CellPress
OPEN ACCESS

Structure

Review

High-resolution cryo-EM of a small protein complex: The structure of the human CDK-activating kinase

Basil J. Greber[1,*]
[1]Division of Structural Biology, The Institute of Cancer Research, London SW3 6JB, UK
*Correspondence: basil.greber@icr.ac.uk
https://doi.org/10.1016/j.str.2024.03.003

SUMMARY

The human CDK-activating kinase (CAK) is a multifunctional protein complex and key regulator of cell growth and division. Because of its critical functions in regulating the cell cycle and transcription initiation, it is a key target for multiple cancer drug discovery programs. However, the structure of the active human CAK, insights into its regulation, and its interactions with cellular substrates and inhibitors remained elusive until recently due to the lack of high-resolution structures of the intact complex. This review covers the progress in structure determination of the human CAK by cryogenic electron microscopy (cryo-EM), from early efforts to recent near-atomic resolution maps routinely resolved at 2Å or better. These results were enabled by the latest cryo-EM technologies introduced after the initial phase of the "resolution revolution" and allowed the application of high-resolution methods to new classes of molecular targets, including small protein complexes that were intractable using earlier technology.

INTRODUCTION

Cyclin-dependent kinases (CDKs) are a family of serine/threonine protein kinases that are characterized by their dependence on a cyclin subunit for activity.[1] This family of kinases encompasses approximately 20 members in human cells.[2] Cyclin binding induces a series of conformational changes in the CDK that are conducive to kinase activity, including the repositioning of a specific helix in the kinase N-lobe, called the C-helix, toward the active site, which leads to active site assembly (Figures 1A and 1B).[3–5] A second critical structural element controlling CDK activity is the T-loop, a dynamic regulatory segment encompassing the sequence between the conserved DFG and APE motifs in the C-terminal kinase lobe (Figures 1A–1C).[1,3–5] The prototypical CDKs function in the control of the cell cycle and are activated by cyclins whose concentrations undergo cyclic oscillations throughout the cell cycle (hence their name). The ensuing waves of CDK activity and inactivity drive the events of the cell cycle.[6] These CDKs are only partially activated by cyclin binding and additionally require phosphorylation of their T-loop for full activity. The enzyme that places these phosphorylations is called the CDK-activating kinase (CAK). The CAK designation is assigned according to function, not sequence or structural homology, and therefore, the CAKs in different organisms are not necessarily close homologs or embedded in complexes of similar composition (see the following text).[7] Notably, the ways in which different members of the CDK family are activated vary. Some CDKs, including the transcriptional CDKs CDK9, 12, and 13, are constitutively associated with partner cyclins whose concentrations do not oscillate like those of the cell-cycle controlling cyclins but still rely on the CAK for T-loop phosphorylation.[8–11] Other CDKs, such as CDK8, do not require activation by the CAK because they are embedded in larger protein complexes and are activated by other subunits of those complexes.[12]

The human CAK, which is the primary subject of this review, is a trimeric protein complex composed of CDK7, cyclin H, and MAT1. CDK7 and cyclin H form a canonical CDK-cyclin pair composed of an enzymatic kinase subunit and a constitutively expressed activatory cyclin.[1,13] MAT1 is a unique subunit of this complex. MAT1 is involved in the assembly and functional activation of the human CAK and directly recruits the CAK to the transcription factor IIH (TFIIH) core complex to form holo-TFIIH.[14–16] Therefore, the human CAK exists in distinct trimeric and TFIIH-associated forms, and it is multifunctional, serving in several critical cellular pathways. When bound to TFIIH, the CAK phosphorylates the C-terminal heptapeptide repeat region of the largest subunit of RNA polymerase II, the so-called Pol II-CTD.[14,17] In contrast, the free, trimeric form of the CAK phosphorylates CDKs at their regulatory T-loop, thereby fully activating several CDKs, including CDKs 1, 2, 4, and 6 that are critical for control of the cell cycle (Figure 1).[13,18–21] Additionally, the human CAK has been suggested to be involved in the regulation of splicing[22] and phosphorylates multiple transcription factors.[11,19,23–25]

Being a CDK itself, CDK7 in the human CAK contains a T-loop that harbors phosphorylation sites at S164 and T170 (S170 and T176 in *Xenopus laevis* CDK7).[15,26] The T170 phosphorylation site corresponds to the conserved T-loop phosphorylation site in other CDKs, while the second phosphorylation site at S164 is unique to CAK and its sequence context (SPNR) corresponds to a consensus CDK phosphorylation site ([S/T]PX[K/R]).[15] Even though a fully integrated, comprehensive view of all the functional consequences of the different combinations of T-loop phosphorylation states in the human CAK is only beginning to emerge, it is

