appeared within 7 to 10 days and were cryoprotected with 15% ethylene glycol in mother-liquor, rapidly suspended in a nylon loop (Hampton Research) and flash-frozen in liquid nitrogen. Crystallization of Cdk12/CycK with NVP2 was performed by Dr. Kanchan Anand.

Crystallization of Cdk12/CycT1 with NVP2

For initial crystallization screenings, NVP2 was added to Cdk12/CycT1 in a tenfold molar excess after size exclusion chromatography, concentrated to 7.5 mg/ml and crystallized using a Gryphon pipetting robot (Art Robins) and commercial crystallization screens (Molecular Dimensions and Jena Bioscience) using sitting drop vapour diffusion in 96-well plates. Crystals appeared after 21 days in various conditions. For optimized crystals, protein was mixed with reservoir solution (0.17 M Mes pH 6.0, 0.4 M MgCl2, 20.5% PEG 3350) in a 2:1 ratio and crystals were grown using sitting drop vapour diffusion in 96-wells at 20°C. Crystals were cryoprotected with 35% (v/v) glycerol in mother-liquor and flash-frozen in liquid nitrogen.

2.2.4.5 Diffraction data collection

Diffraction data were collected from a single loop-mounted crystal each, held in a stream of liquid nitrogen gas at 100 K. The diffraction data sets were collected at the PX1 synchrotron beamline at Swiss Light Source, Villigen, Switzerland, equipped with an Eiger detector or at the beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY). For further details on diffraction data collection see Appendix A.

2.2.4.6 Structure determination and model building

Diffraction data were processed and scaled using XDS[146]. The phase problem was solved by molecular replacement using PHASER[147] and the model was refined by alternating cycles using PHENIX[148]. Manual rebuilding and visual comparisons were made using the graphical program COOT[149] and the stereochemical quality of the model was confirmed using a Ramachandran plot. Protein interfaces and accessible surface areas were calculated with the program PDBe-PISA (v.1.52). All structural figures were prepared with PyMOL (Schrödinger, LLC). For further details on diffraction data collection see Appendix A. Structure determination was performed by Dr. Kanchan Anand with the help of Gregor Hagelueken, PhD.

Structure determination of apo-HIPK3 and HIPK3 with abemaciclib

To correct for anisotropic scattering, apo-HIPK3 diffraction data were submitted to the StarAniso server (Global Phasin Ltd.). Ellipsoidal truncation resulting in low completeness did not hamper the available data resolution for model building and refinements[150,151]. The coordinates of HIPK2 (PDB: 6P5S) were used as a search model for the apo structure of HIPK3. The final apo-HIPK3 structure includes one continuous chain of residue 184 to 550 but showed no ADP in the nucleotide binding site, while the structure of HIPK3 in complex with abemaciclib contains residue 184 to 551 and has been refined to R_{work} and R_{free} values of 24.3/27.3% and 24.6/27.5%, respectively.

Structure determination of DYRK1A with abemaciclib

The structure of the DYRK1A-abemaciclib complex was solved by molecular replacement using the coordinates of DYRK1A (PDB: 2VX3). A clearly defined region of electron density from a non-protein entity appeared at the ATP exit site that was modeled as two citrate molecules, exhibiting a central density in between both. Following a previous report of a crystal structure determined at 1.35 Å resolution[152], refinement of a Li^+-bis-citrate complex gave a good fit to this electron density. The final model of DYRK1A in complex with abemaciclib contains residue 134 to 480 for both chains and has been refined to R_{work} and R_{free} values of 18.6/21.1%.

Structure determination of Cdk12/CycK with NVP2

The coordinates of Cdk12/CycK (PDB: 4NST) were used as a search model for molecular replacement. The final Cdk12/CycK-NVP2 structure was refined to 2.75 Å resolution, R_{work}/R_{free} values of 24.2/30.0% and include residue 716-1040 for Cdk12 and 20-265 for CycK with NVP-2 in the nucleotide binding site.

Structure determination of Cdk12/CycT1 with NVP2

The coordinates of Cdk12 (PDB: 4NST) and CycT1 (PDB: 2PK2) were used as a search model for molecular replacement. The final Cdk12/CycT1-NVP2 structure was refined to 2.03 Å resolution, R_{work}/R_{free} values of 22.9/26.0% and include residue 717-1035 for Cdk12 and 9-250 for CycT1 with NVP-2 in the nucleotide binding site.

3. RESULTS

3.1 Abemaciclib is a potent inhibitor of HIPKs and DYRK1A

3.1.1 Purification of recombinant HIP kinases and DYRK1A

To gain active, recombinant protein kinases, human HIPK1–4 were cloned, expressed and purified from *Sf9* insect cells and DYRK1A from *E.coli* bacterial cells. HIPK1-3 and DYRK1A constructs comprised only the kinase domain because full-length constructs were insoluble. Only HIPK4 was purified as a full-length construct with an MBP-tag for protein stabilization. The exact construct length and purification strategies of all recombinant proteins used in this study are depicted in Appendix B. Peptide mass finger print confirmed protein identities, as well as phosphorylation of the conserved tyrosine residue in the activation loop. In vitro kinase assays were performed to confirm kinase activity with [^{32}P]-γ-ATP and c-Myc as a substrate (Fig.11). To control for background signals, substrate without kinase was included in the measurement, as well as a kinase dead-mutant (kdm) of HIPK1-4 to exclude contamination of the kinase with other kinases from *Sf9* cells, and the wildtype kinase without substrate to control for autophosphorylation. All five wildtype kinases showed robust kinase activity .

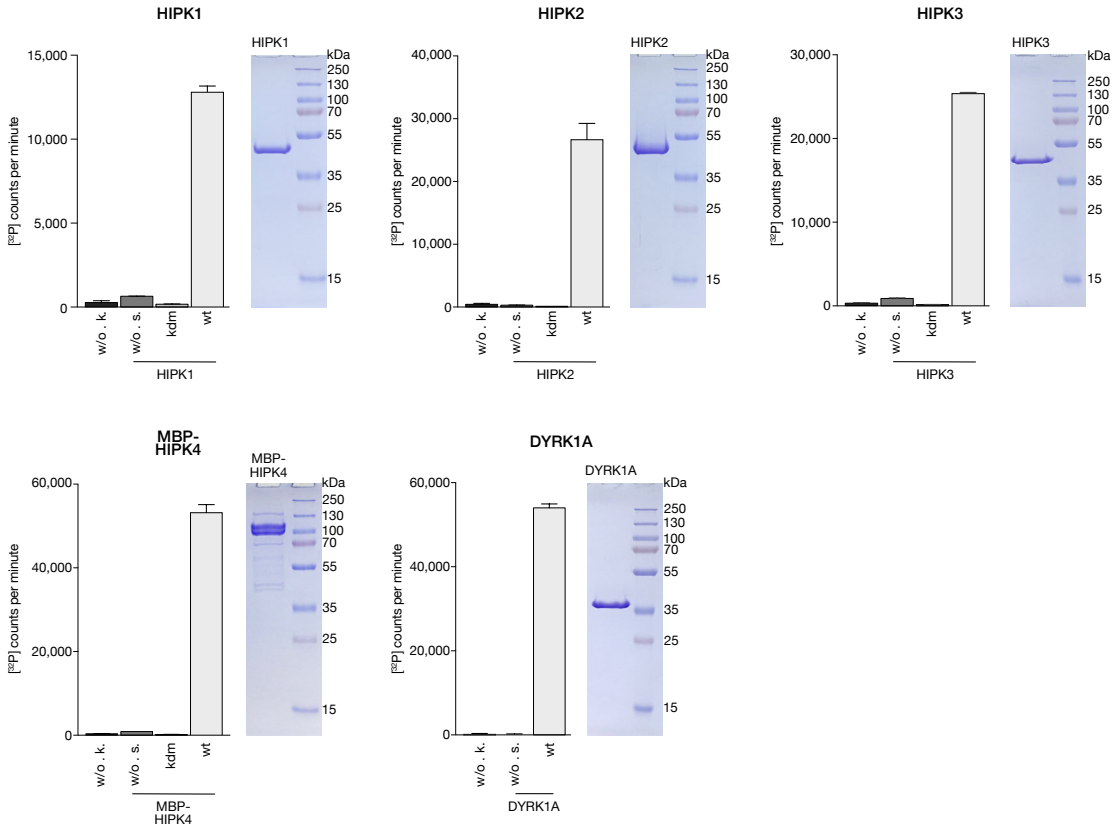

Fig.11 | Purification of active, recombinant HIPK1-4 and DYRK1A kinases. In vitro kinase assays (left) with 0.2 mM [^{32P}]-γ-ATP and 10 µM c-Myc as a substrate demonstrate robust kinase activities for 0.2 µM wildtype kinases (wt) compared to substrate without kinase (w/o. k.), kinase without substrate (w/o. s.), and a kinase-dead mutant (HIPK1$_{D315N}$, HIPK2$_{D324N}$, HIPK3$_{D322N}$, and HIPK4$_{D136N}$). All data are presented as mean ± SD of duplicates. Purity of recombinant protein kinases was analyzed by SDS-PAGE (right).

3.1.2 Structural features of the human DYRK and HIPK family

Crystal structure of the human HIPK3 kinase domain

The human HIPK3 kinase domain (residues 159-562) was purified to homogeneity by affinity and size exclusion chromatography. Crystals were grown by the hanging drop vapor diffusion method in the presence of ADP and $MgCl_2$. The structure was determined to 2.5 Å resolution and an R_{work} of 0.24 and R_{free} of 0.27 with excellent stereochemistry (Fig.12; for crystallographic data collection and refinement statistics see Appendix A, Tab.5). The asymmetric unit cell of the protein crystal was formed by one HIPK3 monomer. Although ADP and $MgCl_2$ were present in the crystallization condition, no crystallographic density was visible, implying that the kinase is in the nucleotide free apo-state. The HIPK3 structure shows a continuous polypeptide chain from residue 184 to 550 and exhibits a classical kinase fold consisting of an N-terminal (residue 197-278) and a C-terminal lobe (residue 279-526) with a fully ordered activation segment. The structure exhibits a sequence identity of 88% compared to the human HIPK2 crystal structure[153] and an RMSD value of 0.58 Å over 290 Cα atoms. Significant deviations occur in the CMGC-specific insert region, the activation loop, and the loop around position 480 connecting helices αN with αH. At the N-terminus, 13 additional residues are visible in the electron density map, preceding the canonical kinase domain and forming an extended stretch in an antiparallel conformation to the β1 strand of the N-lobe. Although these residues adopt a similar conformation as the DH box in DYRK1A and DYRK2, the two central tyrosines which stabilize the association of the DH box with the kinase domain in DYRKs[72], are not conserved in HIPKs (Fig.13).

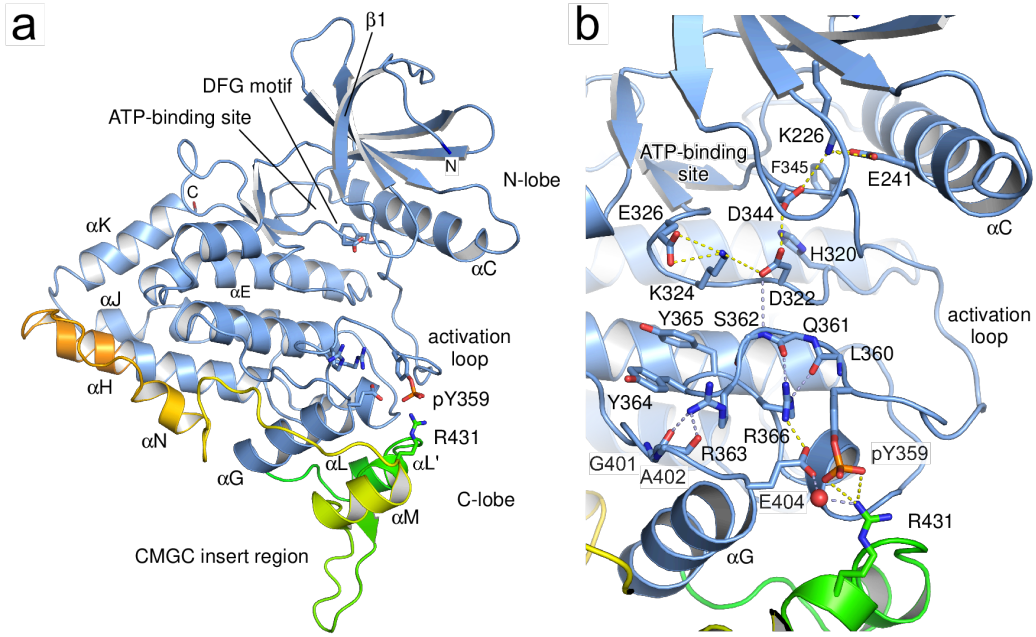

Fig.12 | Crystal structure of the human HIPK3 kinase domain. a Key elements as the αC helix, the DFG motif, and the phosphorylated tyrosine within the activation loop are shown. The CMGC insert region embedded between canonical helices αG and αH is colored from green to orange. **b** Coordination of the kinase active center. Electrostatic and hydrogen bond interactions between the phosphorylated tyrosine of the activation loop, pY359, the RYYR element, the HAD motif with the general base, the DFG motif, the catalytic lysine K226, and the coordinating glutamate E241 of the αC helix are indicated (PDB: 7O7I).

Another 24 residues are visible at the C-terminus extending the chain to position 550 by two extended helices (αJ and αK) that form an α-barrel with helices αI and αE of the C-lobe at the back of the ATP binding site. Although being in the apo state, the HIPK3 structure adopts the active kinase conformation with a phosphorylated tyrosine in the activation loop, the DFG motif, the αC-helix in the "in" position and a fully accessible ATP-binding site.

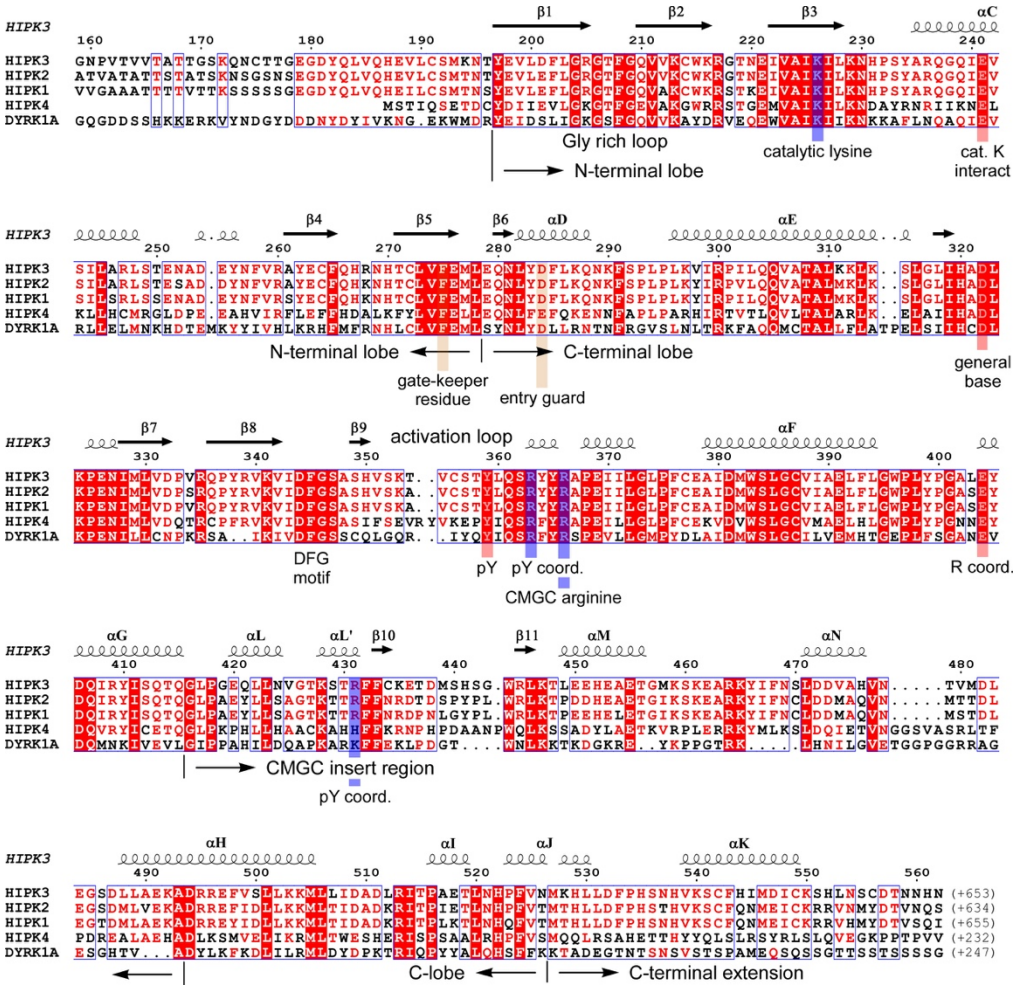

Fig.13 | Sequence alignment of the human HIPKs and DYRK1A kinase domains. The sequence alignment of the kinase domain of human HIPKs and DYRK1A over the entire length of the crystallized HIPK3 construct (159-562) was generated with MultAlin. Secondary structure elements were analyzed with ESPript and are indicated as determined for the apo-HIPK3 structure. Characteristic sequence motifs including the phosphorylated T-loop tyrosine and functional regions are indicated.

DYRK- and HIPK-specific structural features

The activation loop of DYRKs harbors a YxY element whose second tyrosine is auto-phosphorylated for kinase activation[72]. In HIPK1-3 this motif is altered to a STY and to an EPY sequence in HIPK4 (Fig.13). In the HIPK3 structure, the conserved tyrosine Y359 of the STY motif is the only phosphorylated residue. The activating tyrosine is neighboring the "CMGC arginine" R366, which is part of the R_{363}YYR motif and a characteristic feature of the CMGC kinase family[154]. In all known protein structures of DYRKs and of HIPK2, the phosphate group of the conserved

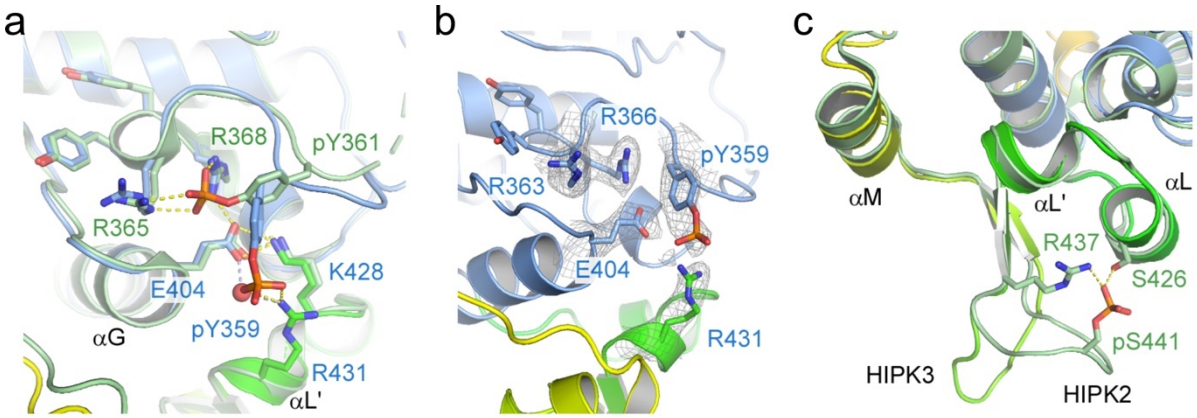

Fig.14 | Coordination of the phosphorylated tyrosine and the CMGC insert region in HIPK2 and HIPK3. a Overlay of the conserved phospho-tyrosine in HIPK3 (blue) and HIPK2 (light green; PDB: 6P5S). Electrostatic and hydrogen bond interactions between neighboring residues are indicated. **b** Close up of the pY359 coordination in the HIPK3 activation segment. The final 2 Fo-Fc density is displayed at 1σ. **c** Overlay of the CMGC insert in HIPK3 and HIPK2. In HIPK2, pS441 makes contacts with residues from the αL-helix whereas no phospho-serine is present in HIPK3.

tyrosine forms strong electrostatic interactions with the guanidine side chains of these two arginines, stabilizing the activation loop in the kinase active conformation. In our HIPK3 structure, however, the phosphate group of pY359 forms a 3.0 Å salt bridge with the side chain of R431, which is part of the HIPK specific insert region (Fig.12b; Fig. 14). The RYYR motif remains at the same position as in DYRKs and HIPK2, but R363 is hydrogen-bonded to the carbonyl group of G401 and R366 forms a tight salt bridge (2.6 Å) to E404. This interaction develops into an electrostatic network with a water-mediated hydrogen bond to R431. In addition, the CMGC arginine R366 forms hydrogen bonds with the main-chain oxygen of L360 and Q361, stabilizing the turn of the activation loop. An intermolecular salt bridge from a symmetry related molecule accompanies the unconventional conformation of the pY359 phosphate group toward R431 (Fig.15). At the tip of the loop between helices αN and αH of the neighboring molecule, D482 forms tight salt bridges to R363 and R366 of the RYYR motif. Additionally, a weak intermolecular salt bridge is formed between the pY359 phosphate group

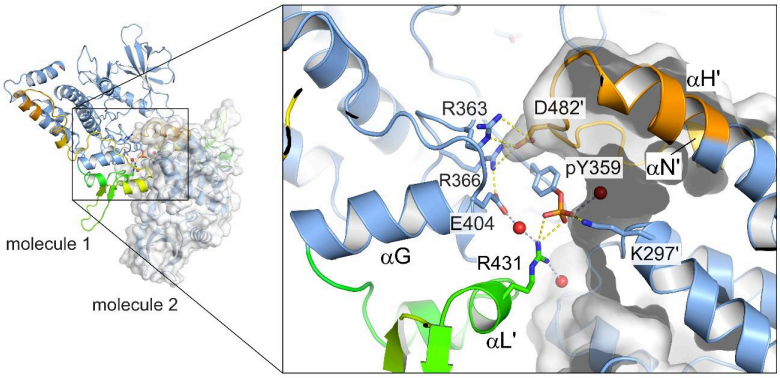

Fig.15 | Crystallographic assembly of HIPK3. In the apo-HIPK3 structure, pY359 of the activation loop forms a salt bridge with R431 and a weak electrostatic interaction with K297' of a symmetry related molecule (surface representation). The carboxy side chain of D482' from the symmetry related molecule forms tight salt bridges with both arginines of the RYYR motif.

and K297 of the symmetry related molecule. The electrostatic surface potential surrounding pY359 in HIPK3 is shown in Fig.14b.

The CMGC insert region is a characteristic feature for the HIPK and DYRK kinase branch and serves as a specific binding platform for signaling partners[154]. HIPKs have the longest insert region of all CMGC kinases with 78 to 84 residues inserted into the C-lobe between the canonical helixes αG and αH (colored green to orange in Fig.12). In HIPK3, the insert region consists of four helices with two short antiparallel β-strands and extends across the base of the C-lobe resulting in a prominent extension of the αH helix. Following two short helices (αL and αL′), the β-hairpin (residues 433–446) extends over a loop of eight residues before turning toward the αG helix and meandering over 40 residues with two long loops and two helices (αM and αN) into the canonical C-lobe. Interestingly, this conformation differs substantially from the known structures of HIPK2, DYRK1A, DYRK2, and DYRK3, highlighting the diversity and specificity of the CMGC insert region in this kinase branch (Fig.16). For example, in HIPK2 pS441 bends the β-hairpin loop toward helix αL, while this loop is three residues shorter in DYRK1A. In DYRK2 and DYRK3, on the other hand, the β-hairpin loop is oriented toward the position of the αM helix in HIPK3. In addition to these divergent conformations, DYRK2 contains two and DYRK3 one phosphorylated residue in the insert region, which further contributes to the specificity of this characteristic region.

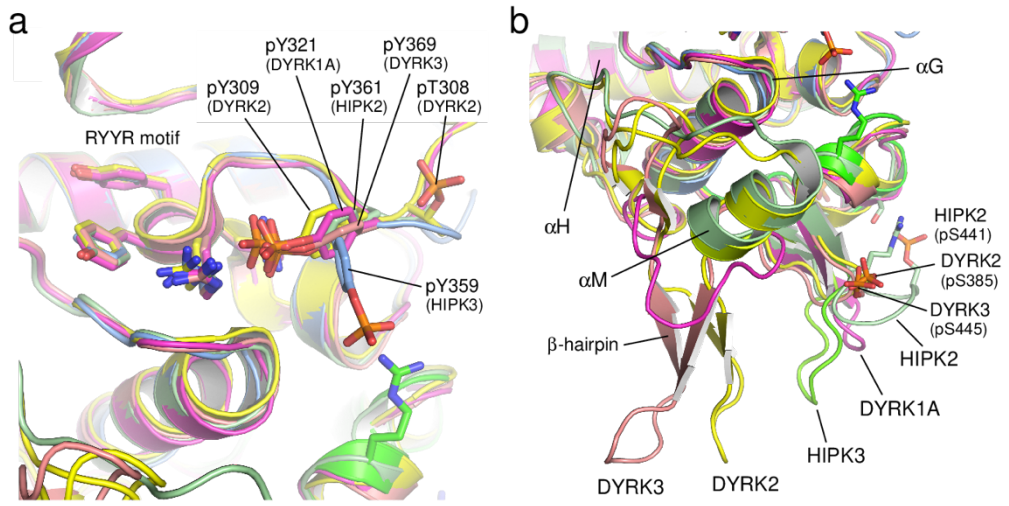

Fig.16 | Coordination of the phosphorylated tyrosine and the CMGC insert region in HIPKs and DYRKs. a In HIPK2 (light green; PDB: 6P5S), DYRK1A (magenta; PDB: 2VX3), DYRK2 (yellow; PDB: 2VX3) and DYRK3 (salmon; PDB: 5Y86), the phospho-tyrosine of the activation loop forms salt bridges with both arginines of the R(Y/F)YR motif. In HIPK3 (blue; PDB: 7O7I) instead, pY359 interacts with R431 of the CMGC insert region. **b** Diverse orientations of the β-hairpin loop in DYRKs and HIPKs as well as additional phospho-residues, highlighting the diversity in the CMGC insert region within this kinase family.

3.1.3 HIPKs phosphorylate proteins of the transcription machinery

In metazoans, gene transcription and co-transcriptional RNA processing are mediated by RNA pol II and a variety of general transcription factors, which are phosphorylated at discrete

checkpoints of the transcription cycle at the initiation, pausing, elongation, and termination phase. Similar to the transcriptional kinase P-TEFb, DYRK1A has been shown to phosphorylate RNA pol II CTD at Ser2 and Ser5 residues. However, whereas P-TEFb is a key regulator of pause-release and therefore involved in the regulation of most genes, expression levels of DYRK1A are more tissue-specific. Moreover, it has been specifically found at genes critical for tissue homeostasis and development[89,155]. Considering the provocative similarities between DYRK1A and the HIPK family, HIPKs may be involved in the direct regulation of the transcription machinery, as it has been shown for DYRK1A. To investigate this in more detail, all four HIPKs, DYRK1A and P-TEFb were tested for their activity toward different recombinant substrates in in vitro kinase assays using [^{32}P]-γ-ATP (Fig.17). Interestingly, we found that DYRK1A and HIPK1-3 show even more pronounced kinase activity towards the C-terminal region of SPT5, a component of the pausing factor DSIF, than P-TEFb. Likewise, all four HIPKs showed robust kinase activity toward the transcription factor and proto-oncogene c-Myc, and toward pol II CTD.

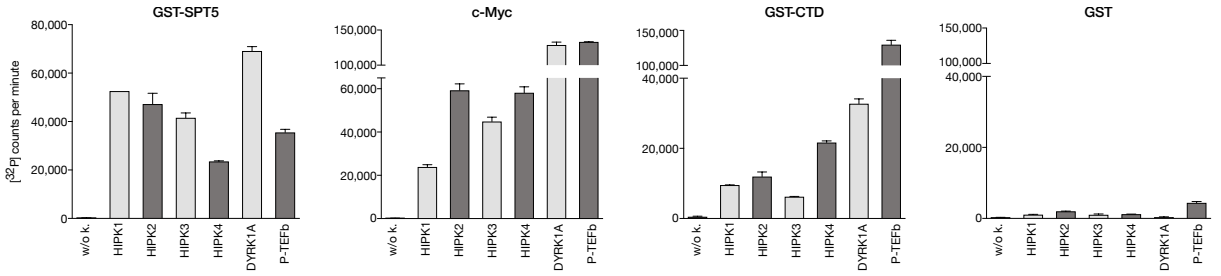

Fig.17 | HIPKs phosphorylate general transcription factors and RNA pol II CTD. HIPK1-4 and DYRK1A show kinase activity toward similar substrates of the transcription machinery as P-TEFb. For radioactive kinase assays, 0.2 mM [^{32}P]-γ-ATP was incubated with 0.2 µM kinase and 10 µM of human GST-SPT5, c-Myc, pol II GST-CTD, and GST as a control for 30 min. All data are presented as mean ± SD of duplicates.

To further define the specificity of HIPKs for the exact phospho-residues on the pol II CTD heptad repeats, we performed immunoblotting using monoclonal antibodies against pTyr1, pSer2, pThr4, pSer5, and pSer7 residues (Fig.18). Similar to DYRK1A, HIPK1-3 exhibit a phosphorylation preference for residues Ser2 and Ser5. In contrast, phosphorylation of Tyr1 and Thr4 was not detectable. Strikingly, HIPK4 seems to be the only HIPK that phosphorylates pol II CTD at Ser7 residues, highlighting the distinct character of this kinase compared to its family members. Consistently, in time-course experiments with gel shift assays (Fig.18b) the migration band of CTD from a hypo- (IIa) to a hyper-phosphorylated (IIo) state was clearly detectable for all four HIPKs and DYRK1A.
In recent studies it has been shown that pre-phosphorylation of a specific residue at Tyr1 or Ser7 primes RNA pol II CTD recognition by Cdk9, Cdk12 and Cdk13[42,43,156,157]. We analyzed the activity of HIPKs and DYRK1A toward a set of synthetic CTD peptides, each peptide comprising three heptad repeats followed by a polyethylene linker and two arginines (Fig.19). In addition to the CTD consensus sequence, CTD peptides were uniformly pre-phosphorylated at position Tyr1, Ser2, Thr4, Ser5, Ser7, or with Lys7, the most frequent alteration in the human CTD consensus repeats[9]. Although background signals were high for HIPK2 and HIPK3, the highest activity of all five kinases was seen toward consensus CTD

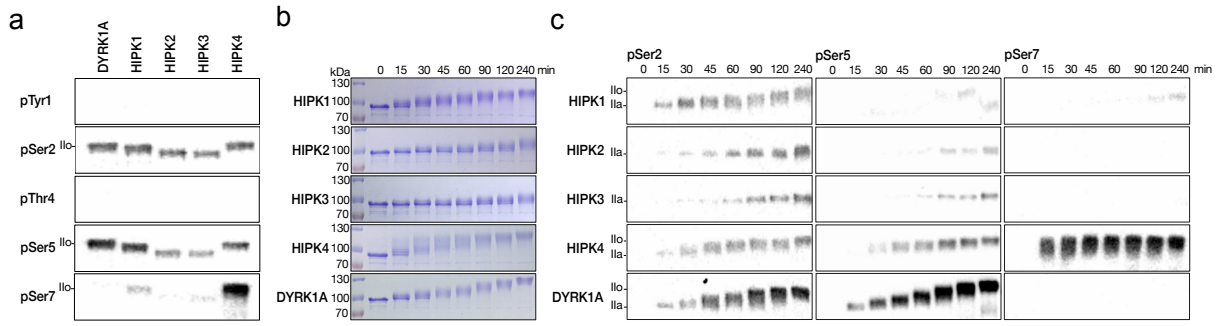

Fig.18 | Specificity of HIPKs and DYRK1A toward pol II CTD heptad residues. a In vitro kinase assays were performed with 0.2 mM ATP, 10 µM GST-CTD$_{[52]}$, and 0.2 µM kinase for 120 min and visualized by immunoblotting using RNA pol II CTD specific phosphoantibodies against pTyr1, pSer2, pThr4, pSer5, and pSer7. **b** Gel shift assays were performed as in (a) but in time-course series and visualized in SDS gels by Coomassie staining. Upon phosphorylation, the GST-CTD$_{[52]}$ band shifts from the IIa to the IIo form. **c** Time-course series were performed as in (b), but were visualized by immunoblotting with pol II CTD phospho-antibodies.

and no activity above background for Ser2, Ser5, and Ser7 pre-phosphorylated peptides. For HIPK4 and DYRK1A, kinase activity toward Lys7 substituted peptides was nearly as high as for consensus CTD, moderate for HIPK1-3, and a lower activity was observed for Tyr1 and Thr4 modified CTD peptides. However, in contrast to Cdk9, Cdk12, and Cdk13, a preference for Ser7 pre-phosphorylated CTD over consensus CTD was not observed for HIPKs and DYRK1A. Taken together, these data support the intriguing possibility that DYRK1A, as well as the HIPK family, are transcription elongation kinases that directly regulate the transcription machinery, as indicated by similar substrate phosphorylation patterns of SPT5 and RNA pol II CTD as it has long been known for P-TEFb.

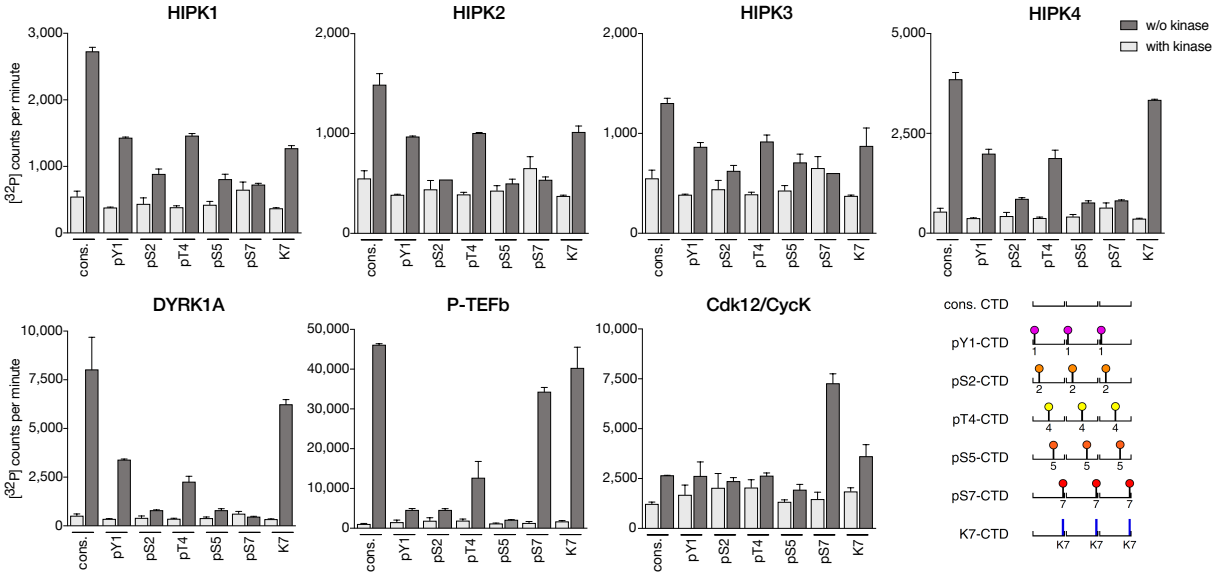

Fig.19 | Specificity of HIPKs and DYRK1A toward pre-phosphorylated pol II CTD residues. In vitro kinase assays were performed with 1 mM [^{32}P]-γ-ATP, 0.5 µM kinase and 150 µM pre-phosphorylated CTD Peptides for 60 min. Each peptide contained three consensus hepta-repeats with either no modification (cons.) or with phosphorylation marks continuously set at one residue of the heptad sequence as depicted in the cartoon. Peptides carried either phosphorylation marks for Tyr1, Ser2, Thr4, Ser5, Ser7 or Lys7. All data are presented as mean ± SD of duplicates.

3.1.4 Abemaciclib is a potent inhibitor of HIPK2, HIPK3 and DYRK1A

HIPK3 inhibitor screen:

HIPK3 has been associated with the pathology of several diseases such as type II diabetes[110], Huntington´s disease[158,159], and prostate cancer[111]. However, to date only few kinase inhibitors have been described to specifically inhibit HIPKs[153,160-161]. With the help of Jinhua Wang and Nathanael Gray from the Dana-Farber Cancer Institute, Harvard Medical School, Boston, and Stanford Cancer Centre, Stanford, USA, we performed a database search to identify potent small-molecular inhibitors of HIPK3 based on a kinome-wide binding assay[162] of over a thousand compounds. Six inhibitors were selected from the Gray laboratory compound library (XMD8-70, XMD8-62-i, JWD-065, HTH-01-091, CVM-05-145-3, and CVM-06-033-2), as well as the well-described kinase inhibitors dinaciclib (SCH727965), palbociclib (PD0332991), abemaciclib (LY2835219), GSK1059615, and AS601245. In addition, the widely used pan-selective CDK inhibitors flavopiridol, (R)-roscovitine, and staurosporine were chosen as well as the Cdk9 inhibitor NVP2. In a first screen, these 15 candidate compounds were tested for their ability to inhibit HIPK3 activity in vitro. A concentration of 0.2 µM HIPK3 was preincubated with 1, 10, and 100 µM of the respective compound and 0.2 mM [^{32}P]-γ-ATP for 5 min followed by the addition of 50 µM c-Myc and incubation for 15 min (Fig.20a). While flavopiridol, roscovitine, and dinaciclib showed almost no inhibitory effect on HIPK3, abemaciclib, GSK1059615, JWD-065, HTH-01-091, and CVM-06-033-2 reduced HIPK3 activity most potently. Staurosporine and palbociclib showed only modest efficacy toward HIPK3.

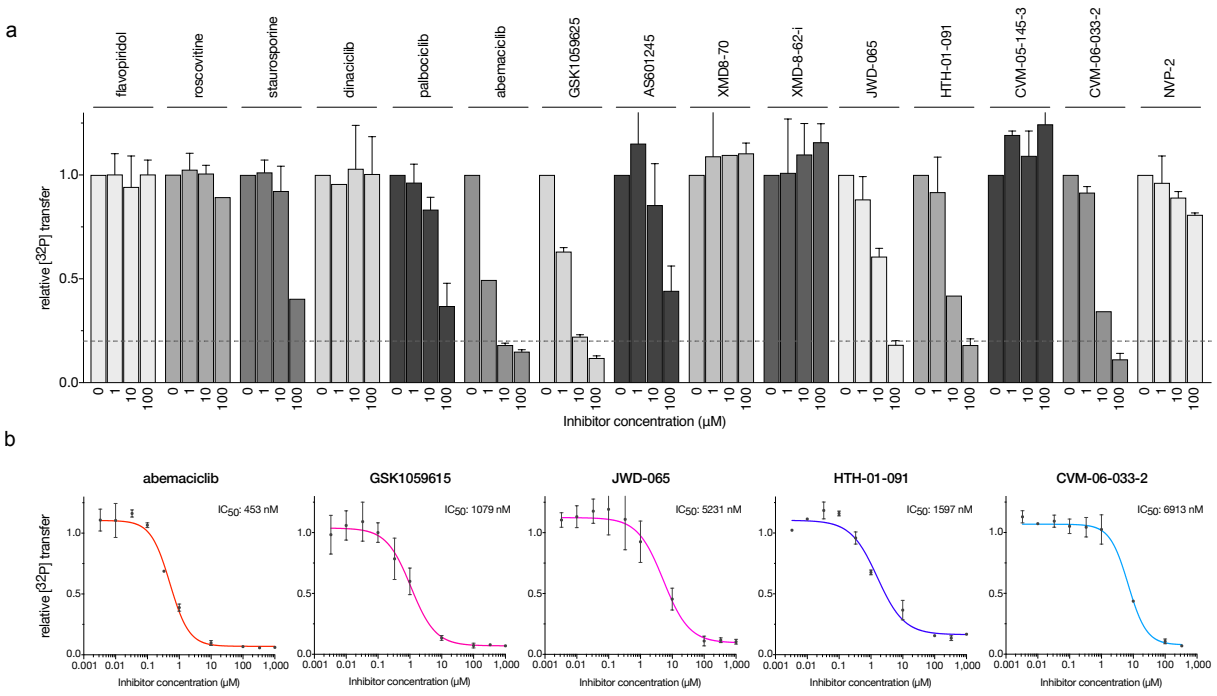

Fig.20 | HIPK3 inhibitor screen. a A panel of 15 small-molecule inhibitors was tested at 1, 10, and 100 µM concentration for HIPK3 inhibition using radioactive kinase activity assays. Compounds were preincubated for 5 min with 0.2 µM kinase and 0.2 mM [^{32P}]-γ-ATP followed by addition of 10 µM c-Myc and another incubation for 15 min. **b** In vitro kinase assays were performed as in (a) with concentration series of the best five small-molecule inhibitors for HIPK3. All data are presented as mean ± SD of duplicates. A sigmoidal fit was used to calculate IC_{50} values.

To analyze the top five hits for HIPK3 inhibition in more detail, concentration series of the compounds ranging from 3.1 nM to 1 mM were incubated with 0.2 µM kinase and 0.2 mM [^{32}P]-γ-ATP concentrations and a sigmoidal fit was used to determine IC_{50} values from these dose-response measurements (Fig.20b). Strikingly, the highest efficacy toward HIPK3 was achieved by abemaciclib with an IC_{50} of 467 nM. Abemaciclib is a Cdk4/Cdk6 inhibitor, which was approved by the FDA in 2017 and is used for the treatment of hormone receptor (HR)-positive and human epidermal growth factor receptor 2 (HER2)-negative metastatic breast cancer[163-164]. Additionally, GSK1059615 (IC_{50}: 1,1 µM) and HTH-01-091 (IC_{50}: 1,6 µM) showed robust inhibition of HIPK3.

To investigate whether the efficacy of abemaciclib toward HIPK3 acts in a significant range in comparison to its original targets Cdk4 and Cdk6, IC_{50} values of abemaciclib toward all four HIPKs, DYRK1A, Cdk4/CycD3, and Cdk6/CycD3 were determined. Retinoblastoma protein 1 (Rb1) was used as a substrate for Cdk4 and Cdk6. Furthermore, the inhibition of Cdk9/CycT1 was tested, a suggested off-target of abemaciclib[165,166] (Fig.21). Interestingly, our in vitro data show inhibition of DYRK1A (IC_{50}: 93 nM), Cdk4/CycD3 (IC_{50}: 78 nM), and Cdk6/CycD3 (IC_{50}: 116 nM) to the same extent. HIPK2 (IC_{50}: 668 nM) was inhibited in a similar range as HIPK3, in contrast to HIPK1 (IC_{50}: 4.53 µM) and HIPK4 (IC_{50}: 10.36 µM). In addition to its described targets Cdk4/6, abemaciclib appears to significantly inhibit members from the HIPK and DYRK family.

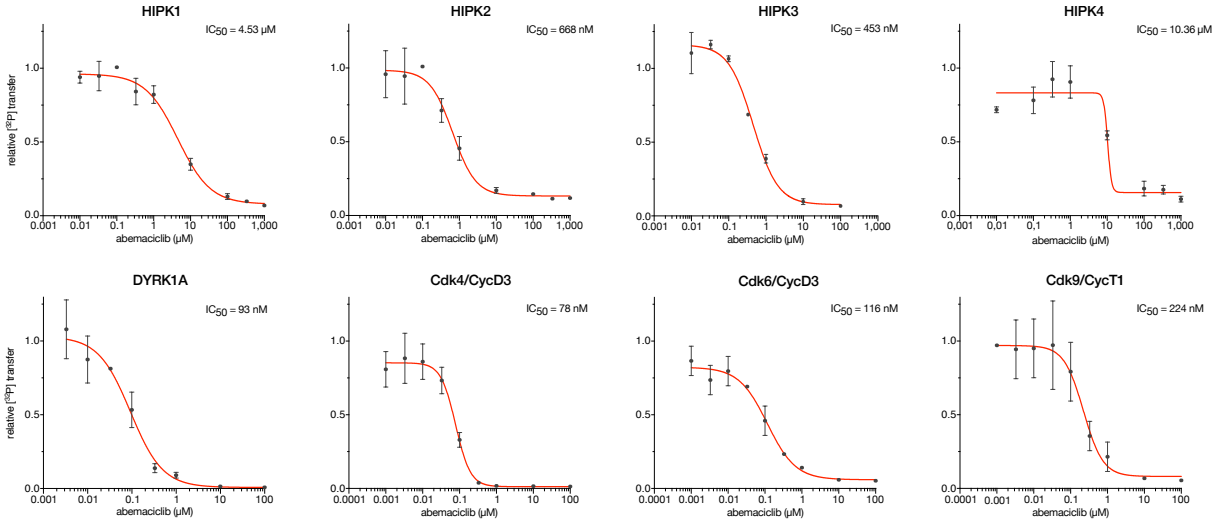

Fig.21 | Abemaciclib is a potent inhibitor of HIPK2, HIPK3 and DYRK1A. Radioactive in vitro kinase assays were performed as in Fig.20 using concentration series of abemaciclib for HIPK1-4, DYRK1A, Cdk4/CycD3, Cdk6/CycD3, and Cdk9/CycT1 inhibition. All data are presented as mean ± SD of duplicates. A sigmoidal fit was used to calculate IC_{50} values.

Binding affinities of abemaciclib toward HIPK2, HIPK3, DYRK1A and Cdk4

To further characterize the interaction of abemaciclib with HIPK1-4, DYRK1A, Cdk4/CycD3, and Cdk9/CycT1, thermal stability shift assays were performed to assess a change in stability of these kinases upon binding abemaciclib at increasing concentrations (Fig.22). Increasing concentrations of abemaciclib (1, 10 or 100 µM) were preincubated with 5 µM kinase in comparison to a DMSO control. HIPK1-3, DYRK1A, and Cdk4/CycD3 were significantly stabilized upon addition of 100 µM abemaciclib with melting temperature shifts of up to

+8.7 °C for HIPK1 and +7.9 °C for Cdk4/CycD3 followed by +7.2 °C, +7.4 °C, and +5.2 °C for HIPK2, HIPK3, and DYRK1A, respectively. In contrast, Cdk9/CycT1 and HIPK4 were only moderately stabilized with increasing melting temperatures of +3.3 °C and +1.9 °C, respectively.

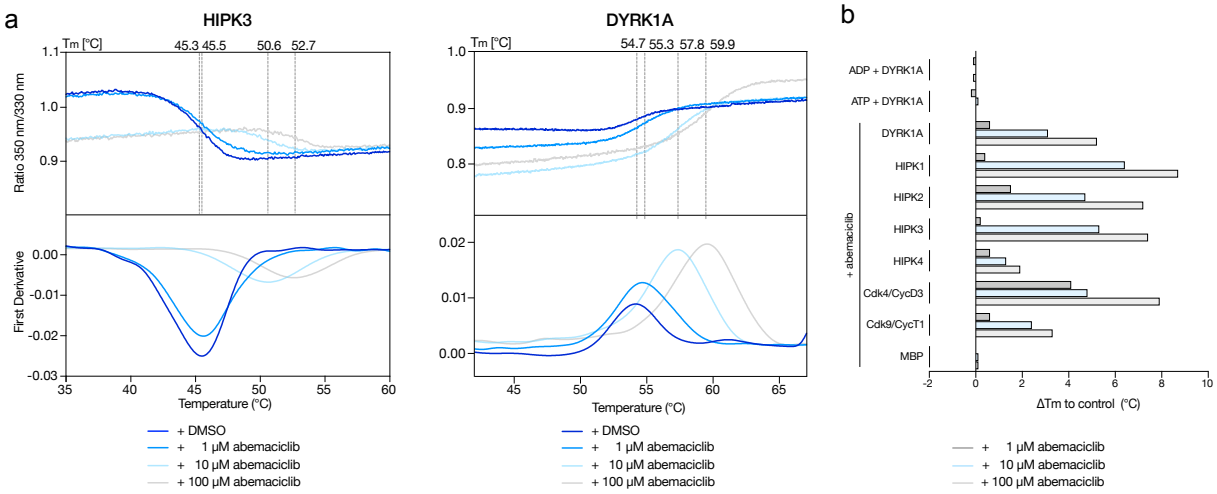

Fig.22 | Thermal stability measurements of HIPK2, HIPK3, DYRK1A and Cdk4/CycD3 with abemaciclib. a Thermal stability measurements were performed using the NanoDSF technique. Proteins were diluted to 5 µM concentration in kinase buffer and incubated with 1, 10, or 100 µM abemaciclib for 10 min. Measurements were performed in duplicates (n = 2 biologically independent samples) and are depicted as mean. **b** Thermal protein stability measurements were performed as in (a) using ADP, ATP, or abemaciclib. ΔTm (max.–min. in °C) is depicted as mean.

Additionally, surface plasmon resonance (SPR) was used to determine dissociation constants (K_D) of abemaciclib with HIPK2, HIPK3, DYRK1A, and Cdk4/CycD3 (Fig.23). The four kinases were immobilized onto a CM5 chip by amine-coupling and flushed with increasing concentrations of abemaciclib (0.55, 1.2, 2.6, 5.8, 12.8, 28.2, 62.4, 136.4, and 300 nM) for single-cycle kinetic measurements. In accordance with the previous results, DYRK1A and Cdk4/CycD3 show dissociation constants in the low nanomolar range with 8.6 and 1.4 nM, respectively, while HIPK2 and HIPK3 show a K_D even in the sub-nanomolar range with 0.8 nM.

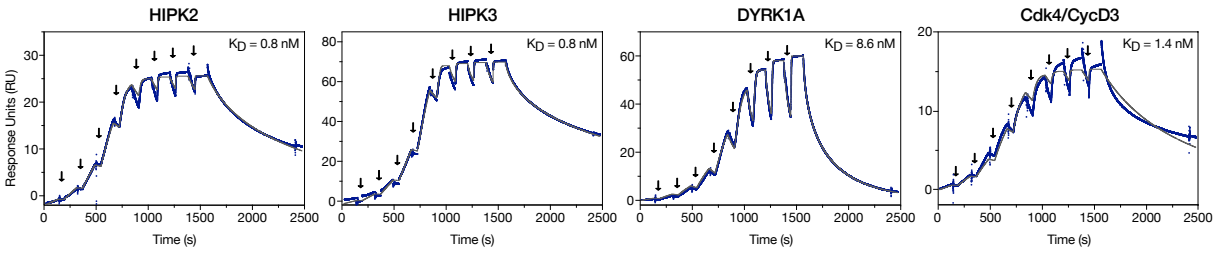

Fig.23 | Binding affinity measurements of abemaciclib with HIPK2, HIPK3, DYRK1A and Cdk4/CycD3. Surface plasmon resonance (SPR) was performed using single-cycle kinetics. Kinases were immobilized to the chip surface by amine coupling. Abemaciclib was injected at increasing concentrations of 0.55, 1.2, 2.6, 5.8, 12.8, 28.2, 62, 136.4, and 300 nM for 120 s followed by dissociation for 900 s. Dissociation constants (K_D) were determined based on a 1:1 interaction model.

Crystal structures of HIPK3 and DYRK1A bound to abemaciclib

For a better understanding of the molecular interactions of abemaciclib with HIPK3 and DYRK1A, both complexes were crystalized and the structures were determined to 2.8 and 1.8 Å resolution, respectively (Fig.24; for crystallographic data collection and refinement statistics see Appendix A, Tab.5). Although abemaciclib is used for the therapeutic treatment of breast cancer, only one crystal structure of abemaciclib bound to Cdk6 is available in the protein data bank at a resolution of 2.3 Å (PDB: 5L2S)[167]. Similar to the Cdk6 structure, abemaciclib interacts intensively with the hinge region of HIPK3 and DYRK1A via its central 2-aminopyrimidine group (Fig.24b). The backbone amino and carbonyl groups of L241 in DYRK1A form hydrogen bonds with the three nitrogen atoms of the aminopyrimidine pyridine unit. The αC-helix in DYRK1A is located in the "in" position, where the coordinating glutamate E203 is located, which together with the catalytic lysine K188 forms H-bonds with the fluorine atom of the benzimidazole head. Toward the backside of the compound, the pyrimidine and

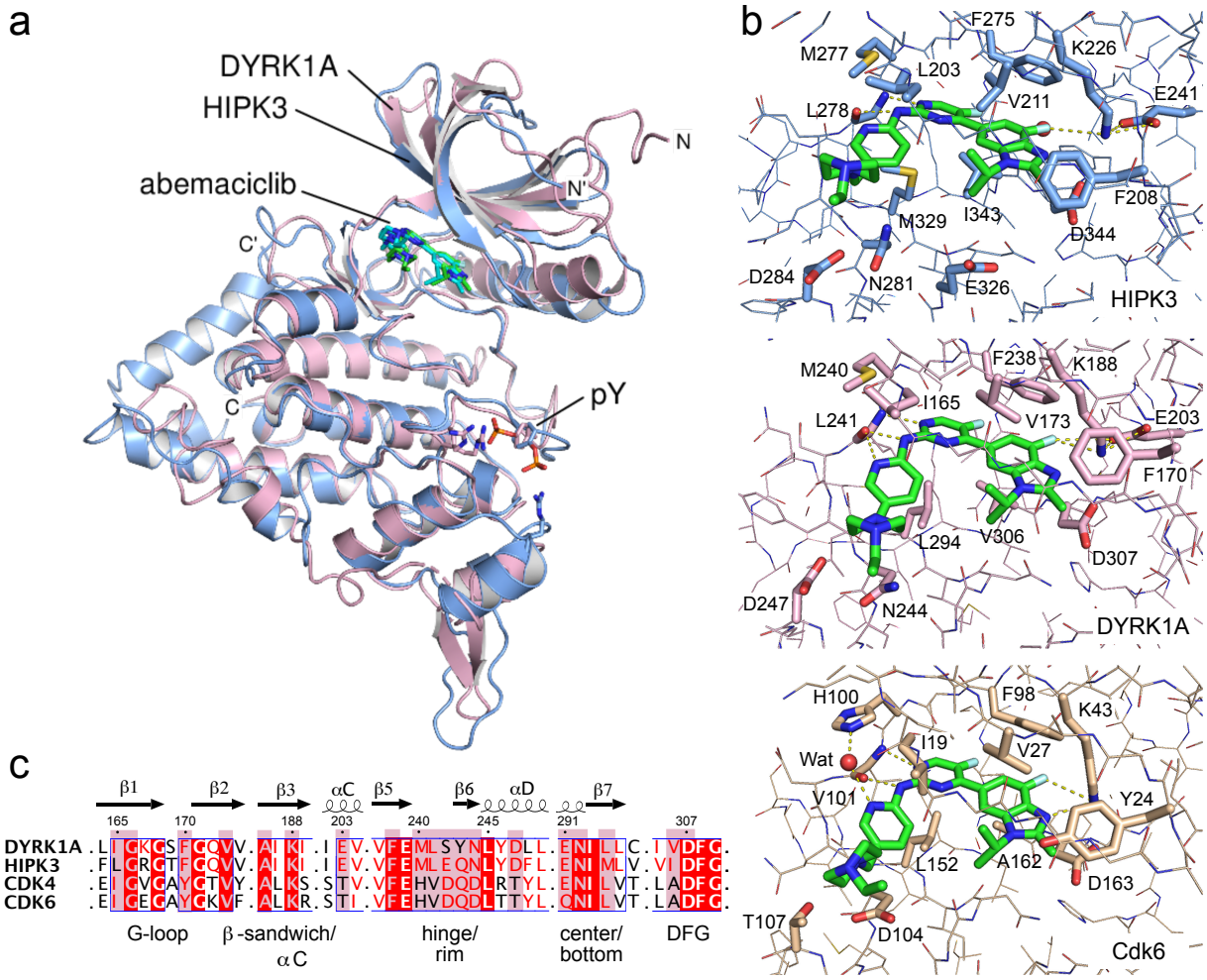

Fig.24 | Crystal structures of DYRK1A and HIPK3 with abemaciclib. a Overlay of the kinase domains of DYRK1A–abemaciclib (rose/green; PDB: 7O7K) and HIPK3–abemaciclib (light blue/cyan; PDB: 7O7J). **b** Molecular interactions of abemaciclib with HIPK3, DYRK1A, and Cdk6 (sand; PDB: 5L2S). Key residues are labeled and hydrogen bonds are indicated as dotted lines. **c** Sequence alignment of DYRK1A, HIPK3, Cdk4, and Cdk6 catalytic center. Residues mediating direct interactions with abemaciclib in DYRK1A are boxed dark rose.

piperazine ring form hydrophobic contacts with I165 at the G-loop, which is fully ordered. The positively charged piperazine ring is stabilized by lying against a solvent-exposed rim consisting of N244 and D247 at the transition between the N- and C-lobe. This latter interaction is thought to be one defining element for the reported specificity of abemaciclib for Cdk4/6 compared to Cdk1/2/3/4[167]. In the Cdk6 structure, these two residues correspond to D104 and T107, both of which are identical in Cdk4 (Fig.24c), whereas in Cdk1/2/3/5 the T107 position of Cdk6 is a lysine that might cause electrostatic repulsion with the piperazine, thereby reducing the efficacy of abemaciclib for Cdk1/2/3/4. In HIPK1-3 and DYRK1A, however, this residue is a negatively charged aspartate, and a larger glutamate in HIPK4, which may be allowing electrostatic association. These two critical residues could explain part of the specificity of abemaciclib within the CMGC kinase group.
The interface between HIPK3 and abemaciclib is almost identical to DYRK1A, in particular the H-bond formation in the hinge region to the 2-aminopyrimidine and pyridine rings are similar. On the other hand, one difference to Cdk6 is the presence of a hydrophobic methionine in HIPK3 (M277) and DYRK1A (M240), whereas Cdk6 contains a histidine (H100) whose imidazole ring forms a water-mediated hydrogen bond to the pyridine nitrogen. While no water molecule is present in a similar coordination in HIPK3 and DYRK1A, a weak hydrogen bond could be formed between the methionine sulfur and the nitrogen in the linker between the aminopyrimidine and the pyridine, which compensates for the water-mediated interaction in Cdk6. Toward the base, the piperazine ring in HIPK3 is in the same saddle conformation as in the other two structures, but is bent upward, which might be caused by L203 at the G-loop, which is a β-branched isoleucine in DYRK1A and Cdk6. Finally, the buried surface of abemaciclib in DYRK1A (536 Å) and HIPK3 (557 Å) is similar to that of Cdk6 (547 Å), although the compound is located slightly deeper in the DYRK1A ATP-binding pocket. Both structures of abemaciclib with these novel off-target kinases, which have similar IC_{50} values and binding affinities as the original targets, Cdk4 and Cdk6, contribute to the understanding of the mode of action of this clinically used drug.