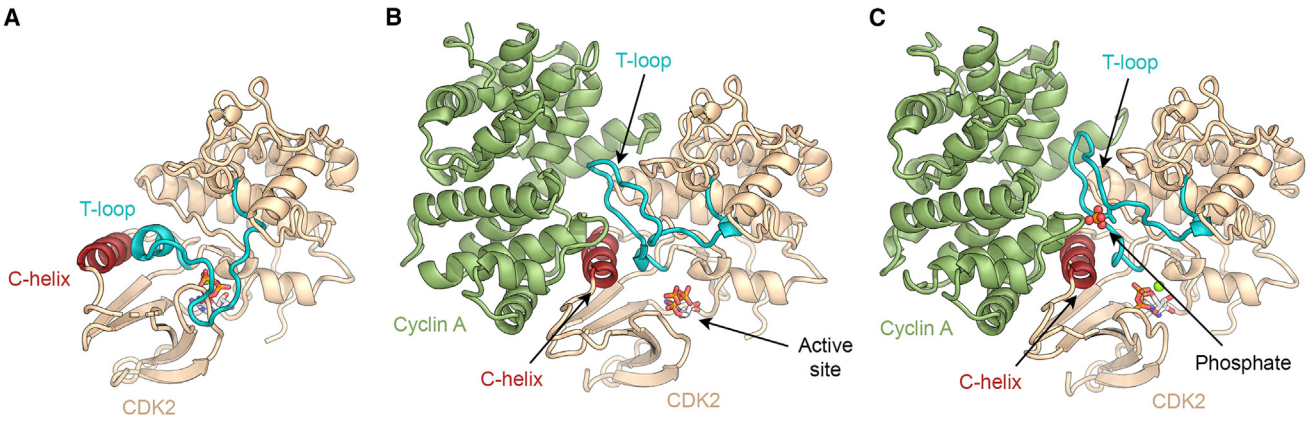

Figure 1. Molecular mechanisms of CDK activation by cyclin binding and T-loop phosphorylation

(A) In their free form, cell cycle CDKs, such as CDK2 (brown), are inactive. The C-helix (red) is removed from the active site, and the T-loop (activation segment, teal) is folded across the active site and prevents substrate access (PDB ID 1HCK).[3]

(B) When a cyclin (cyclin A, green) is bound, the C-helix is relocated toward the active site and the T-loop assumes an extended conformation. In this state, the active site is accessible to substrates (PDB ID 1FIN).[4]

(C) Full activation of cell cycle CDKs requires phosphorylation of the T-loop (T160 in CDK2) by the CAK. The phosphorylated T-loop assumes a distinct conformation, though this transition is more subtle than between the cyclin-bound and unbound forms (PDB ID 1JST).[5] It is worth noting that cyclin binding and phosphorylation do not necessarily occur in that order and may be interdependent.[21,76]

known that T-loop phosphorylation can promote both the activity and the assembly of the human CAK.[15,27,28] T170 phosphorylation has been found to contribute to substantially enhanced CAK activity toward the Pol II CTD (on the order of 3-fold[28] to 25-fold[27]) while only modestly stimulating its CAK activity[15] or leaving it almost unchanged,[19,27,28] depending on the experimental conditions. Initially, conflicting data on S164 phosphorylation suggested that it either aids CAK assembly and stimulates CDK7 activity[27] or represses the function of CAK in transcription initiation by repressing its CTD kinase activity.[29] Recent results have suggested an important role of S164 phosphorylation in CAK assembly and its activation toward the transcriptional Pol II CTD and SPT5 substrates.[28] In spite of the importance of CAK T-loop phosphorylation, neither the assembly nor the activation of the metazoan CAK strictly depends on phosphorylation of its T-loop because MAT1 can at least partially replace these roles by functioning as an assembly and activation factor for CAK.[15,16,19] The mechanistic details of how MAT1 functions in these roles will be described later in this review.

Due to the involvement of the human CAK in key cellular functions and in multiple pathways that are critical for cell growth and proliferation, it has been identified as a promising target for cancer therapeutics.[30] So far, six different CDK7 inhibitors have been clinically developed: ICEC0942, SY5609, XL102, SY1365, Q901, and LY3405105.[31–34] One of the major challenges in the clinical development of CDK-targeting drugs are the selectivity and specificity of the inhibitors.[35] This is because human cells harbor approximately 20 CDKs, some of which share high homology with each other, particularly in the active site region.[2] Furthermore, CDKs are very dynamic, as exemplified by the large structural changes illustrated in Figure 1, which can influence inhibitor binding.[36] Structure-based drug design can support the rational optimization of compounds and the discovery of highly specific enzyme inhibitors in such cases, but this approach requires high-resolution structures of the target complex.