3.2 Novel CDK/Cyclin pairs regulating transcription elongation

3.2.1 Complex formation and activity of novel tCdk/Cyclin pairs in vitro

Cell cycle (c) Cyclins are expressed and degraded depending on the phase of the cell cycle and are thus able to activate multiple different cCDKs. According to the current model, transcriptional (t) Cyclins show stable expression levels independent of cell cycle progression and can therefore form stable complexes with only one specific tCDK partner. Eventhough this long-standing dogma exists, analytical size exclusion chromatography was used as an unbiased approach to re-evaluate human tCDK/Cyclin complex formation between the recombinant tCDKs Cdk7, Cdk8, Cdk9, Cdk11 (isoform p110), and Cdk13, and the tCyclins CycC, CycH, CycK, and CycT1 (Fig.25a). Cdk12, CycL1 and CycL2 were instable in their monomeric form and were therefore excluded from this experiment. Constructs of tCDKs and tCyclins comprised at least the kinase domain including relevant binding regions and both

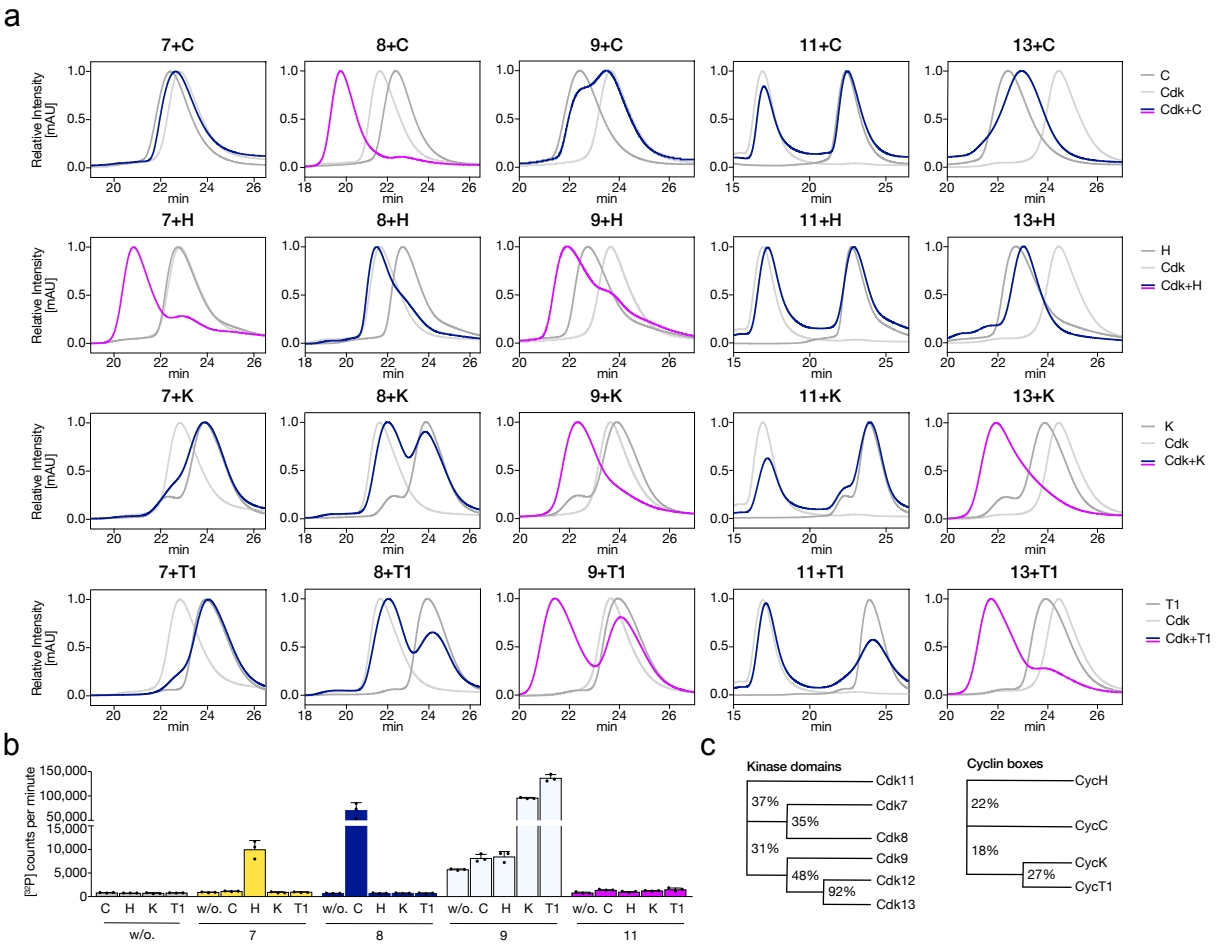

Fig.25 | Interaction and activity of tCDK/Cyclin complexes. a Analytical size exclusion chromatography was performed to analyze stable complex formation between the human tCDKs 7, 8, 9, 11_{p110}, and 13 and the human tCyclins C, H, K, and T1. CDK/Cyclin interaction results in a molecular weight shift (magenta) compared to the retention volume of non-interacting complexes (blue) or monomeric CDKs and Cyclins (grey). **b** In vitro kinase assays were performed with 0.2 mM [^{32P}]-γ-ATP and 10 µM pol II CTD for 20 min to test for the activation of 0.2 µM tCDKs by 0.4 µM Cyclins. All data are presented as mean ± SD of triplicates. **c** Phylogenetic tree illustrating the sequence identity of human tCDK kinase domains and tCyclin boxes (performed with Geneious).

Cyclin boxes, respectively. The exact construct length and purification strategies of all recombinant proteins used in this study are depicted in Appendix B. Apart from the cognate Cdk/Cyclin pairs Cdk7/CycH, Cdk8/CycC, Cdk9/CycT1, and Cdk13/CycK, novel complex assembly was observed for Cdk9/CycH, Cdk9/CycK and Cdk13/CycT1.

To test, whether these novel CDK/Cyclin pairs do not only interact, but also form active kinase complexes, in vitro kinase assays with [^{32}P]-γ-ATP were performed (Fig.25b). Cdk7, Cdk8, Cdk9 and Cdk11$_{p110}$ were tested with and without addition of CycC, CycH, CycK, and CycT1 for their activity towards RNA pol II CTD. Cdk13 was excluded from this experiment, as it requires additional phosphorylation at a conserved T-loop threonine residue for kinase activation. The cognate tCDK/Cyclin pairs Cdk7/CycH, Cdk8/CycC, Cdk9/CycT1 showed robust kinase activity, as did Cdk9/CycK. In contrast, Cdk9/CycH does not form an active kinase complex.

In line with these results, sequence identities of tCDK kinase domains and tCyclin boxes (Fig.25c) reveal a close relationship of Cdk7 and Cdk8 (35%), as well as of CycC and CycH (22%), and on the other hand of Cdk9 (48% shared identity with Cdk12/13), Cdk12 and Cdk13 (92%), and CycK and CycT1 (27%), suggesting two groups of tCDK and tCyclins in accordance with their functions in the transcription cycle: the initiation tCDKs, Cdk7/CycH and Cdk8/CycC, which indeed interact only with their cognate partner, and the elongation tCDKs, Cdk9, Cdk12, and Cdk13 that can be activated by both Cyclins, CycK and CycT1.

To confirm that not only Cdk9 can be activated by both Cyclins, but also Cdk12 and Cdk13, complex formation and kinase activity assays were performed in a different experimental set up. Because Cdk12 and Cdk13 require additional phosphorylation at a conserved T-loop threonine residue for kinase activation, the three tCDKs were co-expressed together with either CycK or CycT1 and a CDK-activating kinase 1 (CAK1) in *Sf9* insect cells and purified by affinity and size exclusion chromatography. SDS-PAGE with Coomassie staining demonstrated a 1:1 stoichiometry of all six complexes (Fig.26a). CDK/Cyclin pairs were tested for their activity towards RNA pol II CTD in in vitro kinase assays with [^{32}P]-γ-ATP, and controlled for autophosphorylation (kinase without substrate) and possible contamination by other kinases from *Sf9* cells (kinase-dead mutation; Cdk9$_{D149N}$, Cdk12$_{D859N}$, Cdk13$_{D837N}$). Robust kinase activity was observed for all six CDK/Cyclin pairs, with higher activities

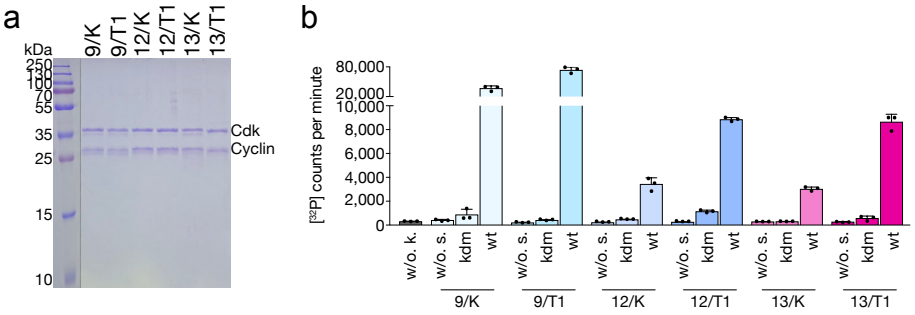

Fig.26 | Purification of catalytically active tCDK/Cyclin complexes. a Cdk9, Cdk12, and Cdk13 were co-expressed together with either CycK or CycT1 and CAK1 in *Sf9* cells. Complex formation was visualized after affinity and size exclusion chromatography by SDS-PAGE. **b** The six Cdk/Cyclin complexes (0.2 µM) were tested without substrate (w/o. s.), as a kinase-dead mutant (kdm; Cdk9$_{D149N}$, Cdk12$_{D859N}$, Cdk13$_{D837N}$) and wild-type (wt) in in vitro kinase assays with 0.2 mM [^{32P}]-γ-ATP and 10 µM pol II CTD. All data are presented as mean ± SD of triplicates.

observed for CycT1 than CycK complexes. However, we assume that this effect probably results from increased protein stability due to construct design rather than allowing any conclusion about in vivo activities.
It has recently been published that phosphorylation of CycT1 at Thr149 is essential for the assembly with Cdk9[168]. However, recombinant CycT1 phospho-mimetic mutants T149A and T149E showed no effect on complex formation and activity of P-TEFb as observed from dissociation constants in SPR experiments and kinase activities in radioactive in vitro kinase assays (Fig.27). Furthermore, purification of CycT1 mutant proteins was difficult to achieve due to strong protein aggregation, which might explain the decreased activation capability and consequently assumed degradation of CycT1 in cell experiments performed in this publication[168,169].

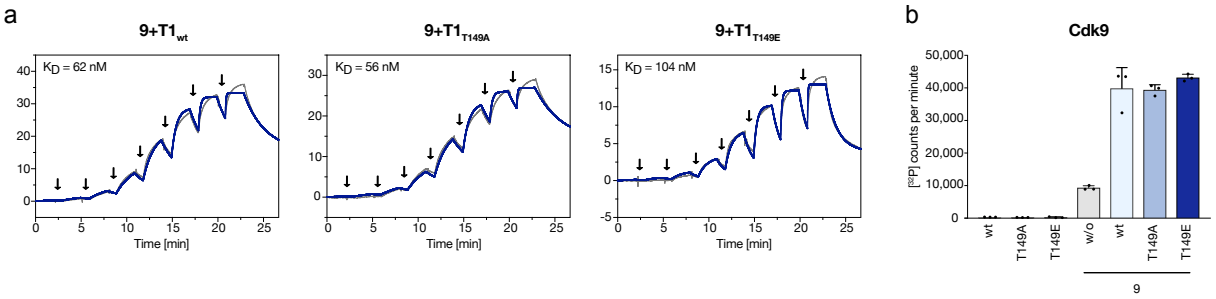

Fig.27 | CycT1 T149 phospho-mutants show no effect on binding and activation of Cdk9. a SPR measurements were performed using single-cycle kinetics. Cdk9 was immobilized on the chip surface by amine coupling. CycT1 wt, T149A or T149E mutant were injected in increasing concentrations of 2, 6, 18, 54, 162, 486, 1458, 4374 nM for 120 sec followed by dissociation for 300 sec. Dissociation constants (K_D) were determined based on a 1:1 interaction model. **b** In vitro kinase assays were performed with 0.2 mM [^{32P}]-γ-ATP and 10 µM pol II CTD to test for the activation of 0.2 µM Cdk9 by 0.4 µM CycT1 wt, T149A or T149E mutant or without (w/o) Cyclin. All data are presented as mean ± SD of triplicates.

In fact, we are not the first to have observed Cdk9/CycK complex formation. Soon after the discovery of P-TEFb in 1992[170], it has been shown that Cdk9 can be activated by CycT1, CycT2 or CycK and that all three complexes can phosphorylate RNA pol II CTD in vitro[171]. A few studies followed, demonstrating that Cdk9/CycK is implicated in the replication stress response in U2OS human osteosarcoma cells independently of Cdk9/CycT1[172], that Cdk9/CycK activates transcription via interaction with RNA, while Cdk9/CycT1 interacts with DNA[173], and the crystal structure of monomeric CycK was resolved and modelled together with Cdk9 to study its Cdk/Cyclin interface[174]. However, in 2010, soon after the discovery of Cdk12 and Cdk13, two studies from the Peterlin and Greenleaf Lab showed that CycK is the cognate Cyclin partner of these newly identified kinases[133,175]. Bartkowiak et al.[175] co-immunoprecipitated endogenous Cdk12 and Cdk13 together with CycK from *D. melanogaster* Kc nuclear extracts, whereas Blazek et al.[133,176] demonstrated that Cdk9 binds exclusively to CycT1, and Cdk12 and Cdk13 exclusively to CycK by overexpression and co-immunoprecipitation of FLAG- and HA-tagged Cdk9, Cdk12 and Cdk13 from HEK293 cells. From this point on, the Cdk9/CycK complex has been virtually ignored in the field. Nonetheless, a recent study found that loss of hnRNP R, a known 7SK interactor, promotes the release of P-TEFb from its regulator 7SK and stabilizes Cdk9 and its association with CycK as demonstrated by co-immunoprecipitation from HeLa cells[177].

Furthermore, together with Megan Insco from the Dana-Farber Cancer Institute, Harvard Medical School and Leonard Zon from Boston Children´s Hospital, Boston, USA, we recently discovered a Cdk13/CycT1 complex that plays a role in nuclear RNA surveillance[178,179]. In this study, Insco et al. identified an enrichment of Cdk13 kinase domain mutations in publicly available melanoma patient data and found that Cdk13 mutations show loss of kinase activity, as well as accelerated melanoma onset in a zebrafish melanoma model and caused human melanoma cells to be more proliferative. They further demonstrated that mutant Cdk13 zebrafish melanomas and human melanoma cells had increased prematurely terminated RNAs (ptRNAs) ending in introns. They found that ptRNAs accumulate post-transcriptionally through lack of degradation[178].

To elucidate how Cdk13 mutations cause ptRNA accumulation, Cdk13 wt was transiently expressed with a V5-tag in A375 melanoma cells, immunoprecipitated from nuclear extracts (Fig.28a) and analyzed by mass spectrometry (MS) analysis (Fig.28b). In comparison to a V5 control, 37 interactors with Cdk13 wt were identified. Interestingly, these MS data showed interaction with CycT1 (Fig.28b-c), whereas cognate CycK was only detected at levels below thresholding and no other Cyclins were detected above background levels. Ontology analysis of Cdk13 wt interactors showed enrichment in "mRNA 3´end processing". Many of the enriched proteins were components or interactors of the polyA tail eXosome Targeting (PAXT) complex, which has crucial functions in targeting ptRNAs for degradation, f.e. ZC3H14 and PABPN1. Due to high levels of ZC3H14 in Cdk13 wt immunoprecipitations, ZC3H14 was further characterized by transient expression and immunoprecipitation from Cdk13 wt and mut A375 cells and analyzed by MS. ZC3H14 from Cdk13 wt cells showed four phosphorylation sites, whereas ZC3H14 from Cdk13 mut cells showed loss of phosphorylation exclusively on Ser475 (Fig.28d), suggesting direct phosphorylation by Cdk13. We demonstrated in radioactive in vitro kinase assays that Cdk13/CycT1 directly phosphorylates full-length ZC3H14 wt but not the phospho-mimetic S475A mut (Fig.28e). Additionally, we confirmed direct phosphorylation of ZC3H14 at S475 by Cdk13/CycT1 in in vitro kinase assays with ATP-γ-S and subsequent alkylation and visualization in immunoblots with an anti-thiophospho antibody of full-length ZC3H14 wt compared to S475A mut (Fig.28f). In addition, Insco et al. rescued PAXT recruitment and activity by phospho-mimetic ZC3H14 S475D expression in Cdk13 mut cells. They showed that the expression of a non-phosphorylatable ZC3H14 decreased PAXT recruitment and activation in Cdk13 wt cells. Recurrent mutations occur in additional PAXT components suggesting a broad, previously unrecognized tumor-suppressive role for nuclear RNA surveillance.

Overall, this study shows that mutation of Cdk13 accelerates melanomagenesis via deficient RNA surveillance pathways. Lack of Cdk13 activity leads to loss of ZC3H14 phosphorylation and deficient recruitment of PAXT components to sites of ptRNA degradation, promoting aberrant RNA accumulation and consequently tumor progression. This is the first study demonstrating that Cdk13 has tumor suppressor functions and that Cdk13 can be activated by CycT1.

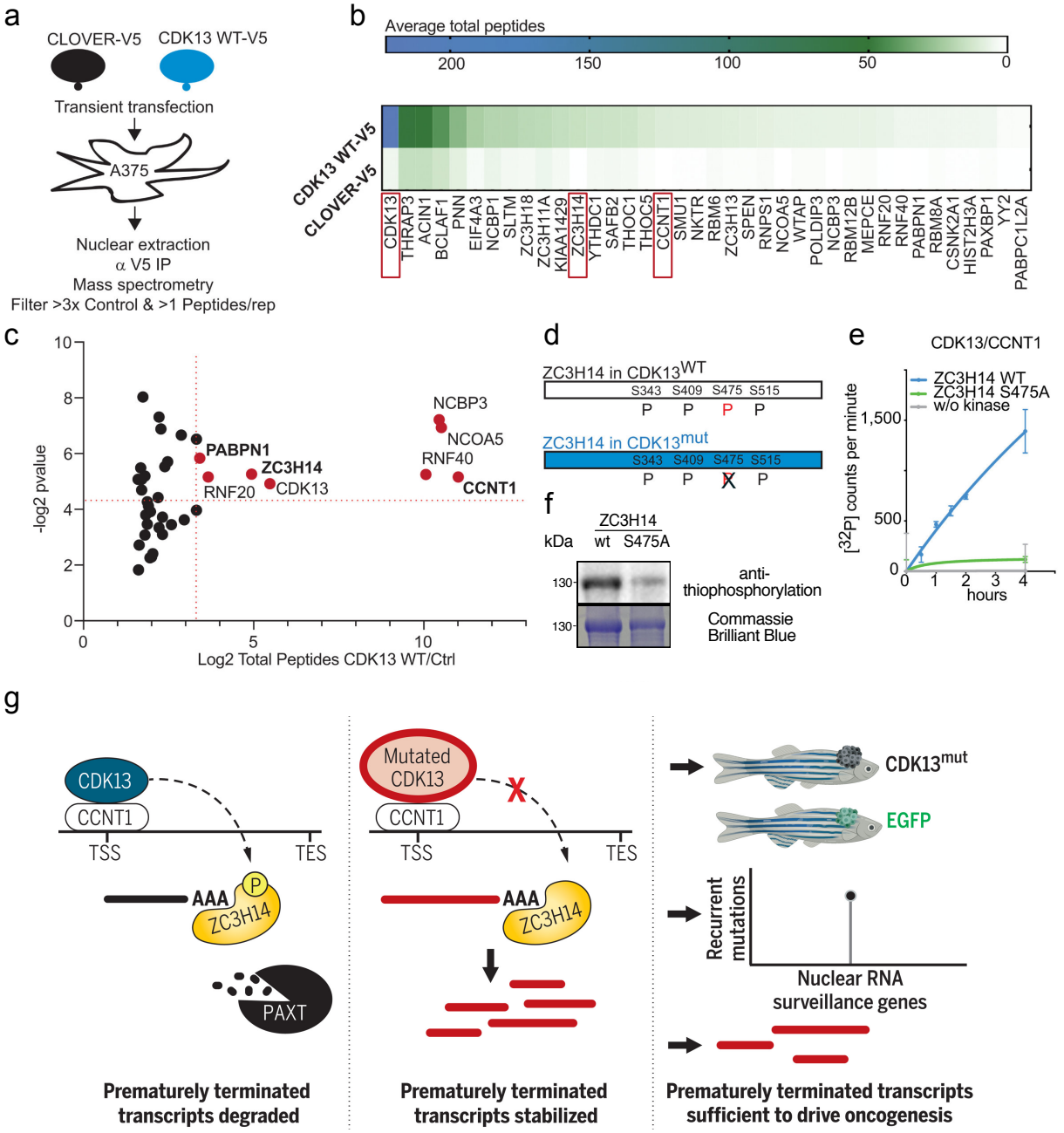

Fig.28 | Cdk13/CycT1 phosphorylates the PAXT complex component ZC3H14 at Ser475. **a** Cdk13 IP-MS scheme. **b** Heatmap of average total peptides from anti-V5 IP-MS of Cdk13 wt-V5 (n=3) or CLOVER-V5 (n=2). **c** Log2-fold change of Cdk13 wt vs. control total peptides by -logP value (two-tailed t-test). Upper right quadrant indicated proteins with a -log P value <0.05 and a log-fold change enrichment over control >3.3. **d** Scheme illustrating ZC3H14 phosphorylation sites identified from ZC3H14-IP from either Cdk13 wt or mut-expressing human melanoma cells. **e** Radioactive in vitro kinase assay of Cdk13/CycT1 phosphorylating full-length ZC3H14 wt and S475A mut. (p=0.007, two-tailed t-test, n=3). **f** In vitro kinase assay of Cdk13/CycT1 phosphorylating full-length ZC3H14 wt and S475A mut with ATP-γ-S, subsequent alkylation, and visualization in immunoblot by anti-thiophospho antibody. **g** Graphical Abstract. In healthy cells, Cdk13/CycT1 phosphorylates ZC3H14 to activate nuclear RNA surveillance on protein-coding genes. Mutant Cdk13 fails to phosphorylate ZC3H14, and aberrant RNAs are stabilized. Patient data and zebrafish melanoma models demonstrate that oncogenic Cdk13 mutations lead to deficient nuclear RNA surveillance and drive melanomagenesis (modified from Insco et al., 2023; experiments a-c were performed by Megan Insco et al.)[178].

3.2.2 Complex formation and activity of novel tCdk/Cyclin pairs in vivo

To exclude the possibility that novel CDK/Cyclin pairs interact exclusively in vitro, co-immunoprecipitation of endogenous CycK and CycT1 was performed from HEK293T and A375 melanoma whole cell lysates. The cognate Cdk9/CycT1, Cdk12/CycK, Cdk13/CycK, as well as the novel Cdk13/CycT1 complex were detected by immunoblotting. These results demonstrate that both Cdk13 pairs can co-exist in cells. Moreover, all five proteins are expressed in both cell lines, suggesting that the formation of Cdk13/Cyclin complexes is depended on a mechanism other than expression levels. This is in line with their functions as essential regulators of transcription and the database Human Protein Atlas (https://www.proteinatlas.org), which specifies ubiquitous expression of these five CDKs and Cyclins across tissues. Although two different cell lines were tested for this experiment with similar results, it cannot be excluded that different Cdk/Cyclin pairs exist in different cell lines. However, whether a Cdk9/CycK or a Cdk12/CycT1 complex exist in cells remains unclear and there are many plausible reasons why they are not present in the tested cell lines, which are discussed in the following sections.

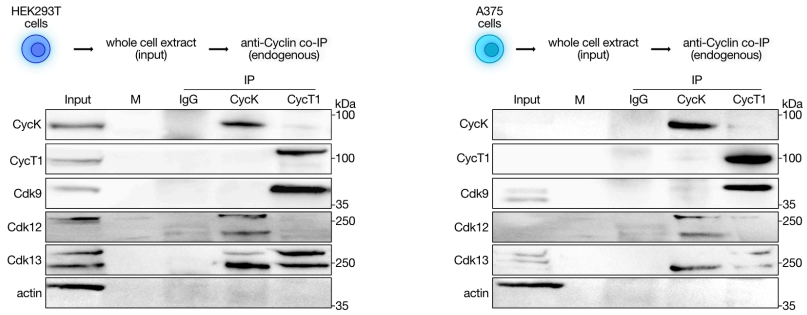

Fig.29 | Cdk13/CycK and Cdk13/CycT1 co-exist in cells. Co-immunoprecipitation of endogenous CycK and CycT1 from HEK293T and A375 whole cell lysates; M: Marker; IgG: rabbit IgG control. This experiment was repeated two times.