Even though the structures of isolated cyclin H and CDK7 were determined using X-ray crystallography two decades ago,[37–39] structures of the fully assembled human CAK remained elusive until recently. Therefore, important medically relevant insight into the molecular basis of inhibitor specificity for CDK7 was lacking, and interesting biological questions, such as the mechanistic details of the function of MAT1, remained unexplored. To investigate the functional mechanisms of CAK activity and regulation, and to facilitate discovery of highly specific next-generation therapeutics via structure-based drug design methodologies, a number of laboratories pursued crystallographic studies of the intact trimeric human CAK complex[28] or its fungal homologs,[40] while we and other groups chose an alternative approach and sought to determine high-resolution structures of these complexes using cryogenic electron microscopy (cryo-EM).[41–44]

This new strategy was becoming available at the time because cryo-EM had matured from a mostly low-resolution method to a structural biology mainstay, in a process sometimes referred to as the "resolution revolution."[45] This revolution was triggered by technical developments, including better cameras that detect electrons directly, better microscopes, and improved software for image analysis.[45,46] However, high-resolution structure determination of very small protein complexes, which includes many potential drug targets, remained challenging for several years even after cryo-EM structure determination of large assemblies had become well established.[47] This review aims to highlight the interplay between technological advances required to address this challenge and the biological insight enabled by them using the example of the human CAK. I will summarize recent efforts to determine the structure of the CAK using cryo-EM and discuss the challenges associated with high-resolution, high-throughput structure determination of small complexes using this method.

Early breakthroughs in cryo-EM of small protein complexes

The specimens resolved to high resolution during the early phase of the "resolution revolution" were often large macromolecular