Following up on the Cdk9/CycK complex, in a recent study from the Johnstone Lab (Fan & Devlin et al., 2020)[136], the Cancer Dependency Map (DepMap) from the Broad Institute was mined for the dependency on CDKs and Cyclins for growth and survival of 62 malignant cancer cell lines of hematological origin from the Avana CRISPR public 19Q1 dataset. This analysis demonstrated the importance of Cdk9 and CycK for the viability of all cancer cell lines tested, in contrast to CycT1 and Cdk12/13 (Fig.30a). Interestingly, all cell lines tested are highly dependent on Cdk9, but not on its cognate activator CycT1, suggesting that either the analyzed cell lines depend on Cdk9 but not on Cdk9 activity, that another kinase compensates for loss of Cdk9 activity, or that Cdk9 may be activated by a Cyclin other than CycT1. To further analyze this observation, CRISPR-mediated gene editing was used to introduce a functional Knock-out (KO) for the Cdk9 and CycT1 gene in HEK293T cells (Fig.30b). Three days after nucleofection with small guide (sg) RNAs and recombinant Cas9, Knock-outs were confirmed by immunoblots from whole cell lysates (Fig.30c). As predicted from the DepMap analysis, $Cdk9_{KO}$ cells died approx. 4 days after nucleofection, in contrast to $CycT1_{KO}$ cells, which showed no difference in cell growth and viability compared to AVVs control cells (data not shown). To specifically analyze the activity of Cdk9 in $CycT1_{KO}$ cells,

AVVs controls and CycT1$_{KO}$ cells were treated with 0.25 µM of the selective Cdk9 degrader THAL-SNS-032 for 5 hours, followed by cell lysis and immunoblotting (Fig.30d). RNA pol II CTD phospho-specific antibodies reveal that CycT1$_{KO}$ cells exhibit similar levels of pol II CTD Ser2 and Ser5 phosphorylation as AVVs controls, and that phospho-levels are dependent on Cdk9, as they are lost after Cdk9 degradation. These data suggest that Cdk9 is still active despite the loss of CycT1, and is presumably activated by another Cyclin so that pause-release of pol II is still maintained. Early studies showed that Cdk9 can be activated by CycT1, CycT2 or CycK[171]. Therefore, further experiments such as Cdk9 co-IPs from CycT1$_{KO}$ cells are required to unravel which Cyclin compensates the loss of CycT1 for Cdk9 activity and whether it is a Cdk9/CycK complex.

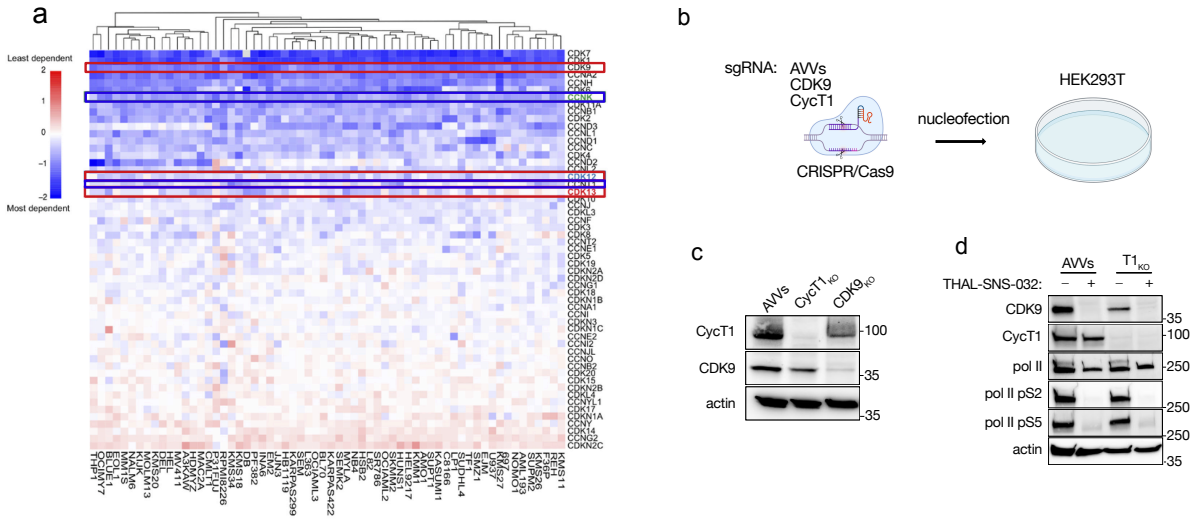

Fig.30 | CycT1 Knock-out in HEK293T cells indicate activation of Cdk9 by a different Cyclin partner. **a** Heat map of dependency scores in 62 hematological cancer cell lines for 55 CDKs and Cyclins from the Avana CRISPR public 19Q1 dataset. CycK and T1 are highlighted in blue, Cdk9, 12 and 13 are highlighted red (modified after Fan & Devlin et al., 2020)[136]. **b** Scheme illustrating generation of Cdk9 and CycT1 Knock-outs (KO) in HEK293T cells with small guide (sg) RNAs and recombinant Cas9 by nucleofection. **c** Immunoblotting confirms Cdk9 and CycT1 Knock-out from cell lysates 3 days after nucleofection. **d** Cdk9 was degraded using 2.5 µM THAL-SNS-032 for 5 hours in HEK293T AVVs control and T1 Knock-out cell lines. Immunoblots show Cdk9 degradation and the phospho-status of pol II confirming that Cdk9 activity is responsible for pol II Ser2 and Ser5 phosphorylation even in the absence of CycT1. This experiment was repeated two times.

3.2.3 Regulation of tCDK/Cyclin pairs by Hexim1 and Brd4

Eventhough Cdk13 complex formation seems to be regulated independently of CycK and CycT1 protein levels, this may not be the case for Cdk9/CycK and Cdk12/CycT1, especially in light of the equilibrium of free and sequestered Cdk9 by the 7SK snRNP complex[36]. In contrast to Cdk12 and Cdk13, Cdk9 regulation has been intensively studied in recent years. Cdk9/CycT1 is sequestered by the 7SK complex through inhibition by Hexim1 or 2 (Hex)[67] and, on the other hand, can assemble with Brd4 or components of the Super Elongation Complex (SEC), which antagonize the inhibition of the 7SK complex and recruit P-TEFb to open chromatin regions. Because Hexim1 has been shown to bind P-TEFb through interactions with CycT1[180,181], in vitro kinase assays with [^{32}P]-γ-ATP were performed with the six CDK/Cyclin pairs, to test whether Cdk9/CycK, as well as Cdk12 and Cdk13/CycT1 pairs

might be similarly regulated by Hexim1 and Brd4 as P-TEFb (Fig.31). In line with the literature, Cdk9/CycT1 is significantly inhibited by Hex1 (approx. -75% activity), activated by addition of Brd4 (approx. +40% activity), and the inhibitory effect of Hex1 is antagonized by Brd4 (approx. 100% activity). However, Cdk9/CycK was inhibited by Hex1 to a much smaller extend (approx. -25% activity) and ,similarly, less well activated (approx. +10% activity) and antagonized (approx. -25% activity) by Brd4, indicating a contribution of Cdk9, and not only CycT1 to the interaction with Hex1. This is consistent with a study demonstrating that Hex1 and 2 compose a proline, tyrosine, arginine, threonine (PYNT) motif that targets and inhibits the ATP binding-pocket of Cdk9[182]. Regulatory effects of the two P-TEFb binding partners are barely pronounced in both Cdk12 and Cdk13/Cyc pairs. Both CycT1 complexes showed poor Hex1 inhibition (approx. -10 to 25% activity) and no positive effect of Brd4 on kinase activity. These data indicate a different regulatory mechanism for Cdk9/CycK, as well as Cdk12 and Cdk13/Cyclin complexes in comparison to P-TEFb. Furthermore, it seems that Hex1 interacts with CycT1, but not necessarily with CycK. As 50-90% of Cdk9/CycT1 is sequestered by Hexim1 in the 7SK snRNP complex[67,68], it might be possible that interaction of Hex1 with Cdk9 prevents complex formation with CycK and might explain why a Cdk9/CycK complex was not detected in co-IPs from HEK293T cells. Further experiments with Hex1 Knock-out cells, or treatment with UV light, actinomycin D, or P-TEFb inhibitors, which have been shown to release P-TEFb from the 7SK complex upon transcriptional inhibition[64], could contribute to a better understanding of Cdk9/CycK complex formation.

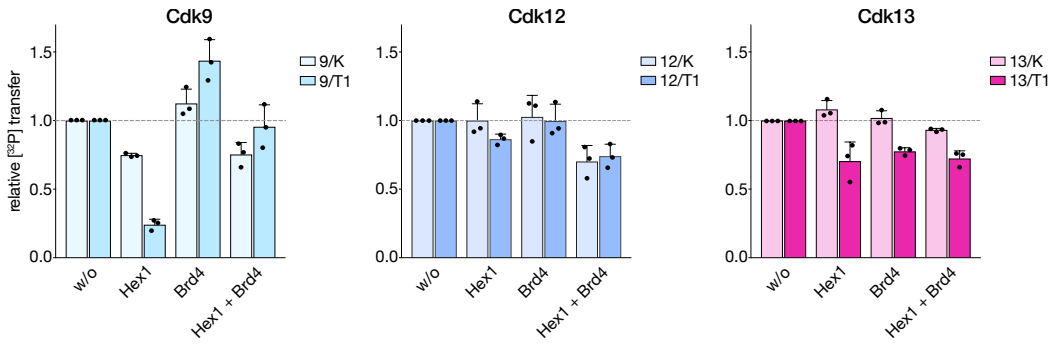

Fig.31 | Regulation of tCDK/Cyc pairs by Hex1 and Brd4. Hexim1 (Hex1) and Brd4 are regulators of Cdk9/CycT1, but show little effects on Cdk9/CycK and Cdk12 or Cdk13/CycK or CycT1 pairs. For radioactive kinase assays, 0.2 mM [^{32}P]-γ-ATP was incubated with 0.2 µM kinase, 10 µM of human c-Myc, 2 µM Hex1 and/or 10 µM Brd4 for 15 min. All data are presented as mean relative to kinase activity without regulator ± SD of triplicates.

3.2.4 Changes of tCyclin levels

It is common sense in the field that tCyclins are stably expressed throughout the cell cycle. On the other hand, cCyclins were discovered and named after their oscillating protein levels depending on the cell division phase and the controlled activation of several different cCDKs. Accordingly, cCyclins are modified by ubiquitination through members of the E3 ligase families of the Anaphase-Promoting Complex (APC/C) or Skp1, Cullins, F-box proteins (SCF) for degradation by the proteasome depending on the respective cell cycle phase[65]. Both E3

ligase families recognize a small sequence motif for ubiquitination, called degron. In silico analysis with degronopedia (https://degronopedia.com), a database for protein degron motifs, identified two to five degron motifs in the amino acid sequence of the cCyclins, CycA2, B1, D1, and E1 (Fig.32). Similarly, analysis of the tCyclins CycC, H, K, L1, L2, and T1 amino acid sequences reveal two to seven degron motifs per tCyclin. The existence of such degron motifs in tCyclins to a similar extend as in cCyclins suggests that even though tCyclins have not been observed to be degraded depending on cell cycle phases, they may be degraded due to other cellular processes, such as stress response, proliferation or global transcriptional changes. Although constant tCyclin levels are a consent in the field, a few experimental studies demonstrate changes in tCyclin levels under different cellular conditions. CycT1 expression levels change during T cell maturation[183,184] and CycT1 is degraded by the E3 ligase SCF^{SKP2} recognizing the PEST domain of CycT1[169] and by the E3 ligase Siah1/2[185]. Also yeast CycC is ubiquitinated, followed by proteasomal degradation upon exposure to various forms of cell stress including oxidative or osmotic stress, nitrogen starvation, ethanol treatment, heat shock, and nutrient deprivation or when grown on non-fermentable carbon, leading to relieve of transcriptional repression of stress response genes[186,187,188]. Similarly, Med13, a component of the Cdk8/CycC/Mediator kinase module, is proteasomally degraded after oxidative stress exposure through SCF E3 ligases[189,190]. However, whether the degradation of tCyclins leads to complex formation with a different Cyclin partner, such as one of the novel Cdk/Cyclin pairs, or is simply a regulatory mechanism for CDK inactivation remains to be determined.

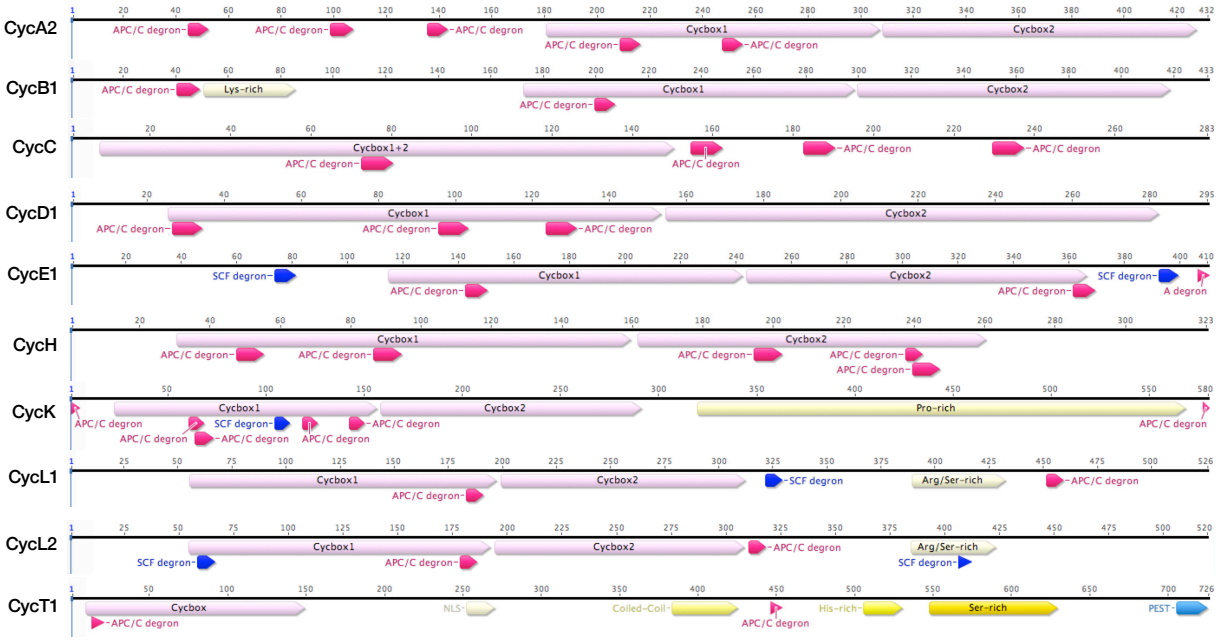

Fig.32 | Cyclin degron motifs for APC/C and SCF E3 ligases. Analysis of degron motifs with degronopredia (https://degronopedia.com) reveals degron motifs for APC/C (pink) and SCF (dark blue) E3 ligases not only in the amino acid sequence of cCyclins, but also in tCyclins, suggesting that they are proteasomally degraded. Cyclin boxes are highlighted in rose and additional domains in yellow. Graphs were designed with Geneious.

Natural proteasomal degradation of CycK in HEK293T cells

In addition to natural Cyclin degradation, in recent years five publications demonstrated the development of molecular glue degraders targeting Cdk12/13 leading to CycK proteasomal degradation[191–195]. Molecular glue degraders cannot be developed by rational design but are identified by chemical screening approaches using CRISPR/Cas9 gene knock-out cell lines with broadly abrogated E3 ligase activity[117]. Interestingly, in all five studies molecular glues were identified targeting a Cullin4 E3 ligase, mediated by the adapter protein DNA damage-binding protein 1 (DDB1). That five independent studies discover compounds targeting CycK for proteasomal degradation using rather serendipitous screening approaches might indicate that CycK is a natural target for E3 ligases. In line with this idea, DDB1 is a predicted E3 ligase of CycK as revealed by in silico analysis with degronopedia.
Furthermore, time course experiments in HEK293T cells with 10 µM of the proteasomal inhibitor bortezomib for 1, 2, or 6 hours reveal increasing protein levels of CycK upon proteasomal inhibition in comparison to a DMSO control (Fig.33). In contrast, CycT1 levels are not affected. Thus it seems, that there might be a constant turn-over of CycK proteasomal degradation in cells. Whether proteasomal degradation might only affect Cdk9/CycK, but not Cdk12/13/CycK complexes needs to be addressed in future experiments. Moreover, it has been shown that CycK has crucial functions in DNA damage response (DDR) and further controls are needed to exclude the possibility that bortezomib induces DDR, which in turn could be responsible for the observed increase in CycK levels. Thus, staining of immunoblots for DDR markers, such as H2AX, would be important.

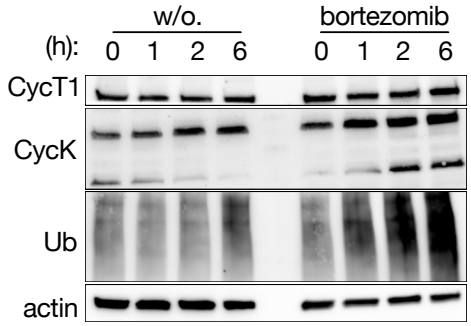

Fig.33 | Natural proteasomal degradation of CycK in HEK293T cells. Cell were treated with 10 µM of the proteasomal inhibitor bortezomib for 0, 1, 2, and 6 hours and whole cell extracts were analyzed by immunoblotting.

Changes of CycT1 protein levels in THP1 cells

As CycT1 expression level have been reported to be upregulated upon T cell maturation[183,184], THP1 cells, which are able to differentiate into macrophage-like cells, were analyzed for CycT1 protein level changes (Fig.34). THP1 resuspension cells were primed by addition of 50 ng/µl phorbol-12-myristate-13-acetate (PMA) into adherent THP1 cells, and after a 12 h rest, were treated with 200 ng/µl lipopolysaccharide (LPS) for 4 h for differentiation into macrophage-like cells. Whole cell lysates of untreated and differentiated THP1 cells were analyzed by immunoblotting for Cdk9, CycK, and CycT1. In contrast to Cdk9 and CycK, CycT1 protein levels increase upon THP1 differentiation. These data indicate that CycT1 levels might be similarly regulated during THP1 differentiation as previously reported for T cell

maturation[183,184]. Whether this effect is caused by upregulation of CycT1, or whether CycT1 might be constantly degraded in undifferentiated THP1 cells, and how this affects Cdk/Cyc pairs needs to be examined in future experiments using proteasomal inhibitiors.

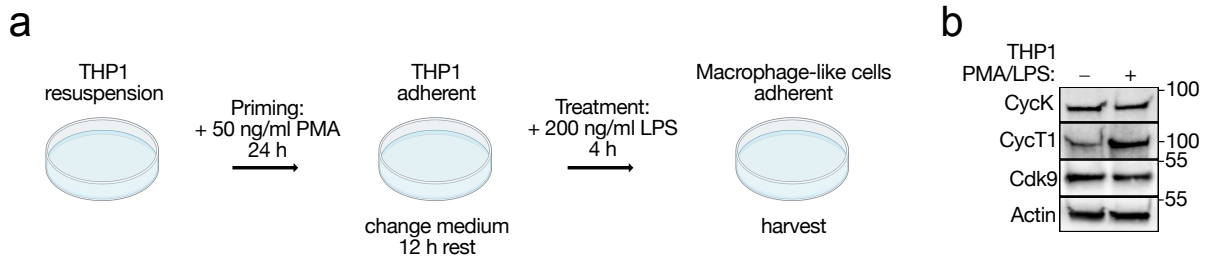

Fig.34 | CycT1 protein levels increase upon THP1 differentiation. a Scheme of THP1 cell differentiation by priming with phorbol-12-myristate-13-acetate (PMA) and treatment with lipopolysaccharide (LPS). **b** Whole cell extracts from treated and untreated THP1 cells were stained in immunoblots for protein levels of Cdk9, CycK and CycT1. This experiment was repeated two times.

Because CycT1 expression or degradation levels seem to be regulated differently in immune cells, a fluorescent ubiquitination-based cell cycle indicator (FUCCI) cell line was used to study changes in CycT1 protein levels throughout the cell cycle in THP1 cells. FUCCI cells were designed for cell cycle studies by expression of mKO2, a red fluorescent protein, fused to human Cdt1 (residue 30-120) specifically in the G1 phase, and mAG, a green fluorescent protein, fused with human Geminin (residue 1-110), expressed in S/G2/M phase (Fig. 35)[196]. As a result, FUCCI cells emit red and green fluorescence in the G1, S and G2/M phase, respectively, allowing to visualize cell cycle phases in individual live cells.
THP1 FUCCI cells were sorted by FACS into different fractions containing cells in the G1, S, and G2/M phase, respectively. Immunoblots of whole cell lysates demonstrate that CycT1 protein levels are increased in the G1 phase compared to the other phases and unsorted cells (bulk). Moreover, it would be interesting to repeat the experiment in a non-immune cell line,

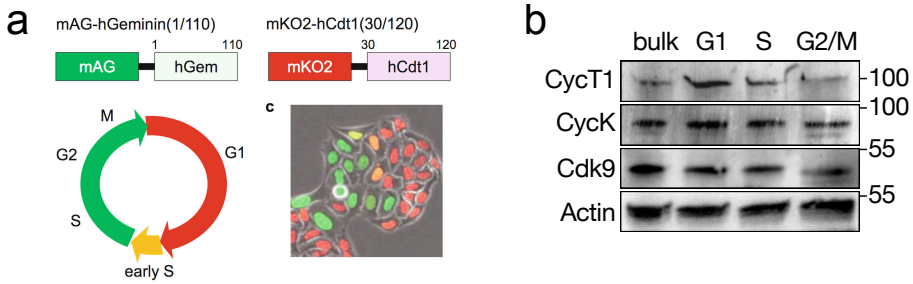

Fig.35 | CycT1 protein levels are elevated in the G1 cell cycle phase of THP1 FUCCI cells. a Scheme of FUCCI cell lines. The cell cycle dependent expression of fluorescent mAG-human Geminin (hGem) in S/G2/M and of mKO2-human Cdt1 in G1 allows visualization cell cycle phases (Coutts & Weston, 2021)[196]. **b** THP1 FUCCI cells were sorted by FACS into fractions of cells in the G1, S, and G2/M phase, respectively, or unsorted (bulk), and stained in immunoblots. FACS sorting was performed by the Peter MacCallum FACS Core Facility and FACS data were analyzed by Jennifer Devlin, PhD, Peter MacCallum Cancer Centre, Melbourne. This experiment was repeated two times.

such as HEK293T cells, to re-evaluate whether tCyclin levels do not change upon cell cycle phases in a FUCCI cell line compared to synchronization of cell cycle by chemical inhibition.

However, whether changes in Cyclin protein levels influence Cdk activity and complex formation needs to be determined in future studies.

3.2.5 ChIP-seq analysis of tCDKs and Cyclins

Chromatin immunoprecipitation sequencing (ChIP-seq) is a common technique to map the location of tCDKs genome-wide or at the gene body of a specific region, but has been rarely performed for tCyclins. To assess potential functional differences between the cognate Cdk13/CycK and novel Cdk13/CycT1 pairs on the gene level, a protocol for total ChIP-seq was established for HEK293T cells using antibodies against pol II, Cdk9, Cdk12, Cdk13, CycK and CycT1 (Fig.36). While ChIP of pol II, Cdk9, Cdk13 and CycT1 was successful, Cdk12 and CycK showed comparatively low numbers of annotated peaks, probably due to poor antibody quality. Analysis of distinct genomic features demonstrate strong similarities between the three CDKs, most of which are located at promoters near the TSS (<=1kb). Cdk9 and Cdk12 show highly identical profiles (even though low peak numbers of Cdk12 have to be taken into consideration). A large fraction of Cdk13 is located downstream of the TSS (<=300kb), whereas Cdk9 is rather mapped to distal intergenic regions. Both Cyclins, as well as pol II show less occupancy of promoter regions close to the TSS, but larger distribution across diverse genomic regions. CycK shows a similar distribution pattern as pol II but is more common on 3´UTR regions than on distal intergenic regions, but overall low peak numbers of CycK have to be taken into consideration. The profile of CycT1 shares high similarity with Cdk9, and is not as common at regions downstream of the TSS (<=300kb) as is Cdk13.

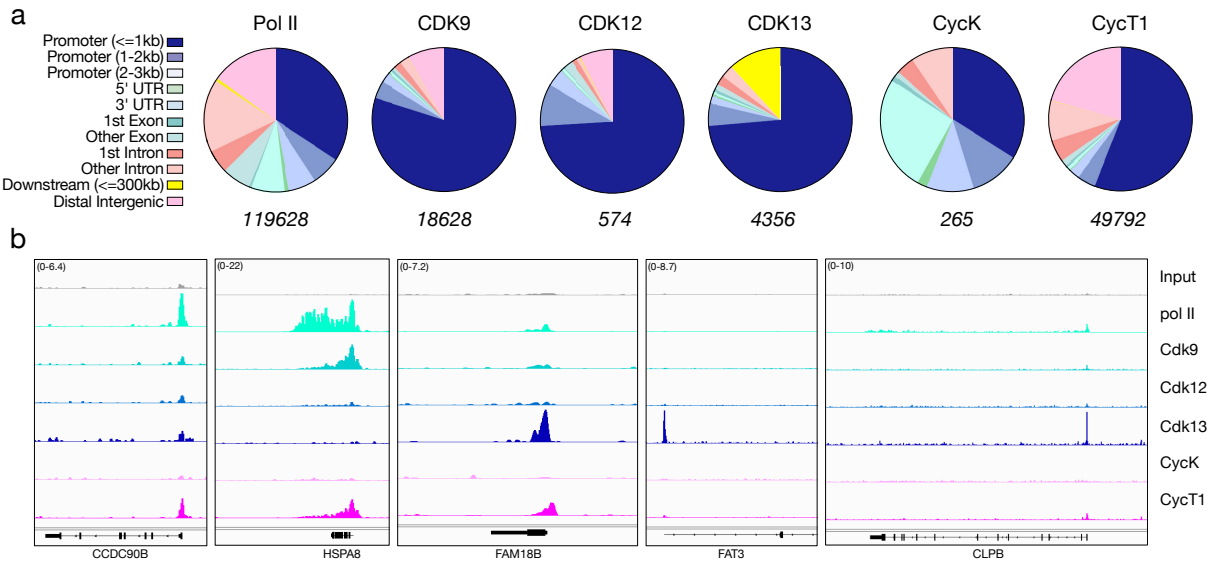

Fig.36 | Genome-wide ChIP-seq analysis of tCDKs and Cyclins. ChIP-seq analysis was performed from HEK293T cells. **a** Graphical representation of how tCDK and tCyclin ChIP regions are distributed over distinct genomic features. The number of annotated peaks is depicted in italic. **b** Representative IGV profiles of selected genes indicating gene-specific differences of Cdk9 and Cdk13/CycT1 pairs.

Analysis of the localization at selected genes indicates gene-specific regulation by the different tCDKs and tCyclins. At CCDC90B all proteins are located at the TSS, whereas

HSPA8 shows specific occupancy of pol II, Cdk9 and CycT1. Nonetheless, FAM18B demonstrates co-localization of Cdk13 and CycT1 at the TSS, whereas pol II and Cdk9 peaks are low. At some genes, such as FAT3 and CLPB, Cdk13 shows high occupancy, whereas all other proteins are absent at these regions, however, the reason is unclear.
To further analyze differences between Cdk13/CycK and Cdk13/CycT1 complexes on the gene level, future experiment will benefit from usage of different antibodies for Cdk12 and CycK, performance of ChIP-seq from HEK293T $T1_{KO}$ cells in comparison to an AVVs control, as well as ChIP-re-ChIP, in which first CycT1 will be immunoprecipitated and then split for an additional immunoprecipitation of either pol II, Cdk9, or Cdk13.