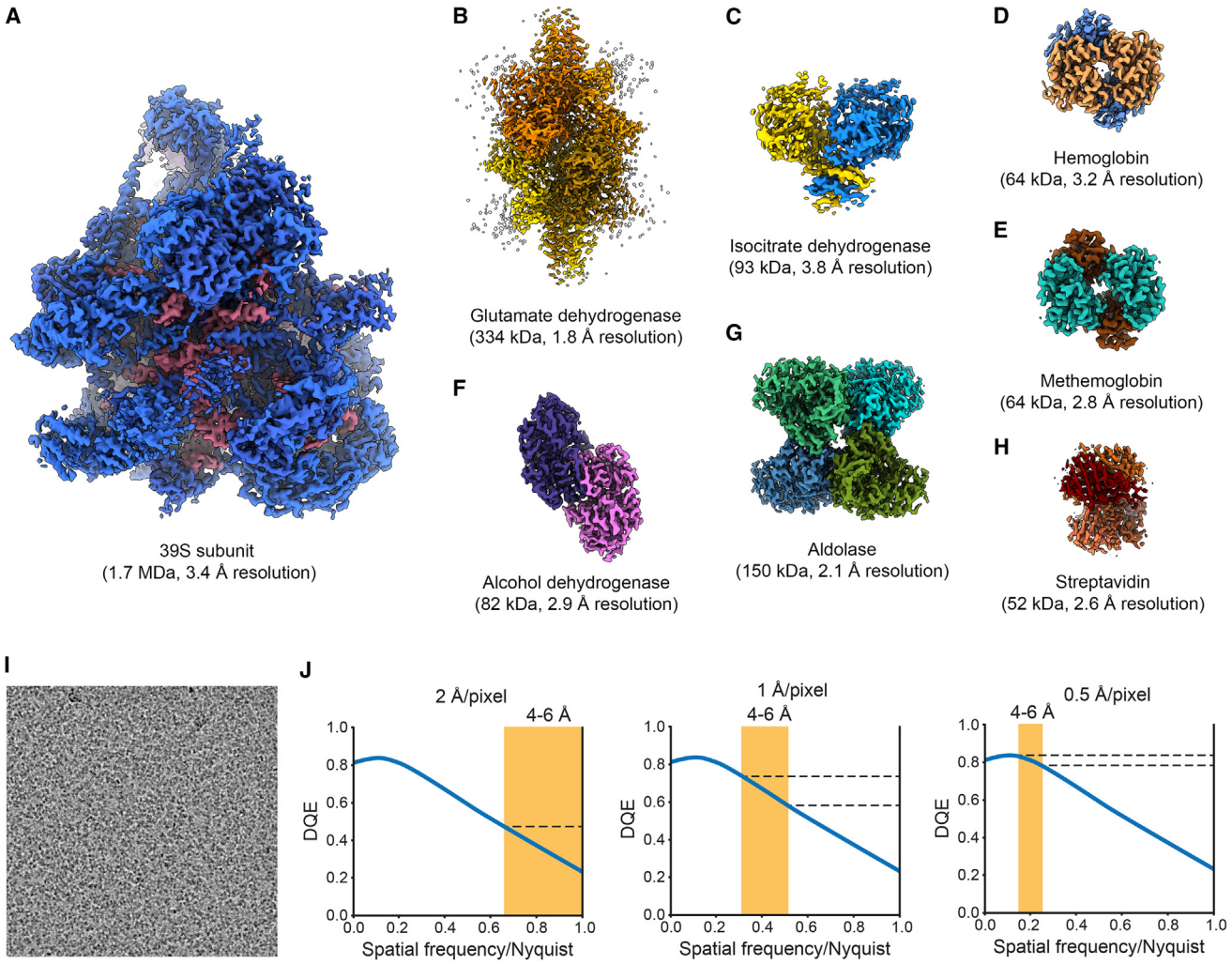

Figure 2. Cryo-EM structure determination of small protein complexes and considerations guiding data collection

(A) The 39S mitoribosomal subunit, a typical representative of early high-resolution cryo-EM structures, depicted for size comparison (EMD-2787, FEI Falcon II, 300 kV, C_1 symmetry).[103]

(B) Glutamate dehydrogenase (EMD-8194, Gatan K2 Summit, 300 kV, D_3 symmetry).[56]

(C) Isocitrate dehydrogenase (EMD-8192, Gatan K2 Summit, 300 kV, C_2 symmetry).[56]

(D) Hemoglobin (EMD-3488, Gatan K2 Summit, 300 kV and VPP, C_2 symmetry).[58]

(E) Methemoglobin (EMD-0407, Gatan K2 Summit, 200 kV, C_2 symmetry).[63]

(F) Alcohol dehydrogenase (EMD-0406, Gatan K2 Summit, 200 kV, C_2 symmetry).[63]

(G) Aldolase (EMD-21023, Gatan K2 Summit, 200 kV, D_2 symmetry).[66]

(H) Streptavidin (EMD-20907, Gatan K2 Summit, 200 kV, D_2 symmetry).[68]

(I) High-magnification electron micrograph showing a high density of CAK particles (approx. 600 particles). Sample micrograph acquired at a pixel size of 0.57 Å, an acceleration voltage of 200 kV, 1.5 μm defocus, and an exposure of 60 electrons/Å^2.

(J) Schematic illustration of the effect of using different pixel sizes on the DQE in the frequency band corresponding to 4–6 Å-resolution information (the DQE curve approximates the performance of electron counting direct detectors[55]). With increasing magnification and decreasing pixel size, frequency bands migrate to higher DQE.

complexes of several hundred kDa or several MDa molecular weight, such as ribosomes, spliceosomal complexes, transcriptional assemblies, or multimeric membrane channels, which, in contrast to small particles (<100 kDa), provide ample signal to drive accurate particle image alignment (Figure 2A).[48–54] Much of this early high-resolution cryo-EM work was accomplished using high-end electron microscopes equipped with highly coherent 300 kV field emission gun (FEG) electron sources and one of two camera models: The Gatan K2 Summit, an electron counting direct detector, and the FEI Falcon II, an integrating direct detector.[55] Both detectors were capable of dose fractionation, collecting a "movie" rather than a single micrograph during each exposure to enable correction for beam-induced motion, and featured strongly improved detective quantum efficiency (DQE) compared to previously used CCD cameras or electron microscopic film.[55] However, even early counting direct detectors such as the Gatan K2 Summit far surpassed integrating direct detectors in the low-frequency bands that aid particle image alignment,[55] providing them with a decisive advantage for work with small macromolecules or their complexes.

In 2016, this detector technology was used to determine the structures of the 334 kDa glutamate dehydrogenase complex at 1.8 Å resolution (Figure 2B) and the 93 kDa isocitrate dehydrogenase dimer at 3.8 Å resolution (Figure 2C).[56] This work was conducted using a 300 kV Thermo Fisher Scientific (TFS) Titan Krios microscope and a Gatan K2 Summit camera coupled with a BioQuantum energy filter and demonstrated that the sub-100 kDa molecular weight range was now amenable for cryo-EM structure determination at resolutions permitting building of atomic models, with a clear path toward applications such as structure-based drug design.

While the aforementioned study had used a traditional defocus-based approach to generate phase contrast in the electron micrographs,[56] alternatives to increase low-resolution contrast while maintaining minimal defocus were explored in order to facilitate work with small particles, which otherwise provide comparably little total signal for particle identification and image alignment.[57] Most notably, the Volta Phase Plate (VPP) was used to visualize the 64 kDa hemoglobin tetramer at 3.2 Å resolution (Figure 2D), thereby pushing both the molecular weight boundary and the resolution achieved for such a small complex into previously uncharted territory.[58] However, the VPP approach, which relies on a thin carbon film placed in the back focal plane of the objective lens to generate a phase shift and increase low-resolution contrast,[59] was later shown to have substantial drawbacks, including reduced high-resolution signal,[60,61] bringing conventional defocus-based approaches back into focus. It is worth noting that the more advanced phase plate concepts, such as the laser phase plate, are currently being explored in the hope to exploit the advantages of contrast enhancement by the phase plate without incurring the drawbacks of the VPP.[62]

In 2018, high-resolution structures of the 64 kDa methemoglobin tetramer at 2.8–3.2 Å resolution (Figure 2E) and the 82 kDa alcohol dehydrogenase dimer at 2.9 Å resolution (Figure 2F) were determined using data acquired on a TFS Talos Arctica—a microscope with a 200 kV FEG initially designed for sample screening purposes—and a Gatan K2 Summit camera.[63] The approach used in this study involved high magnification (resulting in approximately 0.6 Å pixel size) to ensure optimal data quality (see the following text), and the use of very low defocus in the absence of a phase plate. With thin specimens of small complexes, defocus values as low as −0.5 μm produced interpretable data,[63,64] contrary to previous assumptions that large defocus values of several micrometres were required to see small particles in vitreous ice.[65] Subsequent work achieved 2.1 Å resolution for the 150 kDa aldolase tetramer (Figure 2G) and demonstrated the feasibility of sub-2 Å resolution reconstructions of apoferritin using a 200 kV microscope, encouraging the exploration of these lower-cost and more easily accessible systems for high-resolution applications.[66] Sub-2 Å resolution from a 200 kV microscope was also reported using a β-galactosidase specimen and a JEOL CRYO ARM 200 microscope with an in-column energy filter,[67] underscoring the ability of this class of system to produce very high-quality data.

The size limit for high-resolution cryo-EM reconstructions of small protein complexes was later further reduced to 52 kDa by structure determination of the streptavidin tetramer on a graphene film at 2.6 Å resolution (Figure 2H).[68] These pioneering developments using well-behaved and symmetrical test specimens indicated a clear path for application of high-resolution cryo-EM to medically highly relevant small protein targets such as the human CAK, which will be described in the following sections. In parallel to our efforts on CAK, several other sub-100 kDa structures and sub-structures were determined by cryo-EM, as reviewed in a study by Lander et al.,[47] and sub-2.5 Å resolution could be routinely achieved for G-protein coupled receptors (GPCRs), a family of small (~150 kDa) membrane protein complexes.[61] Notably, this latter study systematically investigated the effects of electron exposure, defocus range, use of an energy filter (zero-loss filtering), use of an objective lens aperture, use of a VPP, and use of gold foil grids, and found substantial benefits when using an energy filter, gold grids, and higher electron exposure, and reduced performance when using the VPP.[61]

Structure determination of the human CAK by cryo-EM

Encouraged by the breakthrough results described previously, we decided to apply the technical insights provided by these studies to the high-resolution structure determination of the human CAK. The key parameters we aimed to adapt from these studies were a high protein concentration providing a near-continuous lawn of particles in the micrographs (Figure 2I), the use of high magnification to optimize the DQE in the frequency bands most important for image alignment and reconstruction (Figure 2J),[55,63] and the use of gold foil grids,[69] as pioneered by earlier work. High particle density not only provides large particle numbers in a limited number of micrographs but also appears to correlate with better stability of the thin ice layer required to image particles of this size, which tends to break more readily when particles are sparse. Furthermore, high particle coverage increases the signal required to accurately fit contrast transfer function (CTF) parameters. The use of high magnification shifts the frequency bands that are most important for high-accuracy particle alignment and high-resolution 3D reconstruction toward higher DQE (Figure 2J).[55,63,66] This is beneficial for high-resolution structure determination, and the apparent drawback of collection of fewer particle images per micrograph is partially offset by the shorter exposure times per movie when maintaining a constant electron flux per pixel, which is required for efficient electron counting. To maximize data quality by reducing beam-induced motion, we employed gold foil grids,[69] which have shown to improve data quality in systematic tests of "real world" applications[61] and in our hands had the added benefit of providing the required thin ice films more readily.