3.2.6 Substrate specificity

CDKs recognize substrates with the canonical amino acid sequence S/TPXK/R, where S is the phosphorylatable serine and X is any amino acid[58]. However, recent findings show that tCDKs do not strictly depend on this sequence[17,41]. Cyclins plays a role in CDK substrate recognition as they harbor remote binding pockets recognizing specific substrate binding motifs such as the RXL motif in CycA2[60]. For cell cycle CDKs it has been shown that the Cdk2/CycA complex phosphorylates a broad range of substrates, whereas significantly fewer substrates are recognized by Cdk2/CycE[197]. However, little is known about the involvement of Cyclins in tCDK substrate recognition. Since the consent in the field is dominated by the idea that tCDKs interact with only one particular Cyclin, it has never been investigated whether different Cyclins influence the substrate recognition of their CDK partner. To this end, established tCDK substrates were tested for phosphorylation by different tCDK/Cyclin pairs in in vitro kinase assays with [^{32}P]-γ-ATP (Fig.37). The initiation CDKs, Cdk7/CycH without (7/H) or with its regulator Mat1 (7/H/M), Cdk8/CycC without (8/C) or with its regulator Med12 (8/C/M), and Cdk19/CycC without (19/C) or with Med12 (19/C/M), as well as the elongation CDKs, Cdk9, Cdk12 and Cdk13 with CycK (9/K; 12/K; 13/K) or CycT1 (9/T1; 12/T1; 13/T1), were tested for their activity toward RNA pol II CTD, the transcription factor c-Myc, the elongation factors SPT5 and SPT6, the splicing factors Cell Division Cycle 5 Like (CDC5L), Spliceosome Factor 3B subunit 1 (SF3B1), Serine and arginine-rich splicing factors (SRSF) SRSF1, SRSF2, and SRFS7, as well as the cleavage and polyadenylation factor Cleavage Stimulation Factor 2 (CSTF2), and the termination factor 5´-3´Exoribonuclease 2 (XRN2). Overall, the initiation CDKs show no changes in substrate recognition when associated with their regulators Mat1 and Med12, but have significantly increased activities (Fig.37). Cdk7 is highly specific in its phosphorylation activity for pol II CTD, amplified by the RING-finger protein Mat1, and shows some activity above background toward SPT5. Interestingly, Cdk7/CycH without Mat1 shows higher activity toward the transcription factor c-Myc. Also, the paralogues Cdk8 and Cdk19 have similar substrate preferences with high activities toward pol II CTD and c-Myc, amplified by Mediator subunit Med12. Low activities of both kinases toward SF3B1 are visible above background, in contrast to strong activities toward SPT5. The pausing factor SPT5 is known to be phosphorylated by P-TEFb for pol II pause release, but this assay demonstrates much higher activity levels of Cdk8/19 toward SPT5 than Cdk9, even though levels of pol II activities are comparable between the three CDKs. However, a role of Cdk8/19 in SPT5 phosphorylation or pause release has not been described yet.
The elongation kinases Cdk9, Cdk12 and Cdk13 show highly similar substrate preferences independent of the bound Cyclin, with robust activities toward pol II CTD, c-Myc, SPT5, SPT6,

CDC5L, and SF3B1. In comparison to the initiation kinases Cdk7, Cdk8 and Cdk19, the splicing factor SF3B1 is clearly a substrate of CDKs in the later phases of the transcription cycle, similar to SPT6 and CDC5L, suggesting importance of these substrates in pause release, elongation or RNA processing. Again, higher activities were observed for the three CDKs when assembled with CycT1 in comparison to CycK. This effect probably results from increased protein stability rather than allowing any comparison between the complexes. Interestingly, the only exception from this observation are the Cdk12 and Cdk13/CycK complexes, which show significantly increased phosphorylation toward c-Myc than CycT1 pairs or Cdk9/CycK. The functional relevance of Cdk12 or Cdk13 phosphorylating c-Myc is, however, unknown.

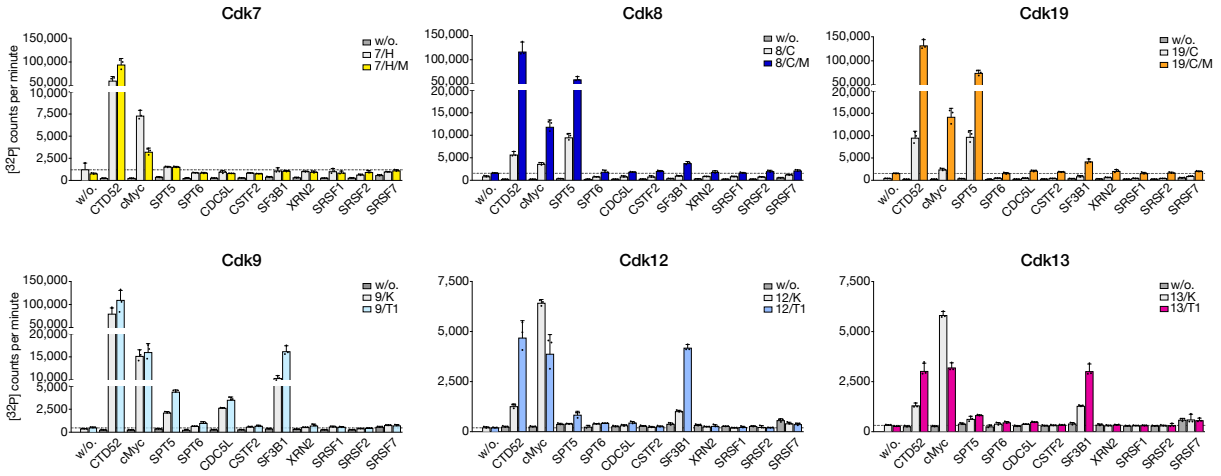

Fig.37 | Substrate phosphorylation by tCDK/Cyclin pairs. The initiation CDKs, Cdk7/CycH without (7/H) or with Mat1 (7/H/M), Cdk8/CycC without (8/C) or with Med12 (8/C/M), and Cdk19/CycC without (19/C) or with Med12 (19/C/M), as well as Cdk9 with CycK (9/K) or CycT1 (9/T1), Cdk12 with CycK (12/K) or CycT1 (12/T1), and Cdk13 with CycK (13/K) or CycT1 (13/T1) were tested for their substrate specificity in in vitro kinase assays. Established tCDK substrates were incubated at 10 µM concentration with 0.2 mM [^{32}P]-γ-ATP and 0.5 µM kinase for 30 min. All data are presented as mean ± SD of triplicates.

tCDK specificity toward RNA pol II CTD heptad residues

To further define the specificity of the different tCDK complexes toward the exact phospho-residues on the pol II CTD heptad repeats, in vitro kinase assays were performed as a time-course followed by Coomassie staining of SDS gels or immunoblotting using monoclonal antibodies against pSer2, pSer5, and pSer7 residues (Fig.38). Phosphorylation of pol II CTD is visible as migration band of pol II CTD from a hypo- (IIa) to a hyper-phosphorylated (IIo) state. Again, the CycK and CycT1 complexes showed no difference in their residue preference for pol II CTD with overall weaker signals for CycK complexes, which again correlates with reduced activity in comparison to CycT1. Interestingly, pSer2 levels were only visible for Cdk12 and Cdk13, but not for other tCDKs. All tCDKs show high Ser5 phosphorylation and robust phosphorylation of Ser7 with reduced levels of Cdk12 and Cdk13 pairs. Also Cdk19 activity toward Ser7 seems to be reduced compared to Cdk8, even though general activities

toward RNA pol II were of similar strength as demonstrated by gel shift and radioactive kinase assays.

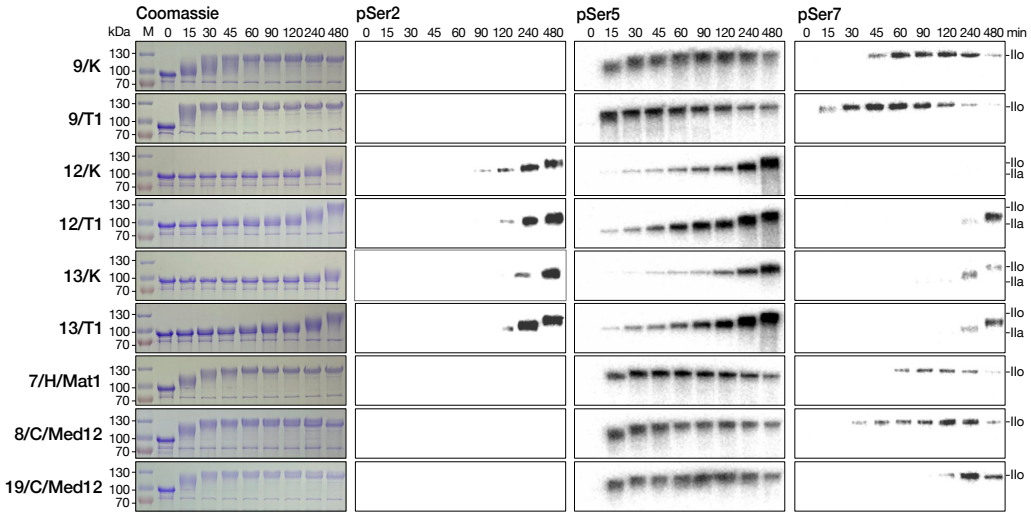

Fig.38 | Specificity of tCDK/Cyclin pairs toward pol II CTD heptad residues. In vitro kinase assays were performed in time-course series with 0.2 mM ATP, 10 µM GST-CTD, and 0.2 µM kinase for indicated time points and visualized in SDS gels by Coomassie staining (left) or by immunoblotting using RNA pol II CTD specific phosphoantibodies against pSer2, pSer5, or pSer7 (right). Upon phosphorylation, pol II CTD bands shift from the hypo- (IIa) to the hyper-phosphorylated (IIo) form. Imaging of blots stained with the same antibody were performed simultaneously.

tCDK specificity toward pre-phosphorylated RNA pol II CTD

During the transcription cycle, RNA pol II CTD is phosphorylated on the different heptad residues depending on the transcription phase. Therefore, CDKs of the later phases of the transcription cycle probably do not encounter unphosphorylated pol II CTD. In recent studies it has been shown that pre-phosphorylation of a specific residue at Tyr1 or Ser7 primes RNA pol II CTD recognition by Cdk9, Cdk12 and Cdk13[42,43,156,157]. Radioactive kinase assays were performed with the different CDK/Cyclin complexes and a set of synthetic CTD peptides, each peptide comprising three heptad repeats followed by a polyethylene linker and two arginines (Fig.39). CTD peptides were either unphosphorylated (consensus) or uniformly pre-phosphorylated at position Tyr1, Ser2, Thr4, Ser5, Ser7, or with lysines at position 7 (K7), the most frequent alteration in the human CTD consensus repeats[9]. Interestingly, the initiation CDKs, Cdk7, Cdk8 and Cdk19 show a clear preference for consensus CTD and Lys7, whereas the elongation CDKs, Cdk9, Cdk12, and Cdk13 prefer CTD peptides with a pre-phosphorylation at Ser7. Additionally, Cdk9 shows robust activity toward consensus and Lys7 CTD peptides, similar to initiation CDKs. Cdk12 and Cdk13 share the exact same phospho-residue preference. Again, the regulators Mat1 and Med12 amplify the catalytic activities of Cdk7, Cdk8, and Cdk19, respectively, whereas CycT1 complexes have overall higher activities in comparison to CycK.
Overall, these data show that the initiation CDKs share similarities in terms of substrate preferences and CTD residue specificity of pol II CTD, as do the elongation CDKs, Cdk12 and Cdk13, whereas the pause-release kinase Cdk9 shares features of both groups. Moreover,

these data indicate that substrate recognition is dominated by the Cdk rather than the bound Cyclin. However, influences of the C-termini of CycK and CycT1 adjacent to these Cyclin box-only constructs cannot be excluded and the importance of the CycT1 His-domain for interaction with DNA[173] and localization to liquid-liquid phase separated droplets has been demonstrated in recent studies[198].

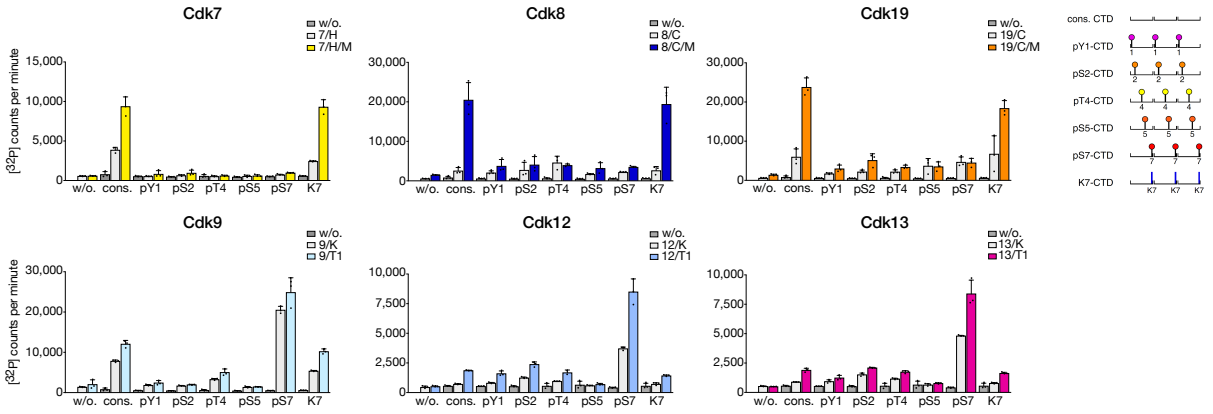

Fig.39 | tCDK/Cyclin specificity toward pre-phosphorylated pol II CTD residues. In vitro kinase assays were performed with 1 mM [^{32}P]-γ-ATP, 0.5 µM kinase and 150 µM pre-phosphorylated CTD Peptides for 60 min. Each peptide contained three consensus hepta-repeats with either no modification (cons.) or with phosphorylation marks continuously set at one residue of the heptad sequence as depicted in the cartoon. Peptides carried either phosphorylation marks for Tyr1, Ser2, Thr4, Ser5, Ser7 or a Lys7. All data are presented as mean ± SD of triplicates.

3.2.7 Crystal structure of the novel Cdk12/CycT1 complex

To better understand why initiation tCDKs can only be activated by their cognate Cyclin, while elongation tCDKs are promiscuous in their Cyclin interaction, the novel Cdk12/CycT1 complex was crystallized together with the small-molecule inhibitor NVP2 and the structure was determined to 2.0 Å resolution (Fig.40). The kinase complex was refined to an R_{work} of 0.23 and R_{free} of 0.26 with excellent stereochemistry (for crystallographic data collection and refinement statistics see Appendix A, Tab.6). An overlay with the crystal structure of cognate Cdk12/CycK in complex with NVP2 (crystalized by Dr. Kanchan Anand), which was determined to a resolution of 2.8 Å and an R_{work} of 0.24 and R_{free} of 0.30, and alignment of both Cdk12 N-lobes reveals a similar overall fold between the two Cdk12/Cyclin pairs, consistent with structures of other tCDK/Cyclins[62]. Both Cdk12 structures exhibit a typical kinase fold consisting of an N- and C-terminal lobe. Due to Cyclin binding, Cdk12 is in the `αC-in´ position promoting an open conformation of the kinase active site. The phosphorylated threonine pT893 within the T-loop of the activation segment is clearly visible in both structures and coordinates the three Cdk12 arginines R858, R882, and R773 of the PITAIRE helix, with R773 being further stabilized by electrostatic interactions of CycT1 E96 or CycK E108 of the conserved KΦEEΦ motif (where Φ is a hydrophobic residue) at the end of the H3 helix, ultimately converging in an activated kinase conformation. Additionally, Cdk12 harbors an extension helix followed by a polybasic cluster which contributes to its catalytic activity, as described before[42].

Most cCyclins share an extended interface with their cognate CDK partners contacting both, the CDK N- and C-lobe, through residues from the Cyclin helices H2 to H5 from the N-terminal Cyclin box and H1´ at the start of the C-terminal Cyclin box[200–203]. In contrast, tCyclins contact their CDKs exclusively with their N-terminal Cyclin box, whereas the C-terminal Cyclin box is not involved in the interface[65]. This is also resembled in both Cdk12 structures, in which the main binding interface involves residues from the αC-helix of the Cdk12 N-lobe with the PITAIRE motif and the H3 and H5 helices of both Cyclins of the N-terminal Cyclin box. Strikingly, the structures of CycK and CycT1 bound to Cdk12 superimpose very well with an RMSD of 0.795 Å with a stronger superimposition of the N-terminal Cyclin boxes and a stronger divergence in the C-terminal Cyclin boxes. This can be explained by the fact that CycT1 contains an N-terminal Cyclin box only, but instead of a typical C-terminal Cyclin box a `Cyclin-like´ domain instead[65]. That CycK is most similar with CycT1 not only functionally but also structurally, has been described extensively over the last decade[65,174]. CycK and CycT1 share a striking sequence similarity from residue 1 to 250 composing their Cyclin boxes, while their C-termini harbor largely unstructured regions of higher variability. The secondary structure, folding and relative orientation of the Cyclin boxes of CycK are virtually identical to CycT1, and the two molecules superimpose with an RMSD of 0.44 Å (without a bound CDK partner)[174].

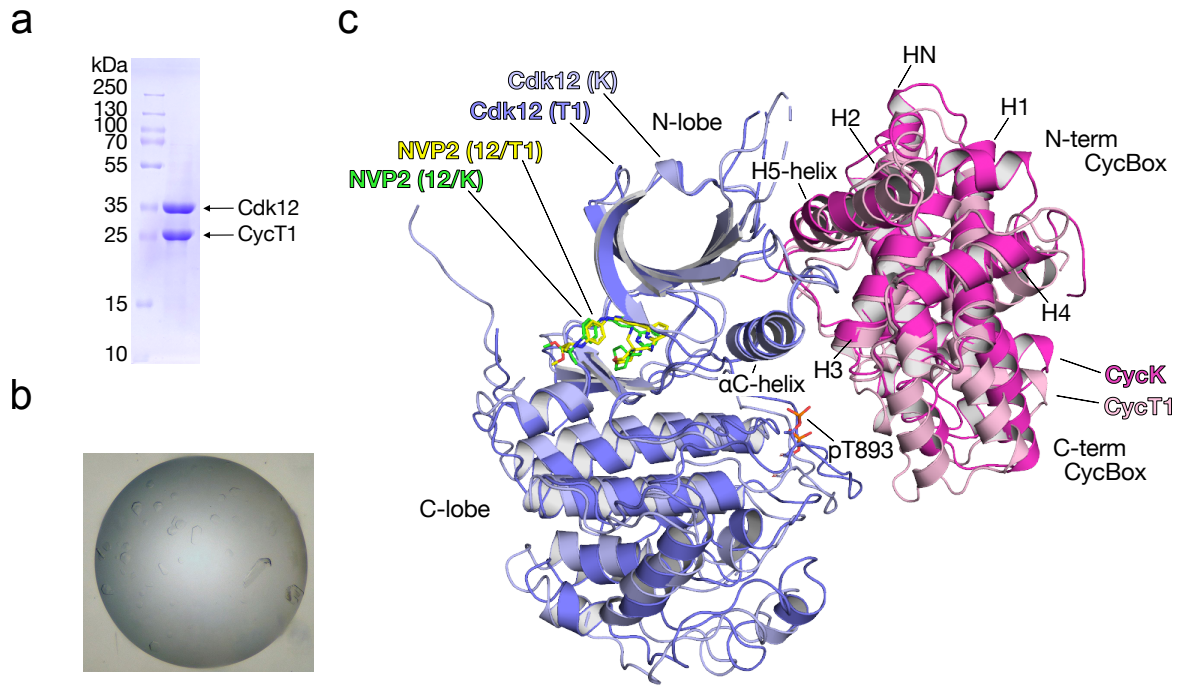

Fig.40 | Crystal structure of the novel Cdk12/CycT1 complex. a Human Cdk12 (residue 714-1063) together with CycT1 (residue 1-272) were purified to homogeneity in a 1:1 stoichiometry. **b** Optimized crystals of Cdk12/CycT1 with NVP2. **c** Overlay of the crystal structures of the novel Cdk12/CycT1 and the cognate Cdk12/CycK in complex with NVP2 at 2.0 Å and 2.8 Å resolution, respectively. The main interactions are made by the Cdk12 αC-helix and the Cyclin H3- and H5-helices. The phosphorylated threonine (pT893) is indicated. Helices are labelled for the N-terminal Cyclin box (CycBox).

When comparing the CDK/Cyclin interfaces of cognate Cdk12/CycK, Cdk9/CycT1 and the novel Cdk12/CycT1 pair in more detail (Fig.41), it becomes clear that the heart of the CDK/Cyclin interface is formed by a hydrophobic pocket, which is build up by two phenylalanine residues, one from the CDK (F767 in Cdk12, F59 in Cdk9) and one from the Cyclin (F153 in CycK, and F146 in CycT1). Analysis of the tCDK/Cyclin interfaces of CycC, H, K, and T1 with their cognate CDK partner and the novel Cdk12/CycT1 pair, reveal three main interfaces of tCDK/Cyclin interaction, in the N-terminal part of the Cyclin box, in the H3-helix, and in the H5-helix (Fig.41b). Comparison of the Cdk1/CycB and Cdk2/CycB structures, as well as of structures of Cdk2/CycA, Cdk2/CycB and Cdk2/CycE, suggests that most influence on the CDK/Cyclin interface results from the identity of the CDK, not the Cyclin[65,200–203]. Similarly, whereas residues from the H3-helix are highly similar between the different Cyclins, including the conserved KΦEEΦ motif, the interacting P**I**TA**I**RE-helix (in Cdk12, P**I**TA**L**RE in Cdk9, N**R**TA**L**RE in Cdk7, and S**M**SA**C**RE in Cdk8) show higher sequence variability, probably contributing to the specificity of the interface. Especially residues in the second and fifth position of the helix vary significantly, demonstrating a high similarity between Cdk9 and Cdk12/13, which might explain, why Cdk9, Cdk12, and Cdk13 are able to be activated by CycK and CycT1, but not by other tCyclin. And, on the contrary, why Cdk7 and Cdk8 can only bind to their cognate partner, CycH or CycC, respectively.

On the other hand, the sequence variability in tCyclins is most prominent in the N-terminal binding region and H5-helix. Interactions made by residues from the Cyclin N-terminal binding regions are highly diverse and seem to be important for the specificity of the tCDK/Cyclin interaction, as described before[65]. Additionally, key residues in the H5-helix are less well conserved. Eventhough from our data it seems that phosphorylation of CycT1 T149 is irrelevant for the interaction with Cdk9 (Fig.27), this residue appears to play a key role in the CDK interaction in its unphosphorylated state, similar to CycK Q156 at this position. In CycC and CycH, however, this residue is a hydrophobic isoleucine (CycC I151, CycH I159), which does not participate in the binding interface with the CDK partner.

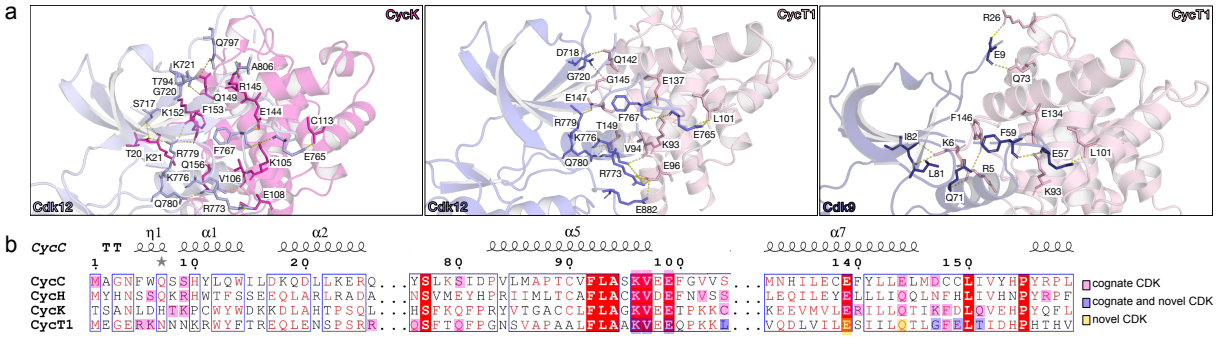

Fig.41 | The interfaces of Cdk12/CycK, Cdk12/CycT1 and Cdk9/CycT1 are highly similar. a Close-up of the CDK/Cyclin interface. Residues are indicated within a distance of smaller than 3.5 Å. **b** Sequence alignment of the CycC, H, K and T1 binding interfaces with their cognate or novel Cdk partners. Residues of the interface within a distance of smaller than 3.5 Å are highlighted. Residues conserved in all four Cyclins are boxed red, and those that are similar have red characters. The sequence alignment was prepared with MultAlin. The secondary structure alignment was prepared with ESPript and indicates secondary structural elements of CycC. PDB codes: Cdk7/CycH (6XBZ), Cdk8/CycC (6R3S), Cdk9/CycT1 (3BLH).

3.2.8 Chemical Inhibition of tCDK/Cyclin pairs

Inhibition of tCDK/Cyclin pairs by small-molecular compounds

Due to transcriptional addiction of many cancer types, tCDKs gained more and more attention as potential drug targets during the last decade as they are direct regulators of the transcription machinery, while, at the same time, kinase inhibitors are a key drug class in the oncology field. To investigate whether different Cyclins influence the inhibition of tCDKs by small-molecular inhibitors, surface plasmon resonance (SPR) measurements were performed with the small-molecular CDK inhibitor NVP2 (Fig.42). All six CDK/Cyclin pairs were immobilized on a CM5 chip by amine-coupling and flushed with increasing concentrations of NVP2 (0.02, 0.08, 0.313, 1.25, 5, 20, 80, 320 nM) for single-cycle kinetics. NVP2 was originally developed for the specific inhibition of Cdk9[129]. However, surface plasmon resonance measurements demonstrate dissociation constants (K_D) of NVP2 in the picomolar range for Cdk9 and in the low nanomolar range for Cdk12 and Cdk13. Overall, the different Cyclins do not show any influence on NVP2 interactions.

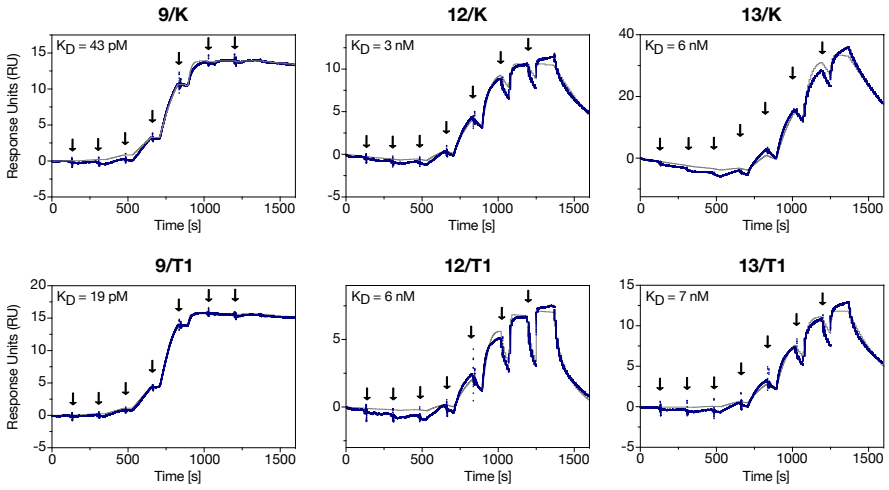

Fig.42 | Binding affinity measurements of NVP2 with novel tCDK/Cyc pairs. Surface plasmon resonance (SPR) was performed using single-cycle kinetics. Kinases were immobilized to the chip surface by amine coupling. NVP2 was injected at increasing concentrations of 0.02, 0.08, 0.313, 1.25, 5, 20, 80, 320 nM for 120 s followed by dissociation for 600 s. Dissociation constants (K_D) were determined based on a 1:1 interaction model.

Additionally, kinase inhibition was determined by radioactive kinase assays (Fig.43). Concentration series of NVP2 ranging from 3.1 nM to 1 mM were incubated with 0.2 µM kinase and 0.2 mM [^{32}P]-γ-ATP concentrations and a sigmoidal fit was used to determine IC_{50} values from these dose-response measurements. In line with the estimated K_D values, Cdk9 (3 nM), Cdk12 (20-45 nM), and Cdk13 (144-202 nM) were inhibited by NVP2 to the same extent, independent of the respective Cyclin. Higher activities of CycT1 complexes may be responsible for slight differences between Cdk12 and Cdk13/Cyclin pairs, because, as generally observed for this assay, lower kinase activities lead to increased background levels, which influences signal outcome.