Applying these considerations to CAK sample preparation and data collection strategies enabled progress toward high resolution. However, overcoming ice breakage, uneven particle distribution, and preferred orientation required extensive grid screening to identify the two high-quality cryo-EM grids from which we determined the structure of the human CAK in complex with ATPγS at 2.8 Å resolution (Figure 3A).[41] Inspired by earlier work using 200 kV instruments for high-resolution structure determination of small complexes,[63,64] this work was performed using a 200 kV TFS Talos Arctica electron microscope equipped with a Gatan K3 electron counting detector. Image processing for these data was comparably straightforward. Notably, the alignment-free classification algorithm implemented in RELION[70] proved to be instrumental in extracting the highest-quality particles from very large datasets (Table 1), thereby improving the quality of the resulting maps substantially.

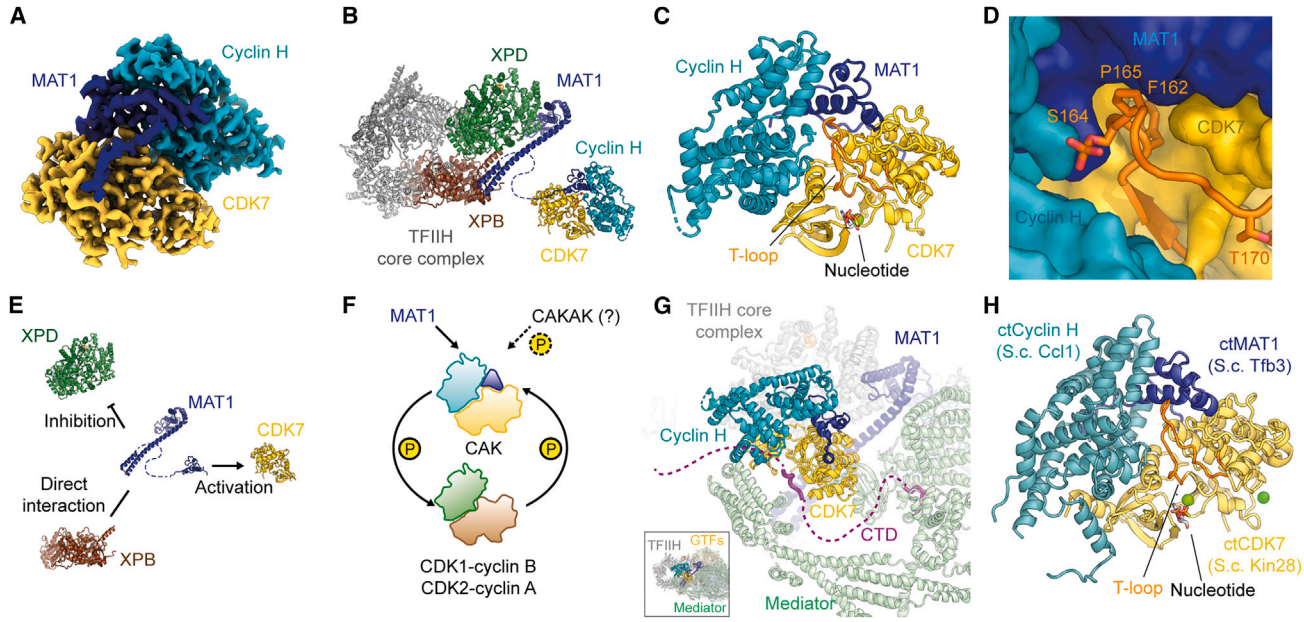

Figure 3. Cryo-EM structure of the human and fungal CAKs

(A) Cryo-EM map of the human CAK at 2.8 Å resolution (EMD-22123),[41] colored by its constituent subunits (CDK7 gold, cyclin H cyan, MAT1 blue). MAT1 is sandwiched between the CDK and the cyclin and stabilizes the assembly.

(B) Structure of holo-TFIIH with the TFIIH core complex (gray; helicases XPD and XPB in green and brown) and the CAK subcomplex in colors as above (PDB IDs 6XBZ, 6NMI).[41,104] MAT1 links the two subcomplexes.

(C) Atomic model of the human CAK (PDB 6XBZ). Its constituent subunits and the locations of the T-loop (activation segment, orange) and the nucleotide binding site are indicated. Note the extended state of the T-loop.

(D) Contacts between MAT1 and the tip of CDK7 T-loop contribute to an activated state of CDK7. The tip of the CDK7 T-loop harbors a second T-loop phosphorylation site at S164, in addition to the conserved CDK T-loop phosphorylation site at CDK7 T170 (which was not phosphorylated in the cryo-EM structure).

(E) Network of MAT1 interactions with the functional centers of TFIIH.

(F) Dual mechanisms of CAK activation by MAT1 and T-loop phosphorylation. CDK7 and CDK2/CDK1-cyclin complexes can phosphorylate and activate each other. CDK7 activation toward CDK substrates by MAT1 is phosphorylation-independent. The possible existence of additional, currently unidentified CAKAKs is indicated with a dashed arrow.

(G) The structure of the human transcription Pol II pre-initiation complex including mediator with a bound Pol II-CTD segment near the CAK active site (PDB ID 7LBM).[81]

(H) X-ray crystal structure of the *C. thermophilum* CAK/TFIIK complex (PDB ID 6Z4X).[40] Names of the homologous budding yeast proteins are given in parentheses.

Structural insight into the function and regulation of the human CAK

The resulting structure revealed the architecture of the human CAK, with a canonical CDK-cyclin pair formed by CDK7 and cyclin H, and with the C-terminal domain of the unique CAK subunit MAT1 sandwiched in between the two aforementioned components (Figure 3A).[41] In this position, MAT1 interacts with both CDK7 and cyclin H and more than doubles the total interaction area in the complex, explaining its role as a CAK assembly factor.[15] The C-terminal domain observed in the CAK structure is tethered to the N-terminal, helical domain of MAT1 via a flexible linker, providing flexible attachment to the TFIIH core complex when CAK engages in its transcriptional role (Figure 3B). In this position, MAT1 interacts with CDK7 as well as the TFIIH helicases XPD and XPB, providing a means for coordination of the activity of these functional centers of TFIIH.

In addition to the known function of MAT1 in XPD repression during transcription initiation,[71–74] the CAK structure revealed that MAT1 forms interactions with the tip of the regulatory T-loop of CDK7 that stabilize an active conformation of the kinase (Figures 3C and 3D). This reinforces the notion of a regulatory role, rather than a purely structural function, for MAT1 (Figure 3E), and it explains how MAT1 promotes the activity of CDK7 even when its two T-loop phosphorylation sites (CDK7 S164 and T170) are not phosphorylated.[15,19,73] More recent work indicates that phosphorylation and MAT1 binding act synergistically to activate CDK7 toward transcriptional substrates and confirms that activity toward CDK substrates is relatively unaffected by phosphorylation.[28] While the mechanistic basis of these differential requirements for activity toward different substrates is currently not clear, our cryo-EM reconstruction[41] and a recent X-ray crystal structure of the human CAK[28] both show that phosphorylated S164 participates in an interaction network involving residues from cyclin H and MAT1 (Figure 3D), thereby providing a rationale for how S164 phosphorylation may affect the interactions and conformation of the CDK7 T-loop.

The phosphorylation-independent mechanism of CDK7 activation toward CDK substrates by MAT1[19] ensures CAK activity throughout the cell cycle. This mechanism may be important because the CAK-activating kinases (CAKAKs) that can phosphorylate the CDK7 T-loop and further activate CAK *in vitro* and in cell extracts are CDK1 and CDK2, both of them CAK substrates themselves (Figure 3F).[75] Because CDK2 alone, rather than CDK2-cyclin complexes, is the preferred substrate of CAK,[76] and stable

Table 1. Comparison of hardware and dataset statistics for different CAK cryo-EM reconstructions discussed in this review

Structure	CAK-ATPγS	CAK-ICEC0942	CAK-ICEC0942
EMDB accession	EMD-22123	EMD-12042	EMD-17508
Reference	Ref.[41]	Ref.[42]	Ref.[43]
Resolution	2.8 Å	2.5 Å	1.9 Å
Microscope	TFS Talos Arctica	TFS Talos Arctica	TFS Titan Krios
FEG type	FEG	FEG	C-FEG
Voltage	200 kV	200 kV	300 kV
Detector	Gatan K3	Gatan K3	TFS Falcon 4i
Energy filter	–	–	TFS Selectris X
Movies (total collected)	6,147 (7,582)	6,985 (8,873)	3,885 (5,367)
Pixel size	0.686 Å	0.686 Å	0.57 Å
Particles picked	8,982,448	10,904,715	3,588,357[a]
Particles after 2D	6,588,462	9,339,254	2,072,743
Particles final	136,859	205,478	269,962

[a]Particles were picked using multiple methods and combined; this number refers to the largest number of particles picked by a single run using a single method.

assembly of CDK1-cyclin complexes appears to be coupled to phosphorylation,[21] CAK activity plays a key role in formation of active CDK1- and CDK2-cyclin complexes. In the absence of MAT1, this could potentially create a molecular chicken-and-egg problem precluding appropriate cell cycle CDK phosphorylation and activation (Figure 3F), though the contribution of CDK1 and CDK2 to CAK activation *in vivo* has not been fully clarified and the existence of additional, mitogen-dependent CAKAKs acting early in the cell cycle has been proposed.[20,28] Furthermore, phosphorylation-independent activation of CAK by MAT1 may ensure that sufficient CAK activity is available to activate its substrates in non-dividing cells, such as CDK9 and CDK11,[9,11,19] which function in transcription and splicing.