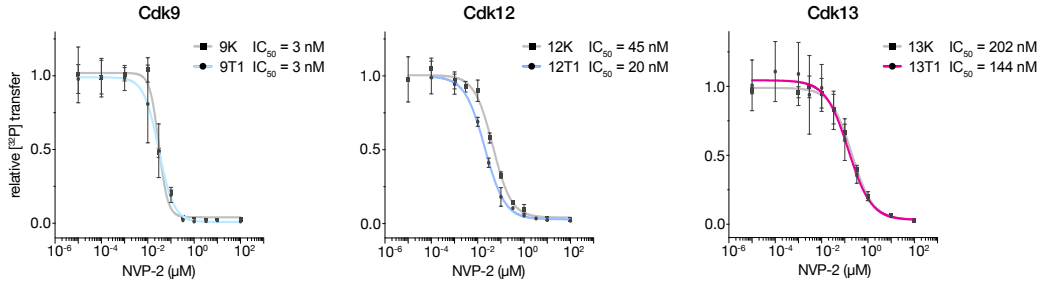

Fig.43 | Similar inhibition of tCDK/Cyclin pairs by NVP2. In vitro kinase assays were performed with concentration series of NVP2. The compound was pre-incubated for 5 min with 0.2 µM kinase and 0.2 mM [^{32P}]-γ-ATP followed by addition of 10 µM c-Myc and incubation for 15 min. All data are presented as mean ± SD of triplicates. A sigmoidal fit was used to calculate IC_{50} values (these measurements were performed by Max Schmitz).

Moreover, NVP2 was co-crystalized with Cdk12/CycK and determined at a resolution of 2.8 Å and Cdk12/CycT1 at 2.0 Å resolution, respectively, to analyze the influence of both Cyclins on the Cdk12 catalytic cleft and interactions with NVP2 (see 3.2.7, Fig.40 and 44). Overall, Cdk12 residues interacting with the inhibitor are highly similar between both structures with some slight deviations. NVP2 interacts intensively with the N-terminal lobe and the hinge region of Cdk12 via its central 2-aminopyrimidine group. The side chain of Cdk12 M816 forms hydrogen bonds with the aminopyrimidine pyridine unit. Also the gatekeeper residue F813 is involved in compound binding. While in Cdk12/CycK NVP2 forms an additional hydrogen bond with the backbone of E735 from the N-lobe, in Cdk12/CycT1 an additional hydrogen bond is formed with the backbone and side chain of D819 from the C-lobe. Whether these differences are a result the interacting Cyclin partner or the influence of crystal packing is not clear.

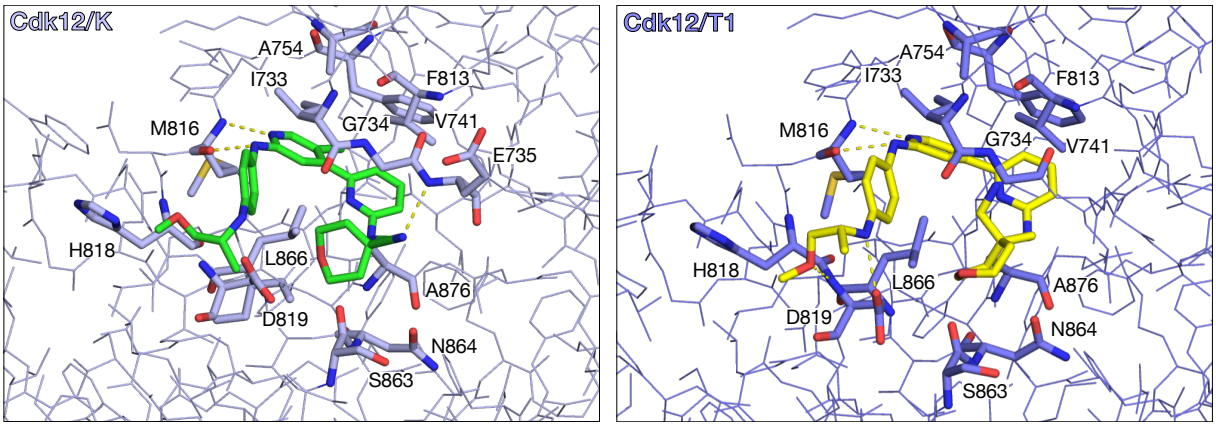

Fig.44 | Crystal structures of Cdk12/CycK and Cdk12/CycT1 in complex with NVP2. Molecular interactions of Cdk12/CycK (light blue; left) with NVP2 (green) are highly similar to Cdk12/CycT1 (dark blue; right) with NVP2 (yellow). Key residues are labeled and hydrogen bonds are indicated as dotted lines (The Cdk12/CycK crystal structure in complex with NVP2 was determined by Dr. Kanchan Anand with the help of Max Schmitz).

Selective degradation of tCDK/Cyclin pairs by TPDs

As an alternative to ATP-competitive small-molecule compounds, targeted protein degraders (TPDs) receive increasing attention as they offer some interesting advantages over classical compounds such as increased target selectivity[113]. Recently, two molecular glue degraders, the CDK inhibitor CR8[192] and dCEMM2[191], have been shown to recruit the Cullin4 adaptor protein DNA damage-binding protein 1 (DDB1) to CDK12/13–CycK, where they induce selective degradation of CycK (Fig.45a). Similarly, the PROTAC THAL-SNS-032 was developed for selective degradation of Cdk9[129]. However, the degradation of CycT1 by CR8 and dCEMM2 via interaction with Cdk12/13 or degradation of the Cyclin partner of Cdk9 was not investigated in the respective studies. Therefore, HEK293T cells were treated with indicated concentrations of TPDs or DMSO for 0.5, 1, 2, 3, and 4 hours, and whole cell extracts were analyzed by immunoblots (Fig.45b). Interestingly, all three TPDs seem to be selective for their respective target and do not induce the degradation of CycT1, either in complex with Cdk13 for CR8 and dCEMM2, nor for THAL-SNS-032, which further confirms that the degradation of the target is highly depend on the lysine residues exposed on the protein surface that are ubiquitinated by the E3 ligase. These data indicate that TPDs are valuable tools to selectively degrade components of novel and cognate CDK/Cyclin complexes.

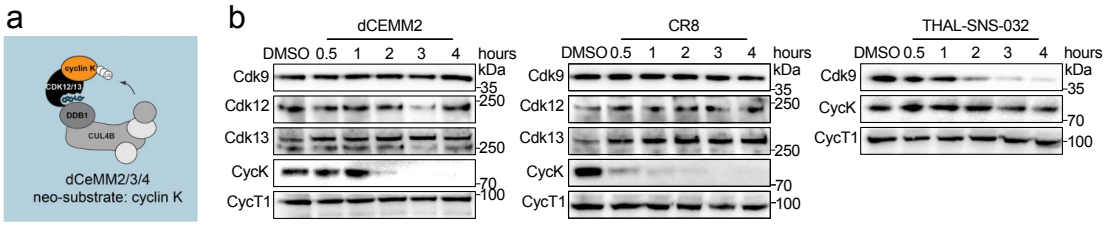

Fig.45 | Targeted protein degraders are selectively targeting components of CDK/Cyclin complexes. **a** Scheme of the molecular glue degrader dCEMM2/3/4 inducing proximity between the DDB1/CUL4 E3 ligase complex and Cdk12/13, resulting in proteasomal degradation of CycK. **b** The targeted protein degraders (TPDs) 2.5 µM dCEMM2, 1 µM CR8, and 0.25 µM THAL-SNS-032 induce degradation specifically to only one Cdk or Cyclin. HEK293T cells were treated with indicated concentrations of TPDs. Cells were lysed after 0.5, 1, 2, 3 and 4 hours and compared to a DMSO control in immunoblots.

4. CONCLUSIONS

4.1 Abemaciclib is a potent inhibitor of HIPKs and DYRK1A

HIPKs and DYRKs are dual-specific kinases, since they auto-phosphorylate a conserved tyrosine in a transient intermediate state during maturation at the ribosome, but phosphorylate substrates at serine/threonine residues[71]. Thus, members of the HIPK and DYRK family are always in a constitutive, at least partially active state. In the crystal structure determined in this thesis, HIPK3 adopts an active conformation with a fully accessible ATP-binding site although the kinase is in the nucleotide-free apo state (Fig.12). The conserved phospho-tyrosine coordinates an extensive salt-bridge and hydrogen-bond network that results in a stably ordered activation loop and the "αC-in" position, ultimately leading to an active kinase conformation (Fig.14). In contrast, the key residue that is crucial for CDK activation is a conserved threonine located in the kinase T-loop, which is phosphorylated by a CDK-activating kinase (CAK) in many CDKs. These differences in HIPK/DYRK and CDK activation are accompanied by a poor conservation of the arginine in the CDK HRD motif, which is an alanine in HIPKs (HAD motif) and a cysteine in DYRKs (HDC motif). In most CDKs, this HRD arginine together with an arginine of the αC-helix (mostly a REI motif) and an arginine of the DFGLAR motif, form salt bridges with the phosphorylated T-loop threonine, thereby coordinating three distant loops and inducing the conformational changes required for CDK activation[62]. In HIPKs and DYRKs instead, the phosphate group of the conserved T-loop tyrosine forms salt bridges with two arginines from a R(F/Y)YR motif, coordinating the kinase active conformation (Fig.12). This was observed for all structures of the DYRK family known to date, as well as for HIPK2 (Fig. 16). For HIPK3 in contrast, both structures determined in this study show that the phosphate moiety of the conserved tyrosine pY359 forms a salt bridge with R431 of the CMGC-specific insert region, which is stabilized by contacts with a symmetry related, neighboring molecule (Fig.15). Further studies are required to clarify whether the unconventional conformation of the HIPK3 tyrosine pY359 is a crystallographic artifact or a peculiarity of HIPK3. Nonetheless, the importance of the auto-phosphorylated tyrosine residue remains a specific feature for the activation of all members of the HIPK and DYRK families.

Another characteristic region of the HIPK and DYRK kinase branch is the CMGC insertion between the αH- and αG-helix that extends over the two helices αN and αM, the β-hairpin loop, and two additional helices, αL and αL' (Fig.12-13;16). Interestingly, HIPKs have the longest insert region of all CMGC kinases and the β-hairpin loop adopts different conformations in the known structures of HIPK2, HIPK3, DYRK1A, DYRK2, and DYRK3. Additionally, HIPK2 and DYRK3 contain one, and DYRK2 contains two phospho-residues in the insert region. Because DYRKs and HIPKs mature in a constitutive active state, further control mechanisms are necessary to specifically regulate kinase activity. Extensive post-translational modification (PTMs) such as phosphorylation, acetylation, ubiquitination, SUMOylation, and caspase cleavage have been observed for these kinase families[108]. Such PTMs have been shown to alter protein stability and subcellular localization, ultimately controlling kinase activity in HIPKs and DYRKs. Furthermore, the CMGC insert region has been shown to serve as a binding platform for such signaling partners[153], which eventually explains the diversity and specificity of the CMGC insert region in the HIPK and DYRK family.

HIPKs show several similarities with DYRK family members, not only in terms of structural features, but also in terms of function[204]. In this thesis, HIPKs have been shown to phosphorylate the general transcription factors SPT5 and c-Myc, as well as the Ser2 and Ser5 residues of RNA pol II CTD in vitro, indicating a role in the direct regulation of the transcription machinery (Fig.17-19). These results are reminiscent to the description of DYRK1A as a transcriptional regulator that acts as a pol II CTD kinase at gene loci, which are functionally associated with translation, RNA processing, and cell cycle control[89]. Early HIPK studies examined the localization of these kinases to nuclear speckles by a speckle-retention signal of low complexity at the C-terminus[103,104,205]. Interestingly, DYRK1A and P-TEFb both possess a low complexity histidine-rich domain, which promotes binding and hyperphosphorylation of pol II CTD in nuclear condensates, mediated by liquid-liquid phase separation (LLPS)[198]. Whether the localization of HIPKs to nuclear speckles is mediated by LLPS and whether this is crucial for pol II CTD hyperphosphorylation in vivo, awaits further investigation in future studies. Furthermore, a recent study demonstrated that DYRK1A activity toward pol II CTD is modulated by the WD40-repeat protein DCAF7 at myogenic loci[155]. Interestingly, DCAF7 is a known interactor of HIPK2 regulating its kinase function in stress responses[206,207] raising the question of whether this effect could be a consequence of direct phosphorylation of pol II and co-factors. Further studies are necessary to investigate whether HIPKs interact with the transcription machinery in vivo, and whether this is restricted to certain genes as it seems to be the case for DYRK1A.

Transcription-regulating kinases have evolved as an important drug target for cancer therapies in recent years[112]. Eventhough HIPKs were identified as tumor suppressors in many cancer types, some studies suggest an oncogenic role of HIPKs[111,141,208]. In addition, members of this kinase family are involved in the pathology of Huntington´s disease, Alzheimer´s disease, diabetes type II, and chronic fibrosis, suggesting the development of HIPK specific inhibitors as potentially valuable therapeutics[161,204]. With the determined crystal structures of apo HIPK3 as well as of HIPK3 and DYRK1A bound to abemaciclib, the structural basis for a rational inhibitor design of this human kinase family was established.
The crystal structures of HIPK3 and DYRK1A in complex with abemaciclib, binding analysis and IC_{50} values demonstrate that HIPK2, HIPK3 and DYRK1A are potent targets of this FDA-approved drug (Fig.20-24). Because ATP-competitive inhibitors target a structurally and functionally similar binding site, poly-pharmacology is a common phenomenon, and such target promiscuity can have both beneficial and detrimental therapeutic consequences[209]. Abemaciclib is a third generation CDK-directed drug with an impressive clinical performance in the treatment of HR+ and HER2- metastatic breast cancer[164-166] and reported to be a selective Cdk4/6 inhibitor[163]. Nonetheless, it has been suggested early on that abemaciclib may have pan-selective kinase inhibition properties that exceed those of related Cdk4/6-targeting drugs such as ribociclib and palbociclib[166,210]. In addition, transcriptional profiling in seven breast cancer cell lines revealed a profile of downregulated genes for abemaciclib beyond the characteristic signature of Cdk4/6 activity[210]. Whether abemaciclib´s therapeutic success benefits from the simultaneous inhibition of the cell cycle regulating kinases Cdk4/6 in combination with off-target effects from the transcription kinases Cdk9, DYRK1A, and HIPK2/3 awaits a detailed examination in future studies.

Moreover, DYRK1A itself is a promising therapeutic target, as dysregulation of DYRK1A activity has been associated with cancer, diabetes type II, Down syndrome, Intellectual Developmental Disorder Autosomal Dominant 7 (MRD7), Fronto-Temporal Degeneration (FTD), dementia, Parkinson´s, Huntington´s and Alzheimer´s disease (AD)[84,211]. In this context, it is remarkable that an FDA-approved drug inhibits DYRK1A to the same extent as Cdk4/6, suggesting that repurposing this approved drug may be a novel therapeutic strategy for the treatment of diseases with DYRK1A involvement. In line with our results, a recent study found that abemaciclib is a potent inhibitor of DYRK1A, DYRK1B, CLK1, CLK4, and GSK3, as demonstrated by radiometric assays determining IC_{50} values and FRET proximity assays[212]. Furthermore, a recent study investigated the effects of abemaciclib on cognitive function and Amyloid β (Aβ)/tau pathology in mouse models of AD[213]. Cdk4 mRNA levels are elevated in the parietal lobe in AD patient post-mortem brains, as is Cdk4 immunoreactivity in the hippocampus and Cdk4/6 are responsible for cell cycle re-entry by regulating G0 to G1 transition, which is markedly increased in the neurons of AD patients. Recent studies revealed that Cdk4 inhibition protects against the effect of Aβ oligomer treatment in rat primary neurons. Thus, the authors investigated whether Cdk4/6 inhibition by abemaciclib might be of therapeutic value for AD pathologies. Indeed, the authors found that abemaciclib treatment resulted in improved memory impairments in Aβ-overexpressing mouse models, as well as inhibition of Aβ accumulation by enhancing the activity and protein levels of Aβ-degrading enzymes, and suppression of tau phosphorylation in tau-overexpressing PS10 mice. Remarkably, the authors point out that these beneficial effects are not exclusively caused by Cdk4/6 inhibition, but are mediated through DYRK1A/STAT3 signalling through the reduction of DYRK1A and/or GSK3β activity[213]. Taken together, these results support the idea that the anticancer drug abemaciclib might be an effective therapy as a pan-selective kinase inhibitor for AD pathologies.

4.2 The promiscuity of elongation tCDKs

According to the current dogma of CDK biology, cCDKs can be activated promiscuously by multiple cCyclins, whereas tCDKs are activated by only one specific tCyclin partner[54]. Our re-evaluation of human tCDK/Cyclin pairs in in vitro experiments revealed not only active kinase complex formation between the cognate Cdk/Cyclin pairs Cdk7/CycH, Cdk8/CycC, Cdk9/CycT1, Cdk12/CycK and Cdk13/CycK, but also between the novel tCDK/Cyclin pairs Cdk9/CycK, Cdk12/CycT1 and Cdk13/CycT1 (Fig.25-26). Analysis of the amino acid sequence identities of tCDK kinase domains and tCyclin boxes (Fig. 25) suggests two groups of tCDKs and tCyclins that coincide not only with their closest relatives but also with functions in the transcription cycle: the initiation tCDKs, Cdk7/CycH and Cdk8/CycC, which indeed interact only with their cognate partner, and the elongation tCDKs, Cdk9, Cdk12, and Cdk13 that can be activated by both Cyclins, CycK and CycT1.

Sequence alignments and analysis of the crystal structure of novel Cdk12/CycT1 in comparison to Cdk12/CycK and other cognate tCDK/Cyclin pairs, reveal an identical overall fold of CycK and CycT1 when bound to Cdk12 (Fig.40-41). The main binding interface involves key residues from the αC-helix of the Cdk12 N-lobe with the PITAIRE motif and highly similar residues of the H3 and H5 helices of the N-terminal Cyclin box of both Cyclins. On the other hand, specificities of CDK/Cyclin interfaces are gained by a high sequence variability in

the Cyclin´s N-terminal region and H5-helix of the first Cyclin box, as well as in the CDK´s P**I**TA**I**RE-helix (in Cdk12, P**I**TA**L**RE in Cdk9, N**R**TA**L**RE in Cdk7, and S**M**SA**C**RE in Cdk8). Especially residues in the second and fifth position of the helix vary significantly, demonstrating a high similarity between Cdk9 and Cdk12/13, in contrast to Cdk7 and Cdk8. These observations might explain, why Cdk9, Cdk12, and Cdk13 are able to be activated by CycK and CycT1, but not by other tCyclins, and, in contrast, why Cdk7 and Cdk8 can only bind their cognate partner, CycH or CycC, respectively.

In co-IP experiments from HEK293T and A375 cells, we demonstrated that both Cdk13 pairs, Cdk13/CycK and Cdk13/CycT1, co-exist in cells (Fig.29). With the help of Jennifer Devlin and Ricky Johnstone from the Peter MacCallum Cancer Centre in Melbourne, Australia, we performed global ChIP-seq analysis, which revealed co-localization of Cdk13 and CycT1 at selected genes, whereas Cdk9 and pol II were not present at these sites (Fig.36). Furthermore, in a study with Megan Insco from the Dana-Farber Cancer Institute, Harvard Medical School and Leonard Zon from Boston Children´s Hospital, Boston, USA, we discovered that Cdk13/CycT1 plays a role in nuclear RNA surveillance in melanoma patients, a zebrafish model and A375 cells[178,179] (Fig. 28). Altogether our data provide substantial evidence that the novel Cdk13/CycT1 pair exists not only in vitro, but also in vivo. However, it remains unclear whether there exists a Cdk9/CycK or a Cdk12/CycT1 complex in cells. It is noteworthy that Cdk9/CycK was already identified in early studies, but predominantly in vitro, and in the few in vivo experiments only under technical weaknesses, which should therefore be taken with caution[133,170-177]. On the one hand, the reasons that Cdk9/CycK and Cdk12/CycT1 are not detected in HEK293T and A375 cells could be simple, such as that they do not exist in these cell lines but are present in others, or that they may have functions in differentiated tissues rather than in proliferating cells. This would also explain why previous studies have not yet identified these novel CDK/Cyclin pairs, since all experiments were performed in proliferating cells[133,170-177]. But on the other hand, also more complex explanations seem plausible, demonstrating how little is still known about this prominent kinase family and that a range of further investigations in a cellular context are necessary. Regarding complex formation, it is not clear yet what determines whether Cdk13 binds to CycK, or CycT1 instead. It seems possible that additional proteins or complexes may participate in the CDK/Cyclin interaction, such as it has long been known for the cCDK/Cyclin regulatory CIP/KIP proteins[62-66]. In this case, expression levels of such adaptor proteins or complex components could determine CDK/Cyclin binding. In contrast to Cdk12 and Cdk13, Cdk9 regulation has been intensively studied and 50-90% of Cdk9/CycT1 is sequestered by Hexim1 in the 7SK snRNP complex[35-37,67,68]. However, from in vitro kinase assays, it seems that Hexim1 interacts with CycT1, but not necessarily with CycK (Fig.31). Thus, it might be possible that the interaction of Hexim1 with Cdk9 prevents complex formation with CycK. This is further supported by a recent study demonstrating that loss of hnRNP R, a known 7SK interactor, promotes the release of P-TEFb from 7SK in HeLa cells, which leads to Cdk9 assembly with CycK[177]. Unfortunately, the study has some technical weaknesses and further experiments with functional Knock-outs of 7SK components, or treatment with UV light, actinomycin D, or CDK inhibitors stimulating release of P-TEFb from the 7SK complex[64] could help to confirm the existence of a Cdk9/CycK complex in vivo. Early studies showed that Cdk9 can be activated by CycT1, CycT2 or CycK[171]. In this thesis, the role of CycT2 in novel tCDK/cyclin complex formation was not investigated. However, CycT1 Knock-out experiments in HEK293T cells combined with

selective Cdk9 degradation revealed that Cdk9 activity is maintained despite the loss of CycT1 (Fig.30). Future experiments are required to unravel which Cyclin compensates for the loss of CycT1 to maintain Cdk9 activity, CycK or CycT2. Furthermore, little is known about CycT2, and the question arises what functions a Cdk9/CycK or Cdk9/CycT2 complex might have, other than CycT1 compensation.

Natural Cyclin degradation by the proteasome is a well-studied phenomenon and provides the basis for promiscuous cCDK regulation throughout the different phases of the cell cycle. Likewise, the promiscuity of tCDK/Cyclin pairs now observed raises the question, whether tCDK/Cyclin complexes might also be regulated via proteasomal degradation, maybe independent from the cell cycle, but due to other cellular stimuli such as cell stress, global transcriptional changes or under certain gene expression programs. Interestingly, our in silico analysis of tCyclin amino acid sequences revealed a similar number of degron motifs for identical E3 ligases of the APC/C and SCF families as for cCyclins (Fig. 32). Although constant tCyclin levels are a consent in the field, a few experimental studies already demonstrated changes in tCyclin levels and proteasomal degradation under different cellular conditions, such as T-cell maturation and cell stress[186-190]. On the other hand, proteasomal degradation of tCyclins already gained a lot of attention, but only in the light of chemical targeting by targeted protein degraders (TPDs)[112–118]. Our in silico predictions suggest DDB1 as an adaptor protein targeting CycK for ubiquitination by a Cullin4 E3 ligase. Interestingly, in five recent studies molecular glue degraders were identified targeting CycK for degradation by a Cullin4 E3 ligase, mediated by DDB1[191–195]. That all five independent studies discover compounds targeting CycK via DDB1 for proteasomal degradation using rather serendipitous screening approaches might indicate that CycK is a natural target for Cullin4 E3 ligases and furthermore, that such screens are biased for natural E3 ligase targets. Also, time course experiments in HEK293T cells with the proteasomal inhibitor bortezomib confirm that CycK is proteasomally degraded, in contrast to CycT1 (Fig.33). Strikingly, all five molecular glue degraders target Cdk12/13, but only induce degradation of CycK, not of the CDK partner. This raises the question, whether under normal cellular conditions a putative Cdk9/CycK complex might be under constant proteasomal degradation of CycK, but not Cdk9, and might have specific functions in cell stress, as it has been shown for CycK in DDR[133] as well as for Cdk9[214–218], and in some studies also as a Cdk9/CycK complex[172]. Thinking a step further, natural E3 ligase targets might be easier targets for the development of TPDs. Interestingly, it has previously been shown that CycC is degraded by the proteasome upon various cell stress stimuli[186-188] and in silico predictions indicate von Hippel-Lindau (VHL) is a natural E3 ligase of CycC (data not shown). As VHL is a well-studied E3 ligase in the TPD field with various small-molecular ligands binding VHL for PROTAC design[219], CycC might a promising target for such VHL-based PROTACs. Also, CycT1 protein levels change during T-cell maturation[183-184] and, according to our data, also upon THP1 differentiation into macrophage-like cells (Fig.34). However, whether CycT1 expression is simply upregulated, or whether CycT1 is also constantly degraded by the proteasome, as indicated by different studies[169,185], in undifferentiated THP1 cells requires further investigation. Overall, however, it is not clear whether the degradation of tCyclins leads to promiscuous complex formation with a different Cyclin partner, such as one of the novel Cdk/Cyclin pairs, or is simply a regulatory mechanism for CDK inactivation.

Although our data provide substantial evidence that a Cdk13/CycT1 complex exists not only in vitro but also in vivo, the functional consequences of the observed Cyclin promiscuity for tCDKs remain enigmatic and further investigation is required to see whether Cdk13/CycT1 has biological functions beyond its role in melanoma[178,179]. From in vitro kinase assays, it seems that CycK and CycT1 pairs show no striking differences in substrate preference (Fig.37-39). The only exception is phosphorylation of c-Myc, which is elevated for Cdk12 and Cdk13/CycK compared to CycT1 pairs. It has long been known that Cdk9 is recruited by c-Myc to promoter regions, resulting in global amplification of existing gene expression programs[38,119]. In contrast, nothing is known about the interactions of CycK, Cdk12, and Cdk13 with c-Myc, but since c-Myc is a proto-oncogene, which is dysregulated in 50% of all tumors[4,119], it offers an interesting starting point for future experiments. Furthermore, ChIP-seq analysis was performed to see whether the two Cdk13 pairs might be regulating different sets of genes and whether there are differences to Cdk9 and Cdk12/Cyclin pairs. At many genes, pol II does not co-localize with Cdk13/CycT1. Thus, it might be possible that Cdk13/CycT1 exerts other functions than direct pol II regulation, as f.e. observed in the study from Insco et al.[178], regulating the PAXT complex and co-transcriptional processes such as the clearance of prematurely terminated RNAs. Because Cdk12 and CycK antibodies were of poor quality, and to the point that this thesis was written, we were still waiting for further sequencing data from the Genomics Core facility in Melbourne, no conclusions can be drawn about functional differences between the two Cdk13 pairs yet. Especially, ChIP-seq of CycT1 followed by re-ChIP with Cdk9 and Cdk13 could give important insight into functional roles of the novel Cdk13/CycT1 complex. Despite a putative regulation of different gene sets by both Cdk13 pairs, or participation in different complexes regulating f.e. different co-transcriptional RNA processes, liquid-liquid phase separation (LLPS) is emerging as a new concept in the field of transcriptional regulation[10,220] and CycT1 and pol II CTD have been shown to participate in nuclear condensates through formation of liquid-liquid droplets[198,221]. Intrinsically unstructured regions are crucial for LLPS and CycT1 possesses such a low complexity domain rich in histidine residues, which promotes co-localization to nuclear condensates[198]. Similarly, CycK harbors an intrinsically unstructured proline-rich region at the C-terminus (data not shown). From the distinct character of histidine or proline residues, it might be possible that both Cyclins are able to drive LLPS of different types of nuclear condensates and might consequently direct their bound Cdk partner into different LLPS droplets, which differ in their protein content, and thus substrates. However, whether CycK is indeed able to phase-separate and participate in liquid-liquid droplet formation awaits further investigation.

Overall, the observation that elongation tCDKs can be activated promiscuously by different Cyclins challenges a long-standing dogma in CDK biology. However, we must admit that our observations of tCDK promiscuity raises more questions than we have been able to answer yet, but paves the way for exciting new directions in the field of transcriptional regulation for years to come.

5. REFERENCES

1. Roeder, R. G. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 783–791 (2019).
2. Lee, T. I. & Young, R. A. Transcriptional Regulation and Its Misregulation in Disease. *Cell* **152**, 1237–1251 (2013).
3. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional Addiction in Cancer. *Cell* **168**, 629–643 (2017).
4. Vervoort, S. J. *et al.* Targeting transcription cycles in cancer. *Nat. Rev. Cancer* **22**, 5–24 (2022).
5. Kornberg, R. D. Chromatin Structure: A Repeating Unit of Histones and DNA: Chromatin structure is based on a repeating unit of eight histone molecules and about 200 DNA base pairs. *Science* **184**, 868–871 (1974).
6. Schoenfelder, S. & Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
7. Luger, K. Crystal structure of the nucleosome core particle at 2.8 A° resolution. **389**, (1997).
8. Fuda, N. J., Ardehali, M. B. & Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461**, 186–192 (2009).
9. Eick, D. & Geyer, M. The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code. *Chem. Rev.* **113**, 8456–8490 (2013).
10. Harlen, K. M. & Churchman, L. S. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **18**, 263–273 (2017).
11. Chapman, R. D., Heidemann, M., Hintermair, C. & Eick, D. Molecular evolution of the RNA polymerase II CTD. *Trends Genet.* **24**, 289–296 (2008).
12. Schilbach, S., Wang, H., Dienemann, C. & Cramer, P. Yeast PIC-mediator structure with RNA polymerase II C-terminal domain. *Proc. Natl. Acad. Sci.* **120**, e2220542120 (2023).
13. Schier, A. C. & Taatjes, D. J. Everything at once: cryo-EM yields remarkable insights into human RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* **28**, 540–543 (2021).
14. Hsin, J.-P. & Manley, J. L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137 (2012).
15. Zaborowska, J., Egloff, S. & Murphy, S. The pol II CTD: new twists in the tail. *Nat. Struct. Mol. Biol.* **23**, 771–777 (2016).
16. Vos, S. M. *et al.* Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature* **560**, 607–612 (2018).
17. Krajewska, M. *et al.* CDK12 loss in cancer cells affects DNA damage response genes through premature cleavage and polyadenylation. *Nat. Commun.* **10**, 1757 (2019).
18. Farnung, L. & Vos, S. M. Assembly of RNA polymerase II transcription initiation complexes. *Curr. Opin. Struct. Biol.* **73**, 102335 (2022).
19. He, Y. *et al.* Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016).
20. Aibara, S., Schilbach, S. & Cramer, P. Structures of mammalian RNA polymerase II pre-initiation complexes. *Nature* **594**, 124–128 (2021).
21. Rengachari, S., Schilbach, S., Aibara, S., Dienemann, C. & Cramer, P. Structure of the human Mediator–RNA polymerase II pre-initiation complex. *Nature* **594**, 129–133 (2021).
22. Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
23. Li, Y.-C. *et al.* Structure and noncanonical Cdk8 activation mechanism within an Argonaute-containing Mediator kinase module. *Sci. Adv.* **7**, eabd4484 (2021).

24. Galbraith, M. D., Donner, A. J. & Espinosa, J. M. CDK8: A positive regulator of transcription. *Transcription* **1**, 4–12 (2010).
25. Poss, Z. C. *et al.* Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. *Cell Rep.* **15**, 436–450 (2016).
26. Greber, B. J. *et al.* The cryoelectron microscopy structure of the human CDK-activating kinase. *Proc. Natl. Acad. Sci.* **117**, 22849–22857 (2020).
27. Greber, B. J. *et al.* The cryo-electron microscopy structure of human transcription factor IIH. *Nature* **549**, 414–417 (2017).
28. Fisher, R. P. Secrets of a double agent: CDK7 in cell-cycle control and transcription. *J. Cell Sci.* **118**, 5171–5180 (2005).
29. Fisher, R. P. Cdk7: a kinase at the core of transcription and in the crosshairs of cancer drug discovery. *Transcription* **10**, 47–56 (2019).
30. Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 167–177 (2015).
31. Booth, G. T., Parua, P. K., Sansó, M., Fisher, R. P. & Lis, J. T. Cdk9 regulates a promoter-proximal checkpoint to modulate RNA polymerase II elongation rate in fission yeast. *Nat. Commun.* **9**, 543 (2018).
32. Larochelle, S. *et al.* Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. *Nat. Struct. Mol. Biol.* **19**, 1108–1115 (2012).
33. Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).
34. Vos, S. M., Farnung, L., Urlaub, H. & Cramer, P. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature* **560**, 601–606 (2018).
35. Egloff, S. CDK9 keeps RNA polymerase II on track. *Cell. Mol. Life Sci.* **78**, 5543–5567 (2021).
36. Fujinaga, K., Huang, F. & Peterlin, B. M. P-TEFb: The master regulator of transcription elongation. *Mol. Cell* **83**, 393–403 (2023).
37. C. Quaresma, A. J., Bugai, A. & Barboric, M. Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb. *Nucleic Acids Res.* **44**, 7527–7539 (2016).
38. Rahl, P. B. *et al.* c-Myc Regulates Transcriptional Pause Release. *Cell* **141**, 432–445 (2010).
39. Xu, Y. *et al.* Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex. *Nat. Commun.* **8**, 15741 (2017).
40. Bernecky, C., Plitzko, J. M. & Cramer, P. Structure of a transcribing RNA polymerase II–DSIF complex reveals a multidentate DNA–RNA clamp. *Nat. Struct. Mol. Biol.* **24**, 809–815 (2017).
41. Ehara, H. *et al.* Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* **357**, 921–924 (2017).
42. Bosken, C. A. *et al.* The structure and substrate specificity of human Cdk12/Cyclin K. *Nat. Commun.* 14 (2014).
43. Greifenberg, A. K. *et al.* Structural and Functional Analysis of the Cdk13/Cyclin K Complex. *Cell Rep.* **14**, 320–331 (2016).
44. Nojima, T., Dienstbier, M., Murphy, S., Proudfoot, N. J. & Dye, M. J. Definition of RNA Polymerase II CoTC Terminator Elements in the Human Genome. *Cell Rep.* **3**, 1080–1092 (2013).
45. Mayer, A. *et al.* CTD Tyrosine Phosphorylation Impairs Termination Factor Recruitment to RNA Polymerase II. *Science* **336**, 1723–1725 (2012).
46. Kuehner, J. N., Pearson, E. L. & Moore, C. Unravelling the means to an end: RNA polymerase II transcription termination. *Nat. Rev. Mol. Cell Biol.* **12**, 283–294 (2011).
47. Sansó, M. *et al.* P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes Dev.* **30**, 117–131 (2016).
48. Vervoort, S. J. *et al.* The PP2A-Integrator-CDK9 axis fine-tunes transcription and can be targeted therapeutically in cancer. *Cell* **184**, 3143-3162.e32 (2021).

49. Manning, G. The Protein Kinase Complement of the Human Genome. *Science* **298**, 1912–1934 (2002).
50. Kanev, G. K. *et al.* The Landscape of Atypical and Eukaryotic Protein Kinases. *Trends Pharmacol. Sci.* **40**, 818–832 (2019).
51. Attwood, M. M., Fabbro, D., Sokolov, A. V., Knapp, S. & Schiöth, H. B. Trends in kinase drug discovery: targets, indications and inhibitor design. *Nat. Rev. Drug Discov.* **20**, 839–861 (2021).
52. Endicott, J. A., Noble, M. E. M. & Johnson, L. N. The Structural Basis for Control of Eukaryotic Protein Kinases. *Annu. Rev. Biochem.* **81**, 587–613 (2012).
53. Adams, J. A. Kinetic and Catalytic Mechanisms of Protein Kinases. *Chem. Rev.* **101**, 2271–2290 (2001).
54. Malumbres, M. Cyclin-dependent kinases. *Genome Biol.* **15**, 6 (2014).
55. Cao, L. *et al.* Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. *BMC Evol. Biol.* **14**, 10 (2014).
56. Guo, Z. & Stiller, J. W. Comparative genomics of cyclin-dependent kinases suggest co-evolution of the RNAP II C-terminal domain and CTD-directed CDKs. *BMC Genomics* **5**, 69 (2004).
57. Brown, N. R., Noble, M. E. M., Endicott, J. A. & Johnson, L. N. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat. Cell Biol.* **1**, 438–443 (1999).
58. Johnson, J. L. *et al.* An atlas of substrate specificities for the human serine/threonine kinome. *Nature* **613**, 759–766 (2023).
59. Knighton, D. R. *et al.* Crystal Structure of the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase. *Science* **253**, 407–414 (1991).
60. De Bondt, H. L. *et al.* Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595–602 (1993).
61. Majumdar, A. *et al.* Allostery governs Cdk2 activation and differential recognition of CDK inhibitors. *Nat. Chem. Biol.* **17**, 456–464 (2021).
62. Wood, D. J. & Endicott, J. A. Structural insights into the functional diversity of the CDK–cyclin family. *Open Biol.* **8**, 180112 (2018).
63. Echalier, A., Endicott, J. A. & Noble, M. E. M. Recent developments in cyclin-dependent kinase biochemical and structural studies. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1804**, 511–519 (2010).
64. Russo, A. A., Jeffrey, P. D., Patten, A. K. & Massague, J. Crystal structure of the p27Kipl cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. **382**, (1996).
65. Tatum, N. J. & Endicott, J. A. Chatterboxes: the structural and functional diversity of cyclins. *Semin. Cell Dev. Biol.* **107**, 4–20 (2020).
66. Brown, N. R., Noble, M. E. M., Endicott, J. A. & Johnson, L. N. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat. Cell Biol.* **1**, 438–443 (1999).
67. Yik, J. H. N. *et al.* Inhibition of P-TEFb (CDK9/Cyclin T) Kinase and RNA Polymerase II Transcription by the Coordinated Actions of HEXIM1 and 7SK snRNA. *Mol. Cell* **12**, 971–982 (2003).
68. Nguyen, V. T., Kiss, T., Michels, A. A. & Bensaude, O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**, 322–325 (2001).
69. Han, J. *et al.* Deep Evolutionary Conservation of an Intramolecular Protein Kinase Activation Mechanism. *PLoS ONE* **7**, e29702 (2012).
70. Kentrup, H. *et al.* Dyrk, a Dual Specificity Protein Kinase with Unique Structural Features Whose Activity Is Dependent on Tyrosine Residues between Subdomains VII and VIII. *J. Biol. Chem.* **271**, 3488–3495 (1996).
71. Lochhead, P. A., Sibbet, G., Morrice, N. & Cleghon, V. Activation-Loop Autophosphorylation Is Mediated by a Novel Transitional Intermediate Form of DYRKs. *Cell* **121**, 925–936 (2005).

72. Soundararajan, M. *et al.* Structures of Down Syndrome Kinases, DYRKs, Reveal Mechanisms of Kinase Activation and Substrate Recognition. *Structure* **21**, 986–996 (2013).
73. Saul, V. V. *et al.* HIPK2 kinase activity depends on cis-autophosphorylation of its activation loop. *J. Mol. Cell Biol.* **5**, 27–38 (2013).
74. Himpel, S. *et al.* Specificity Determinants of Substrate Recognition by the Protein Kinase DYRK1A. *J. Biol. Chem.* **275**, 2431–2438 (2000).
75. Aranda, S., Laguna, A. & Luna, S. de la. DYRK family of protein kinases: evolutionary relationships, biochemical properties, and functional roles. *FASEB J.* **25**, 449–462 (2011).
76. Kim, E. J. *et al.* Dyrk1A Phosphorylates α-Synuclein and Enhances Intracellular Inclusion Formation. *J. Biol. Chem.* **281**, 33250–33257 (2006).
77. Becker, W. Emerging role of DYRK family protein kinases as regulators of protein stability in cell cycle control. *Cell Cycle* **11**, 3389–3394 (2012).
78. Chang, C. & Hsia, K. More than a zip code: global modulation of cellular function by nuclear localization signals. *FEBS J.* **288**, 5569–5585 (2021).
79. Kinstrie, R. *et al.* Characterization of a Domain That Transiently Converts Class 2 DYRKs into Intramolecular Tyrosine Kinases. *Sci. Signal.* **3**, (2010).
80. Rechsteiner, M. & Rogers, S. W. PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.* **21**, 267–271 (1996).
81. Lindberg, M. F. & Meijer, L. Dual-Specificity, Tyrosine Phosphorylation-Regulated Kinases (DYRKs) and cdc2-Like Kinases (CLKs) in Human Disease, an Overview. *Int. J. Mol. Sci.* **22**, 6047 (2021).
82. Becker, W. & Sippl, W. Activation, regulation, and inhibition of DYRK1A: Activation, regulation, and inhibition of DYRK1A. *FEBS J.* **278**, 246–256 (2011).
83. Arbones, M. L., Thomazeau, A., Nakano-Kobayashi, A., Hagiwara, M. & Delabar, J. M. DYRK1A and cognition: A lifelong relationship. *Pharmacol. Ther.* **194**, 199–221 (2019).
84. Deboever, E., Fistrovich, A., Hulme, C. & Dunckley, T. The Omnipresence of DYRK1A in Human Diseases. *Int. J. Mol. Sci.* **23**, 9355 (2022).
85. Guimera, J., Casas, C., Estivill, X. & Pritchard, M. HumanMinibrainHomologue (MNBH/DYRK1): Characterization, Alternative Splicing, Differential Tissue Expression, and Overexpression in Down Syndrome. *Genomics* **57**, 407–418 (1999).
86. Belgardt, B.-F. & Lammert, E. DYRK1A: A Promising Drug Target for Islet Transplant–Based Diabetes Therapies. *Diabetes* **65**, 1496–1498 (2016).
87. Martí, E. *et al.* Dyrk1A expression pattern supports specific roles of this kinase in the adult central nervous system. *Brain Res.* **964**, 250–263 (2003).
88. Fernández-Martínez, P., Zahonero, C. & Sánchez-Gómez, P. DYRK1A: the double-edged kinase as a protagonist in cell growth and tumorigenesis. *Mol. Cell. Oncol.* **2**, e970048 (2015).
89. Di Vona, C. *et al.* Chromatin-wide Profiling of DYRK1A Reveals a Role as a Gene-Specific RNA Polymerase II CTD Kinase. *Mol. Cell* **57**, 506–520 (2015).
90. Nguyen, T. L. *et al.* Correction of cognitive deficits in mouse models of Down syndrome by a pharmacological inhibitor of DYRK1A. *Dis. Model. Mech.* dmm.035634 (2018) doi:10.1242/dmm.035634.
91. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
92. Ferrer, I. *et al.* Constitutive Dyrk1A is abnormally expressed in Alzheimer disease, Down syndrome, Pick disease, and related transgenic models. *Neurobiol. Dis.* **20**, 392–400 (2005).
93. Branca, C. *et al.* Dyrk1 inhibition improves Alzheimer's disease-like pathology. *Aging Cell* **16**, 1146–1154 (2017).
94. Velazquez, R. *et al.* Chronic Dyrk1 Inhibition Delays the Onset of AD-Like Pathology in 3xTg-AD Mice. *Mol. Neurobiol.* **56**, 8364–8375 (2019).

95. García-Cerro, S., Rueda, N., Vidal, V., Lantigua, S. & Martínez-Cué, C. Normalizing the gene dosage of Dyrk1A in a mouse model of Down syndrome rescues several Alzheimer's disease phenotypes. *Neurobiol. Dis.* **106**, 76–88 (2017).
96. Melchior, B. *et al.* Tau pathology reduction with SM07883, a novel, potent, and selective oral DYRK1A inhibitor: A potential therapeutic for Alzheimer's disease. *Aging Cell* **18**, (2019).
97. Lee, Y. H., Im, E., Hyun, M., Park, J. & Chung, K. C. Protein phosphatase PPM1B inhibits DYRK1A kinase through dephosphorylation of pS258 and reduces toxic tau aggregation. *J. Biol. Chem.* **296**, 100245 (2021).
98. Souchet, B. *et al.* Inhibition of DYRK1A proteolysis modifies its kinase specificity and rescues Alzheimer phenotype in APP/PS1 mice. *Acta Neuropathol. Commun.* **7**, 46 (2019).
99. Kumar, K. *et al.* DYRK1A Inhibitors as Potential Therapeutics for β-Cell Regeneration for Diabetes. *J. Med. Chem.* **64**, 2901–2922 (2021).
100. Ackeifi, C. *et al.* Pharmacologic and genetic approaches define human pancreatic β cell mitogenic targets of DYRK1A inhibitors. *JCI Insight* **5**, e132594 (2020).
101. Kisaka, J. K., Ratner, L. & Kyei, G. B. The Dual-Specificity Kinase DYRK1A Modulates the Levels of Cyclin L2 To Control HIV Replication in Macrophages. *J. Virol.* **94**, e01583-19 (2020).
102. Booiman, T., Loukachov, V. V., Van Dort, K. A., Van 'T Wout, A. B. & Kootstra, N. A. DYRK1A Controls HIV-1 Replication at a Transcriptional Level in an NFAT Dependent Manner. *PLOS ONE* **10**, e0144229 (2015).
103. Kim, Y. H., Choi, C. Y., Lee, S.-J., Conti, M. A. & Kim, Y. Homeodomain-interacting Protein Kinases, a Novel Family of Co-repressors for Homeodomain Transcription Factors. *J. Biol. Chem.* **273**, 25875–25879 (1998).
104. Schmitz, M. L., Rodriguez-Gil, A. & Hornung, J. Integration of stress signals by homeodomain interacting protein kinases. 12.
105. Arai, S. *et al.* Novel homeodomain-interacting protein kinase family member, HIPK4, phosphorylates human p53 at serine 9. *FEBS Lett.* **581**, 5649–5657 (2007).
106. Ritter, O. & Schmitz, M. L. Differential intracellular localization and dynamic nucleocytoplasmic shuttling of homeodomain-interacting protein kinase family members. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1866**, 1676–1686 (2019).
107. He, Q. *et al.* Characterization of Human Homeodomain-interacting Protein Kinase 4 (HIPK4) as a Unique Member of the HIPK Family. 16 (2010).
108. de la Vega, L. *et al.* Control of nuclear HIPK2 localization and function by a SUMO interaction motif. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1813**, 283–297 (2011).
109. Saul, V. V. & Schmitz, M. L. Posttranslational modifications regulate HIPK2, a driver of proliferative diseases. *J. Mol. Med.* **91**, 1051–1058 (2013).
110. Shojima, N. *et al.* Depletion of homeodomain-interacting protein kinase 3 impairs insulin secretion and glucose tolerance in mice. *Diabetologia* **55**, 3318–3330 (2012).
111. Curtin, J. F. & Cotter, T. G. JNK Regulates HIPK3 Expression and Promotes Resistance to Fas-mediated Apoptosis in DU 145 Prostate Carcinoma Cells. 12.
112. Ferguson, F. M. & Gray, N. S. Kinase inhibitors: the road ahead. *Nat. Rev. Drug Discov.* **17**, 353–377 (2018).
113. Mullard, A. Targeted protein degraders crowd into the clinic. *Nat. Rev. Drug Discov.* **20**, 247–250 (2021).
114. Cromm, P. M. & Crews, C. M. Targeted Protein Degradation: from Chemical Biology to Drug Discovery. *Cell Chem. Biol.* **24**, 1181–1190 (2017).
115. Dikic, I. Proteasomal and Autophagic Degradation Systems. *Annu. Rev. Biochem.* **86**, 193–224 (2017).
116. Dale, B. *et al.* Advancing targeted protein degradation for cancer therapy. *Nat. Rev. Cancer* **21**, 638–654 (2021).

117. Kozicka, Z. & Thomä, N. H. Haven't got a glue: Protein surface variation for the design of molecular glue degraders. *Cell Chem. Biol.* **28**, 1032–1047 (2021).
118. Henley, M. J. & Koehler, A. N. Advances in targeting 'undruggable' transcription factors with small molecules. *Nat. Rev. Drug Discov.* **20**, 669–688 (2021).
119. Wolf, E. & Eilers, M. Targeting MYC Proteins for Tumor Therapy. *Annu. Rev. Cancer Biol.* **4**, 61–75 (2020).
120. Gonda, T. J. & Ramsay, R. G. Directly targeting transcriptional dysregulation in cancer. *Nat. Rev. Cancer* **15**, 686–694 (2015).
121. Moret, N. *et al. A resource for exploring the understudied human kinome for research and therapeutic opportunities*. http://biorxiv.org/lookup/doi/10.1101/2020.04.02.022277 (2020) doi:10.1101/2020.04.02.022277.
122. Finan, C. *et al.* The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* (2017).
123. Asghar, U., Witkiewicz, A. K., Turner, N. C. & Knudsen, E. S. The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nat. Rev. Drug Discov.* **14**, 130–146 (2015).
124. Hui, R., Boer, R., Lim, E., Yeo, B. & Lynch, J. CDK4/6 inhibitor plus endocrine therapy for hormone receptor-positive, HER2-negative metastatic breast cancer: The new standard of care. *Asia Pac. J. Clin. Oncol.* **17**, 3–14 (2021).
125. Kwiatkowski, N. *et al.* Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).
126. Olson, C. M. *et al.* Development of a Selective CDK7 Covalent Inhibitor Reveals Predominant Cell-Cycle Phenotype. *Cell Chem. Biol.* **26**, 792-803.e10 (2019).
127. Cee, V. J., Chen, D. Y.-K., Lee, M. R. & Nicolaou, K. C. Cortistatin A is a High-Affinity Ligand of Protein Kinases ROCK, CDK8, and CDK11. *Angew. Chem. Int. Ed.* **48**, 8952–8957 (2009).
128. Stransky, N., Cerami, E., Schalm, S., Kim, J. L. & Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **5**, 4846 (2014).
129. Olson, C. M. *et al.* Pharmacological perturbation of CDK9 using selective CDK9 inhibition or degradation. *Nat. Chem. Biol.* **14**, 163–170 (2018).
130. Bogdanova, N. V. *et al.* A Splice Site Variant of CDK12 and Breast Cancer in Three Eurasian Populations. *Front. Oncol.* **9**, 493 (2019).
131. Quereda, V. Therapeutic Targeting of CDK12/CDK13 in Triple-Negative Breast Cancer. *Cancer Cell* 22 (2019).
132. Tien, J. F. *et al.* CDK12 regulates alternative last exon mRNA splicing and promotes breast cancer cell invasion. *Nucleic Acids Res.* **45**, 6698–6716 (2017).
133. Blazek, D. *et al.* The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. *Genes Dev.* **25**, 2158–2172 (2011).
134. Dubbury, S. J., Boutz, P. L. & Sharp, P. A. CDK12 regulates DNA repair genes by suppressing intronic polyadenylation. *Nature* **564**, 141–145 (2018).
135. Johnson, S. F. *et al.* CDK12 Inhibition Reverses De Novo and Acquired PARP Inhibitor Resistance in BRCA Wild-Type and Mutated Models of Triple-Negative Breast Cancer. *Cell Rep.* **17**, 2367–2381 (2016).
136. Fan, Z. *et al.* CDK13 cooperates with CDK12 to control global RNA polymerase II processivity. *Sci. Adv.* **6**, eaaz5041 (2020).
137. Dong, X. *et al.* CDK13 RNA Over-Editing Mediated by ADAR1 Associates with Poor Prognosis of Hepatocellular Carcinoma Patients. *Cell. Physiol. Biochem.* **47**, 2602–2612 (2018).
138. Zhang, T. *et al.* Covalent targeting of remote cysteine residues to develop CDK12 and CDK13 inhibitors. *Nat. Chem. Biol.* **12**, 876–884 (2016).
139. Paruch, K. *et al.* Discovery of Dinaciclib (SCH 727965): A Potent and Selective Inhibitor of Cyclin-Dependent Kinases. *ACS Med. Chem. Lett.* **1**, 204–208 (2010).

140. Boni, J., Rubio-Perez, C., López-Bigas, N., Fillat, C. & De La Luna, S. The DYRK Family of Kinases in Cancer: Molecular Functions and Therapeutic Opportunities. *Cancers* **12**, 2106 (2020).
141. D'Orazi, G., Rinaldo, C. & Soddu, S. Updates on HIPK2: a resourceful oncosuppressor for clearing cancer. *J. Exp. Clin. Cancer Res.* **31**, 63 (2012).
142. Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. & McCawley, P. The establishment of two cell lines from the insectspodoptera frugiperda (lepidoptera; noctuidae). *In Vitro* **13**, 213–217 (1977).
143. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
144. Bieniossek, C., Imasaki, T., Takagi, Y. & Berger, I. MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem. Sci.* **37**, 49–57 (2012).
145. Allen, J. J. *et al.* A semisynthetic epitope for kinase substrates. *Nat. Methods* **4**, 511–516 (2007).
146. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
147. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
148. Adams, P. D. *et al. PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
149. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
150. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
151. Diederichs, K. & Karplus, P. A. Better models by discarding data? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1215–1222 (2013).
152. Stevenson, C. E., Sargent, F., Buchanan, G., Palmer, T. & Lawson, D. M. Crystal structure of the molybdenum cofactor biosynthesis protein MobA from Escherichia coli at near-atomic resolution. *Struct. Lond. Engl. 1993* **8**, 1115–1125 (2000).
153. Agnew, C. *et al.* The crystal structure of the protein kinase HIPK2 reveals a unique architecture of its CMGC-insert region. *J. Biol. Chem.* **294**, 13545–13559 (2019).
154. Kannan, N. & Neuwald, A. F. Evolutionary constraints associated with functional specificity of the CMGC protein kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2α. *Protein Sci.* **13**, 2059–2077 (2004).
155. Yu, D., Cattoglio, C., Xue, Y. & Zhou, Q. A complex between DYRK1A and DCAF7 phosphorylates the C-terminal domain of RNA polymerase II to promote myogenesis. *Nucleic Acids Res.* **47**, 4462–4475 (2019).
156. Czudnochowski, N., Bösken, C. A. & Geyer, M. Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nat. Commun.* **3**, 842 (2012).
157. Mayfield, J. E. *et al.* Tyr1 phosphorylation promotes phosphorylation of Ser2 on the C-terminal domain of eukaryotic RNA polymerase II by P-TEFb. *eLife* **8**, e48725 (2019).
158. Yu, M. *et al.* Suppression of MAPK11 or HIPK3 reduces mutant Huntingtin levels in Huntington's disease models. *Cell Res.* **27**, 1441–1465 (2017).
159. Fu, Y., Sun, X. & Lu, B. HIPK3 modulates autophagy and HTT protein levels in neuronal and mouse models of Huntington disease. *Autophagy* **14**, 169–170 (2018).
160. Cozza, G. *et al.* Synthesis and Properties of a Selective Inhibitor of Homeodomain–Interacting Protein Kinase 2 (HIPK2). *PLoS ONE* **9**, e89176 (2014).
161. Němec, V. *et al.* Highly selective inhibitors of protein kinases CLK and HIPK with the furo[3,2-b]pyridine core. *Eur. J. Med. Chem.* **215**, 113299 (2021).
162. Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **29**, 1046–1051 (2011).
163. Goetz, M. P. *et al.* MONARCH 3: Abemaciclib As Initial Therapy for Advanced Breast Cancer. *J. Clin. Oncol.* **35**, 3638–3646 (2017).

164. Cunningham, N. C. & Turner, N. C. Understanding divergent trial results of adjuvant CDK4/6 inhibitors for early stage breast cancer. *Cancer Cell* **39**, 307–309 (2021).
165. Gelbert, L. M. *et al.* Preclinical characterization of the CDK4/6 inhibitor LY2835219: in-vivo cell cycle-dependent/independent anti-tumor activities alone/in combination with gemcitabine. *Invest. New Drugs* **32**, 825–837 (2014).
166. O'Leary, B., Finn, R. S. & Turner, N. C. Treating cancer with selective CDK4/6 inhibitors. *Nat. Rev. Clin. Oncol.* **13**, 417–430 (2016).
167. Chen, P. *et al.* Spectrum and Degree of CDK Drug Interactions Predicts Clinical Performance. *Mol. Cancer Ther.* **15**, 2273–2281 (2016).
168. Huang, F. *et al.* Reversible phosphorylation of cyclin T1 promotes assembly and stability of P-TEFb. *eLife* **10**, e68473 (2021).
169. Huang, F., Feng, Y., Peterlin, B. M. & Fujinaga, K. P-TEFb is degraded by Siah1/2 in quiescent cells.
170. Marshall, N. F. & Price, D. H. Control of Formation of Two Distinct Classes of RNA Polymerase II Elongation Complexes. *Mol. Cell. Biol.* **12**, 2078–2090 (1992).
171. Fu, T.-J., Peng, J., Lee, G., Price, D. H. & Flores, O. Cyclin K Functions as a CDK9 Regulatory Subunit and Participates in RNA Polymerase II Transcription. *J. Biol. Chem.* **274**, 34527–34530 (1999).
172. Yu, D. S. *et al.* Cyclin-dependent kinase 9–cyclin K functions in the replication stress response. *EMBO Rep.* **11**, 876–882 (2010).
173. Lin, X., Taube, R., Fujinaga, K. & Peterlin, B. M. P-TEFb Containing Cyclin K and Cdk9 Can Activate Transcription via RNA. *J. Biol. Chem.* **277**, 16873–16878 (2002).
174. Baek, K., Brown, R. S., Birrane, G. & Ladias, J. A. A. Crystal Structure of Human Cyclin K, a Positive Regulator of Cyclin-dependent Kinase 9. *J. Mol. Biol.* **366**, 563–573 (2007).
175. Bartkowiak, B. *et al.* CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.* **24**, 2303–2316 (2010).
176. Kohoutek, J. & Blazek, D. Cyclin K goes with Cdk12 and Cdk13. *Cell Div.* **7**, 12 (2012).
177. Ji, C. *et al.* HNRNP R negatively regulates transcription by modulating the association of P-TEFB with 7SK and BRD4. *EMBO Rep.* **23**, e55432 (2022).
178. Insco, M. L. *et al.* Oncogenic *CDK13* mutations impede nuclear RNA surveillance. *Science* **380**, eabn7625 (2023).
179. Fisher, R. P. Tumor suppression by RNA surveillance. *Science* **380**, 240–241 (2023).
180. Czudnochowski, N., Vollmuth, F., Baumann, S., Vogel-Bachmayr, K. & Geyer, M. Specificity of Hexim1 and Hexim2 Complex Formation with Cyclin T1/T2, Importin α and 7SK snRNA. *J. Mol. Biol.* **395**, 28–41 (2010).
181. Schulte, A. *et al.* Identification of a Cyclin T-Binding Domain in Hexim1 and Biochemical Analysis of Its Binding Competition with HIV-1 Tat. *J. Biol. Chem.* **280**, 24968–24977 (2005).
182. Michels, A. A. & Bensaude, O. Hexim1, an RNA-controlled protein hub. *Transcription* **9**, 262–271 (2018).
183. Garriga, J. *et al.* Upregulation of cyclin T1/CDK9 complexes during T cell activation. *Oncogene* **17**, 3093–3102 (1998).
184. Marshall, R. M., Salerno, D., Garriga, J. & Graña, X. Cyclin T1 Expression Is Regulated by Multiple Signaling Pathways and Mechanisms during Activation of Human Peripheral Blood Lymphocytes. *J. Immunol.* **175**, 6402–6411 (2005).
185. Kiernan, R. E. *et al.* Interaction between Cyclin T1 and SCF ^{SKP2} Targets CDK9 for Ubiquitination and Degradation by the Proteasome. *Mol. Cell. Biol.* **21**, 7956–7970 (2001).
186. Cooper, K. F., Mallory, M. J. & Strich, R. Oxidative Stress-Induced Destruction of the Yeast C-Type Cyclin Ume3p Requires Phosphatidylinositol-Specific Phospholipase C and the 26S Proteasome. *Mol. Cell. Biol.* **19**, 3338–3348 (1999).

187. Krasley, E., Cooper, K. F., Mallory, M. J., Dunbrack, R. & Strich, R. Regulation of the Oxidative Stress Response Through Slt2p-Dependent Destruction of Cyclin C in *Saccharomyces cerevisiae*. *Genetics* **172**, 1477–1486 (2006).
188. Willis, S. D., Hanley, S. E., Beishke, T., Tati, P. D. & Cooper, K. F. Ubiquitin–proteasome-mediated cyclin C degradation promotes cell survival following nitrogen starvation. *Mol. Biol. Cell* **31**, 1015–1031 (2020).
189. Davis, M. A. *et al.* The SCF–Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with Mediator. *Genes Dev.* **27**, 151–156 (2013).
190. Stieg, D. C. *et al.* A complex molecular switch directs stress-induced cyclin C nuclear release through SCF ^{Grr1} -mediated degradation of Med13. *Mol. Biol. Cell* **29**, 363–375 (2018).
191. Mayor-Ruiz, C. *et al.* Rational discovery of molecular glue degraders via scalable chemical profiling. *Nat. Chem. Biol.* **16**, 1199–1207 (2020).
192. Słabicki, M. *et al.* The CDK inhibitor CR8 acts as a molecular glue degrader that depletes cyclin K. *Nature* **585**, 293–297 (2020).
193. Dieter, S. M. *et al.* Degradation of CCNK/CDK12 is a druggable vulnerability of colorectal cancer. *Cell Rep.* **36**, 109394 (2021).
194. Gracias, D. *et al.* Identification of a Novel Cyclin K Degrader Molecule with Selective Anti-AML Activity. *Blood* **140**, 10684–10684 (2022).
195. Lv, L. *et al.* Discovery of a molecular glue promoting CDK12-DDB1 interaction to trigger cyclin K degradation. *eLife* **9**, e59994 (2020).
196. *Cell Cycle Oscillators: Methods and Protocols*. vol. 2329 (Springer US, 2021).
197. Endicott, J. A., Noble, M. E. M. & Johnson, L. N. The Structural Basis for Control of Eukaryotic Protein Kinases. *Annu. Rev. Biochem.* **81**, 587–613 (2012).
198. Lu, H. *et al.* Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**, 318–323 (2018).
199. Bösken, C. A. *et al.* The structure and substrate specificity of human Cdk12/Cyclin K. *Nat. Commun.* **5**, 3505 (2014).
200. Honda, R. *et al.* The structure of cyclin E1/CDK2: implications for CDK2 activation and CDK2-independent roles. *EMBO J.* **24**, 452–463 (2005).
201. Russo, A. A., Jeffrey, P. D. & Pavletich, N. P. Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat. Struct. Mol. Biol.* **3**, 696–700 (1996).
202. Brown, N. R. *et al.* CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nat. Commun.* **6**, 6769 (2015).
203. Brown, N. R. *et al.* Cyclin B and Cyclin A Confer Different Substrate Recognition Properties on CDK2. *Cell Cycle* **6**, 1350–1359 (2007).
204. Murphy, J. M. The long-awaited structure of HIPK2. *J. Biol. Chem.* **294**, 13560–13561 (2019).
205. Rinaldo, C., Siepi, F., Prodosmo, A. & Soddu, S. HIPKs: Jack of all trades in basic nuclear activities. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1783**, 2124–2129 (2008).
206. Glenewinkel, F. *et al.* The adaptor protein DCAF7 mediates the interaction of the adenovirus E1A oncoprotein with the protein kinases DYRK1A and HIPK2. *Sci. Rep.* **6**, 28241 (2016).
207. Ritterhoff, S. *et al.* The WD40-repeat protein Han11 functions as a scaffold protein to control HIPK2 and MEKK1 kinase functions. *EMBO J.* **29**, 3750–3761 (2010).
208. Hofmann, T. G., Glas, C. & Bitomsky, N. HIPK2: A tumour suppressor that controls DNA damage-induced cell fate and cytokinesis. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **35**, 55–64 (2013).
209. Klaeger, S. *et al.* The target landscape of clinical kinase drugs. *Science* **358**, (2017).
210. Hafner, M. *et al.* Multiomics Profiling Establishes the Polypharmacology of FDA-Approved CDK4/6 Inhibitors and the Potential for Differential Clinical Activity. *Cell Chem. Biol.* **26**, 1067-1080.e8 (2019).
211. Demuro, S., Di Martino, R. M. C., Ortega, J. A. & Cavalli, A. GSK-3β, FYN, and DYRK1A: Master Regulators in Neurodegenerative Pathways. *Int. J. Mol. Sci.* **22**, 9098 (2021).

212. Lindberg, M. F. *et al.* Comparative Efficacy and Selectivity of Pharmacological Inhibitors of DYRK and CLK Protein Kinases. *J. Med. Chem.* **66**, 4106–4130 (2023).
213. Lee, H. & Hoe, H.-S. Inhibition of CDK4/6 regulates AD pathology, neuroinflammation and cognitive function through DYRK1A/STAT3 signaling. *Pharmacol. Res.* **190**, 106725 (2023).
214. Anshabo, A. T., Milne, R., Wang, S. & Albrecht, H. CDK9: A Comprehensive Review of Its Biology, and Its Role as a Potential Target for Anti-Cancer Agents. *Front. Oncol.* **11**, 678559 (2021).
215. Kciuk, M., Gielecińska, A., Mujwar, S., Mojzych, M. & Kontek, R. Cyclin-dependent kinases in DNA damage response. *Biochim. Biophys. Acta BBA - Rev. Cancer* **1877**, 188716 (2022).
216. Storch, K. & Cordes, N. The impact of CDK9 on radiosensitivity, DNA damage repair and cell cycling of HNSCC cancer cells. *Int. J. Oncol.* **48**, 191–198 (2016).
217. Zhang, H. *et al.* SIRT2 directs the replication stress response through CDK9 deacetylation. *Proc. Natl. Acad. Sci.* **110**, 13546–13551 (2013).
218. Zhou, K. *et al.* Disrupting the Cdk9/Cyclin T1 heterodimer of 7SK snRNP for the Brd4 and AFF1/4 guided reconstitution of active P-TEFb. *Nucleic Acids Res.* **50**, 750–762 (2022).
219. Diehl, C. J. & Ciulli, A. Discovery of small molecule ligands for the von Hippel-Lindau (VHL) E3 ligase and their use as inhibitors and PROTAC degraders. *Chem. Soc. Rev.* **51**, 8216–8257 (2022).
220. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13–23 (2017).
221. Boehning, M. *et al.* RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).

6. APPENDIX

6.1 Appendix A

Tab.5 | Crystallographic data collection and refinement statistics of HIPK3 and DYRK1A.

	Apo-HIPK3	HIPK3-abemaciclib	DYRK1A-abemaciclib
Data collection[a]			
Beamline	SLS X06SA	SLS X06SA	SLS X06SA
Wavelength (Å)	0.9999	0.9999	1.0000
Space group	P 3_2 2 1	P 3_2 2 1	P 2_1 2_1 2_1
Unit cell:			
a, b, c (Å)	80.28, 80.28, 181.99	81.06, 81.06, 178.82	76.70, 109.98, 112.49
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90
Resolution range (Å)	45.71- 2.50 (2.59 -2.5)	45.44- 2.81 (2.91- 2.81)	41.93 - 1.82 (1.89 - 1.82)
Unique reflections	17,854 (236)	17,230 (1679)	85,615 (8,460)
Multiplicity	17.0 (10.5)	16.3 (17.4)	13.4 (13.6)
Completeness:			
spherical (%)	73.69 (9.84)	99.91 (100.0)	99.59 (99.35)
ellipsoidal (%)	93.8		
Mean I/sigma(I)	12.38 (0.85)	23.68 (1.79)	16.36 (1.41)
R_{meas}	0.126 (5.02)	0.06367 (1.596)	0.07891 (1.718)
$CC_{1/2}$	0.89 (0.38)	1 (0.829)	1 (0.836)
Reflections used:			
in refinement	17,854 (236)	17,222 (1,679)	85,539 (8,447)
used for R-free	938 (16)	862 (84)	4,279 (422)
Refinement			
Model content	A: HIPK3 (184-550)	A: HIPK3 (184-551): abemaciclib	A: DYRK1A (134-480, Δ411-412): abemaciclib, 2 citrates, 1 Li B: DYRK1A (133-480, Δ407-411): abemaciclib, 2 citrates, 1 Li
# of atoms	2,9323	2,940	5,688
# of ligands	4	69	224
# of solvent	51	15	375
Solvent content (%)	49	64	57
R_{work}	0.2429 (0.5255)	0.2464 (0.3327)	0.1856 (0.3321)
R_{free}	0.2730 (0.5908)	0.2748 (0.3842)	0.2111 (0.3434)
RMS:			
deviation bonds (Å)	0.002	0.007	0.009
deviation angles (°)	0.44	1.25	1.09
Ramachandran:			
favored (%)	95.86	92.29	96.44
allowed (%)	3.87	6.61	3.41

Average B-factor	81.57	115.63	48.45
Macromolecules	81.61	115.86	48.08
Ligands	84.80	98.03	50.98
Solvent	78.93	113.47	52.58
PDB accession code	707I	707J	707K

[a] Values in parentheses are for the highest resolution shell. R_{free}-value is equivalent to the R-value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.

Tab.6 | Crystallographic data collection and refinement statistics of the Cdk12/CycK-NVP2 and Cdk12/CycT1-NVP2 complex.

	Cdk12/CycK-NVP2	Cdk12/CycT1-NVP2
Data collection [a]		
Beamline	SLS X06SA	SLS X06SA
Wavelength (Å)	0.9763	0.8266
Space group	P 3_1 2 1	P 1 2_1 1
Unit cell:		
a, b, c (Å)	112.517 112.517 100.328	67.811 124.506 83.808
α, β, γ (°)	90 90 120	90 92.777 90
Resolution range (Å)	43.83 - 2.751 (2.849 - 2.751)	35.39 - 2.032 (2.105 - 2.032)
Unique reflections	19430 (21)	88009 (5079)
Multiplicity	29.6 (30.1)	5.2 (4.2)
Completeness (%)	82.17 (1.10)	95.19 (57.10)
Mean I/sigma(I)	9.96 (0.10)	8.81 (0.51)
R_{meas}	0.2805 (11.92)	0.1367 (2.484)
$CC_{1/2}$	0.999 (0.333)	0.997 (0.257)
Reflections used:		
in refinement	15996 (21)	84960 (5080)
used for R-free	1609 (3)	2009 (122)
Refinement		
Model content	A: Cdk12 (716-1040): NVP2 B: CycK (20-265)	A: Cdk12 (718-1035): NVP2 B: CycT1 (9-250) C: Cdk12 (717-1035): NVP2 D: CycT1 (9-248)
# of atoms	4404	8853
# of ligands	36	78
# of solvent	1	271
R_{work}	0.2424 (1.0322)	0.2293 (0.3724)
R_{free}	0.2997 (1.0969)	0.2599 (0.3752)
RMS:		
deviation bonds (Å)	0.003	0.003
deviation angles (°)	0.56	0.55
Ramachandran:		
favored (%)	90.63	97.70
allowed (%)	8.11	2.12

Average B-factor	134.5	47.70
Macromolecules	134.5	47.59
Ligands	131.8	42.35
Solvent	86.2	52.72
PDB accession code	-	-

[a] Values in parentheses are for the highest resolution shell. R_{free}-value is equivalent to the R-value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.

6.2 Appendix B

Tab.7 | Purification strategies of all recombinant proteins used in this thesis.

[1] these proteins were used for the DYRK1A and HIPK project.

[2] these proteins were used for the novel Cdk/Cyclin pair project.

Protein	Project	PDB accession code	DNA cloning source	Vector backbone	Affinity tag	Protease cleavage site	Residues	Mutations
Kinases								
DYRK1A	1	Q13627	Addgene #38913	pNIC28-Bsa4	N-term. His	tev	127-485	wt
HIPK1	1	Q86Z02	GeneArt (Regensburg)	pACEBac1	N-term. MBP	tev	154-554	wt; kdm (D135N)
HIPK2	1	Q9H2X6	GeneArt	pACEBac1	N-term. GST	tev	160-562	wt; kdm (D324N)
HIPK3	1	Q9H422	Addgene #23467	pACEBac1	N-term. GST	tev	159-562	wt; kdm (D322N)
HIPK4	1	Q8NE63	Addgene #23760	pACEBac1	N-term. MBP	tev	2-616	wt; kdm (D136N)
Cdk4 CycD3	1	P11802 P30281	HEK cDNA HEK cDNA	pACEBac1 plDK	N-term. His N-term. GST	tev tev	1-303 1-292	wt wt
Cdk9 CycT1	1	P50750 O60563	Barboric Lab Peterlin Lab	pACEBac1 pGEX-4T1	N-term. His N-term. GST	tev tev	1-372 1-272	wt wt
Cdk12 CycK CAK	1	Q9NYV4 O75909 P43568	GeneArt GeneArt GeneArt	pACEBac1 plDK plDC	N-term. GST N-term. GST -	tev tev -	714-1063 1-267 1-368	wt wt wt (*S.cerevisiae*)

Protein	Expression & Purification strategy
DYRK1A	*E.coli,* IPTG induction, expression 20°C overnight Standard His-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion chromatography on a Superdex75 column.
HIPK1	*Sf9* insect cells Standard MBP-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion chromatography on a Superdex200 column.
HIPK2	*Sf9* insect cell Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion chromatography on a Superdex200 column.
HIPK3	*Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion chromatography on a Superdex200 column.
HIPK4	*Sf9* insect cells Standard MBP-affinity chromatography, no tag-cleavage; standard size exclusion chromatography on a Superdex200 column.
Cdk4 **CycD3**	*Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion chromatography on a Superdex200 column.
Cdk9 **CycT1**	Cdk9: *Sf9* insect cells; Standard His-affinity chromatography followed by Tev cleavage overnight; CycT1: *E.coli,* IPTG induction, expression 20°C overnight; Standard GST-affinity chromatography followed by Tev-cleavage overnight; after affinity chromatography both proteins were concentrated and mixed in a 1:1 molar ratio and applied to size exclusion. chromatography on a Superdex200 column.
Cdk12 **CycK** **CAK**	Vector Cre-recombination; *Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion chromatography on a Superdex200 column.

Protein	Project	PDB accession code	DNA cloning source	Vector backbone	Affinity tag	Protease cleavage site	Residues	Mutations
Kinases								
Cdk7	2	P50613	GeneArt	pACEBac1	N-term. GST	tev	2-346	wt
Cdk8	2	P49336	Addgene #68856	pACEBac1	N-term. GST	tev	1-464	wt
Cdk9	2	P50750	Peterlin Lab	pACEBac1	N-term. His	tev	1-372	wt
Cdk11B$_{p110}$	2	P21127	GeneArt	pACEBac1	N-term. GST	tev	1-795	wt
Cdk13	2	Q14004	GeneArt	pACEBac1	N-term. GST	tev	694-1039	wt
CycC **Med12**	2	P24863 Q93074	Addgene #8961 Addgene #49240	pACEBac1 pACEBac1	- C-term. Strep	- -	1-283 1-100	wt wt
CycH	2	P51946	GeneArt	pGEX-4T1	N-term. GST	tev	2-346	wt
CycK	2	O75909	GeneArt	pET28a	N-term. MBP	tev	1-267	wt

Protein	Expression & Purification strategy
Cdk7	*Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk8	*Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk9	*Sf9* insect cells Standard His-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk11B$_{p110}$	*Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk13	*Sf9* insect cells Standard GST-affinity chromatography followed by Tev-on-column cleavage overnight; standard size exclusion on a Superdex200 column.
CycC **Med12**	*Sf9* insect cells, co-expression (CycC alone is not stable) Standard Strep-affinity chromatography; standard size exclusion chromatography on a Superdex200 column.
CycH	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
CycK	*E.coli,* autoinduction, expression 20°C overnight Standard MBP-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.

Protein	Project	PDB accession code	DNA cloning source	Vector backbone	Affinity tag	Protease cleavage site	Residues	Mutations
Kinases								
CycT1	2	O60563	Peterlin Lab	pGEX-4T1	N-term. GST	tev	1-272	wt; T149A; T149E
Cdk7 **CycH**	2	P50613 P51946	GeneArt GeneArt	pACEBac1 pIDK	N-term. GST -	tev -	1-346 1-323	wt wt
Cdk7 **CycH** **Mat1**	2	P50613 P51946 P51948	GeneArt GeneArt GeneArt	pACEBac1 pIDK pIDC	N-term. GST - -	tev - -	1-346 1-323 1-309	wt wt wt
Cdk8 **CycC**	2	P49336 P24863	Addgene #68856 Addgene #8961	pACEBac1 pACEBac1	- C-term. Strep	- -	1-464 1-283	wt wt
Cdk8 **CycC** **Med12**	2	P49336 P24863 Q93074	Addgene #68856 Addgene #8961 Addgene #49240	pACEBac1 pACEBac1 pACEBac1	- - C-term. Strep	- - -	1-464 1-283 1-100	wt wt wt
Cdk9 **CycK**	2	P50750 O75909	Peterlin Lab Blazek Lab	pACEBac1 pACEBac1	N-term. His N-term. GST	tev tev	1-372 1-267	wt; kdm (D149N) wt
Cdk9 **CycT1**	2	P50750 O60563	Peterlin Lab Peterlin Lab	pACEBac1 pACEBac1	N-term. His N-term. GST	tev tev	1-372 1-272	wt; kdm (D149N) wt
Cdk12 **CycK**	2	Q9NYV4 O75909	GeneArt GeneArt	pACEBac1 pACEBac1	N-term. GST N-term. GST	tev tev	714-1063 1-267	wt; kdm (D895N) wt

Protein	Expression & Purification strategy
CycT1	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk7 **CycH**	Vector Cre-recombination; *Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk7 **CycH** **Mat1**	Vector Cre-recombination; *Sf9* insect cells Standard GST-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk8 **CycC**	*Sf9* insect cells Standard Strep-affinity chromatography; standard size exclusion chromatography on a Superdex200 column.
Cdk8 **CycC** **Med12**	*Sf9* insect cells Standard Strep-affinity chromatography; standard size exclusion chromatography on a Superdex200 column.
Cdk9 **CycK**	*Sf9* insect cells; co-expression with pACEBac1-CAK Standard His-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk9 **CycT1**	*Sf9* insect cells; co-expression with pACEBac1-CAK Standard His-affinity chromatography followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk12 **CycK**	*Sf9* insect cells; co-expression with pACEBac1-CAK Standard GST-affinity chromatography followed by Tev on-column cleavage overnight; standard size exclusion on a Superdex200 column.

Protein	Project	PDB accession code	DNA cloning source	Vector backbone	Affinity tag	Protease cleavage site	Residues	Mutations
Kinases								
Cdk12 **CycT1**	2	Q9NYV4 O60563	GeneArt Peterlin Lab	pACEBac1 pACEBac1	N-term. GST N-term. GST	tev tev	714-1063 1-272	wt; kdm (D895N) wt
Cdk13 **CycK**	2	Q14004 O75909	GeneArt GeneArt	pACEBac1 pACEBac1	N-term. GST N-term. GST	tev tev	694-1039 1-267	wt; kdm (D837N) wt
Cdk13 **CycT1**	2	Q14004 O60563	GeneArt Peterlin Lab	pACEBac1 pACEBac1	N-term. GST N-term. GST	tev tev	694-1039 1-272	wt; kdm (D837N) wt
Cdk19 **CycC**	2	Q9BWU1 P24863	Addgene #68858 Addgene #8961	pACEBac1 pACEBac1	- C-term. Strep	- -	1-502 1-283	wt wt
Cdk19 **CycC** **Med12**	2	Q9BWU1 P24863 Q93074	Addgene #68858 Addgene #8961 Addgene #49240	pACEBac1 pACEBac1 pACEBac1	- - C-term. Strep	- - -	1-502 1-283 1-100	wt wt wt

Protein	Expression & Purification strategy
Cdk12 **CycT1**	*Sf9* insect cells; co-expression with pACEBac1-CAK Standard GST-affinity chromatography followed by Tev on-column cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk13 **CycK**	*Sf9* insect cells; co-expression with pACEBac1-CAK Standard GST-affinity chromatography followed by Tev-on-column cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk13 **CycT1**	*Sf9* insect cells; co-expression with pACEBac1-CAK Standard GST-affinity chromatography followed by Tev-on-column cleavage overnight; standard size exclusion on a Superdex200 column.
Cdk19 **CycC**	*Sf9* insect cells Standard Strep-affinity chromatography followed by standard size exclusion on a Superdex200 column.
Cdk19 **CycC** **Med12**	*Sf9* insect cells Standard Strep-affinity chromatography followed by standard size exclusion on a Superdex200 column.

Protein	Project	PDB accession code	DNA cloning source	Vector backbone	Affinity tag	Protease cleavage site	Residues	Mutations
Substrates								
pol II CTD	1 + 2	P24928	BioScience (UK)	pGEX-6P1	N-term. GST	3C	1587-1970	wt
SPT5	1 + 2	O00267	HEK cDNA	pET28a	N-term. GST	tev	748-1087	wt
c-Myc	1 + 2	P01106	BioScience	pPROEx-HTa	N-term. His	tev	17-167	wt
Rb1	1	P06400	Addgene #82275	pGEX-6P1	N-term. GST	tev	761-928	wt
SPT6	2	Q7KZ85	GeneArt	pGEX-4T1	N-term. GST	tev	1323-1534	wt
CDC5L	2	Q99459	GeneArt	pGEX-4T1	N-term. GST	tev	370-505	wt
CSTF2	2	P33240	GeneArt	pGEX-4T1	N-term. GST	tev	509-577	wt
SF3B1	2	O75533	Addgene #68858	pGEX-4T1	N-term. GST	tev	113-461	wt

Protein	Expression & Purification strategy
pol II CTD	*E.coli,* IPTG induction, expression 20°C overnight Standard GST-affinity chromatography, no tag cleavage; standard size exclusion on a Superdex200 column.
SPT5	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
c-Myc	*E.coli,* IPTG induction, expression 37°C 4 hours Inclusion body purification followed by His-affinity chromatography, no tag cleavage, standard size exclusion chromatography on a Superdex75 column
Rb1	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
SPT6	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
CDC5L	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
CSTF2	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
SF3B1	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, followed by Tev-cleavage; standard size exclusion on a Superdex200 column.

Protein	Project	PDB accession code	DNA cloning source	Vector backbone	Affinity tag	Protease cleavage site	Residues	Mutations
Substrates								
Xrn2	1 + 2	P24928	HEK cDNA	pET28a	N-term. His-Thioredoxin	tev	1-783	wt
SRSF1	1 + 2	Q07955	Addgene #99020	pGEX-4T1	N-term. GST	tev	1-249	wt
SRSF2	1 + 2	Q01130	GeneArt	pGEX-4T1	N-term. GST	tev	1-221	wt
SRSF7	1	Q16629	GeneArt	pGEX-4T1	N-term. GST	tev	1-238	wt
ZC3H14	2	Q6PJT7	Zon Lab	pET28a	N-term. MBP	tev	1-736	wt; S475A
Regulators								
Hexim1	2	O94992	Barboric Lab	pGEX-4T1	N-term. GST	tev	200-359	wt
Brd4	2	O60885	GeneArt	pGEX-4T1	N-term. GST	tev	1338-1400	wt; murine

Protein	Expression & Purification strategy
Xrn2	*E.coli,* autoinduction, expression 20°C overnight Standard His-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
SRSF1	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
SRSF2	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
SRSF7	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, no Tev-cleavage; standard size exclusion on a Superdex200 column.
ZC3H14	*E.coli,* autoinduction, expression 20°C overnight Ammonium sulfate precipitation followed by MBP-affinity chromatography, no Tev-cleavage; standard size exclusion chromatography on a Superdex200 column.
Hexim1	*E.coli,* autoinduction, expression 20°C overnight Standard GST-affinity chromatography, followed by Tev-cleavage overnight; standard size exclusion on a Superdex200 column.
Brd4	*E.coli,* autoinduction, expression 20°C overnight. Standard GST-affinity chromatography, followed by Tev-cleavage overnight; standard size exclusion on a Superdex75 column.