Structural insight into the molecular mechanism of the kinase function of CAK in transcription initiation, which is the second major role of the human CAK, came from work by several laboratories that resolved structures of RNA polymerase II pre-initiation complexes with and without mediator, in which low-resolution densities for the CAK or its fungal homologs (see the following text) could be visualized in the vicinity of TFIIH and RNA polymerase II[77,78] or bound to mediator.[79,80] Subsequent higher-resolution cryo-EM structures of the RNA polymerase II pre-initiation complex associated with TFIID, mediator, or +1 nucleosomes permitted unambiguous docking of the CAK structure into the density.[81–85] In these structures, the CAK was again localized in different positions, depending on the combination of the co-activators TFIID and mediator present. In some instances, density for the Pol II CTD near the CDK7 active site could be visualized in the cryo-EM maps (Figure 3G).[81,84] These studies therefore revealed composition-dependent plasticity in CAK positioning within the pre-initiation complex and explained how mediator aids in Pol II-CTD phosphorylation by providing a structural framework that brings the enzyme and its substrates into proximity (Figure 3G).

Comparison to fungal structures

In the budding yeast *Saccharomyces cerevisiae*, the CAK—i.e., the enzyme responsible for CDK activation—is a monomeric kinase, called Cak1/Civ1[86–88] and the trimeric complex that is homologous to the human CAK is known as transcription factor IIK (TFIIK) because it shares only the transcriptional activity but not the CAK function with the human CAK.[89,90] While metazoan and budding yeast cells employ entirely distinct CAKs, both the CDK7 homolog Mcs6 and the Cak1 homolog Csk1 function as partially redundant CAKs in the fission yeast *Schizosaccharomyces pombe*, which thus occupies an intermediate state in its CDK activation mechanism.[91,92] Structures of *Chaetomium thermophilum* CAK (ctCDK7-ctCyclinH-ctMAT1) and budding yeast TFIIK (Kin28-Ccl1-Tfb3) published around the same time as our initial structure of the human CAK confirmed that the overall architecture and assembly principles of human CAK and its fungal homolog are conserved.[40,44]

The X-ray crystal structures of the *C. thermophilum* ctCDK7-ctCyclinH-ctMAT1 complex in its free and ATPγS-bound forms (Figure 3H) were resolved at 2.6–3.0 Å, similar to the human cryo-EM structure.[40] This study strongly supported the notion that CDK7/Kin28 is activated by T-loop interactions formed with MAT1/Tfb3, suggesting conservation of this activatory mechanism between the human and fungal complexes, and suggested that the conformation of the base of the adenine nucleotide in the Kin28 active site is different from the human CDK7 complex.[40] A subsequent cryo-EM reconstruction of yeast TFIIK (Kin28-Ccl1-Tfb3) was resolved to 3.5 Å, confirming the conserved architecture of the complex and providing a basis for initial integrative modeling of the position of TFIIK near mediator in the RNA polymerase II pre-initiation complex, but precluding detailed interpretation of ligand interactions.[44]

CAK interactions with small-molecule inhibitors

Given our interest in visualizing small-molecule interactions with the human CAK, we initially determined the structure of a covalent adduct between CAK and the inhibitor THZ1 at 3.3 Å resolution, along with the nucleotide-bound structure described previously (Figures 4A–4C).[41] THZ1 is a covalent CAK inhibitor that was shown to strongly suppress super-enhancer direct transcription, to which many cancer cell lines are addicted.[93] THZ1 was initially reported to be highly selective for CDK7 based on its ability to modify the non-conserved C312 near the active site,[93] but was later found to also modify and inhibit CDK12 and CDK13, which also harbor active site-proximal cysteine residues.[94] Our CAK-THZ1 structure showed the inhibitor head group bound in the active site pocket where it contacts the CDK7 "hinge" region and prevents adenosine nucleotide binding (Figure 4B). The warhead forming a covalent adduct with CDK7 C312 was not clearly resolved in the density due to flexibility, though approximate docking based on weak density was possible. Flexibility of the warhead portion of THZ1 was expected, considering that the related inhibitor THZ531 bound to CDK12 shows two entirely different warhead conformations in the X-ray crystal structure of this complex.[95]

We subsequently also determined the structure of the human CAK bound to the clinically developed inhibitor ICEC0942 (Figures 4D and 4E).[42] This compound is a non-covalent high-affinity competitive inhibitor binding the CDK7 active site.[31] The CAK-ICEC0942 structure was again determined at 200 kV using

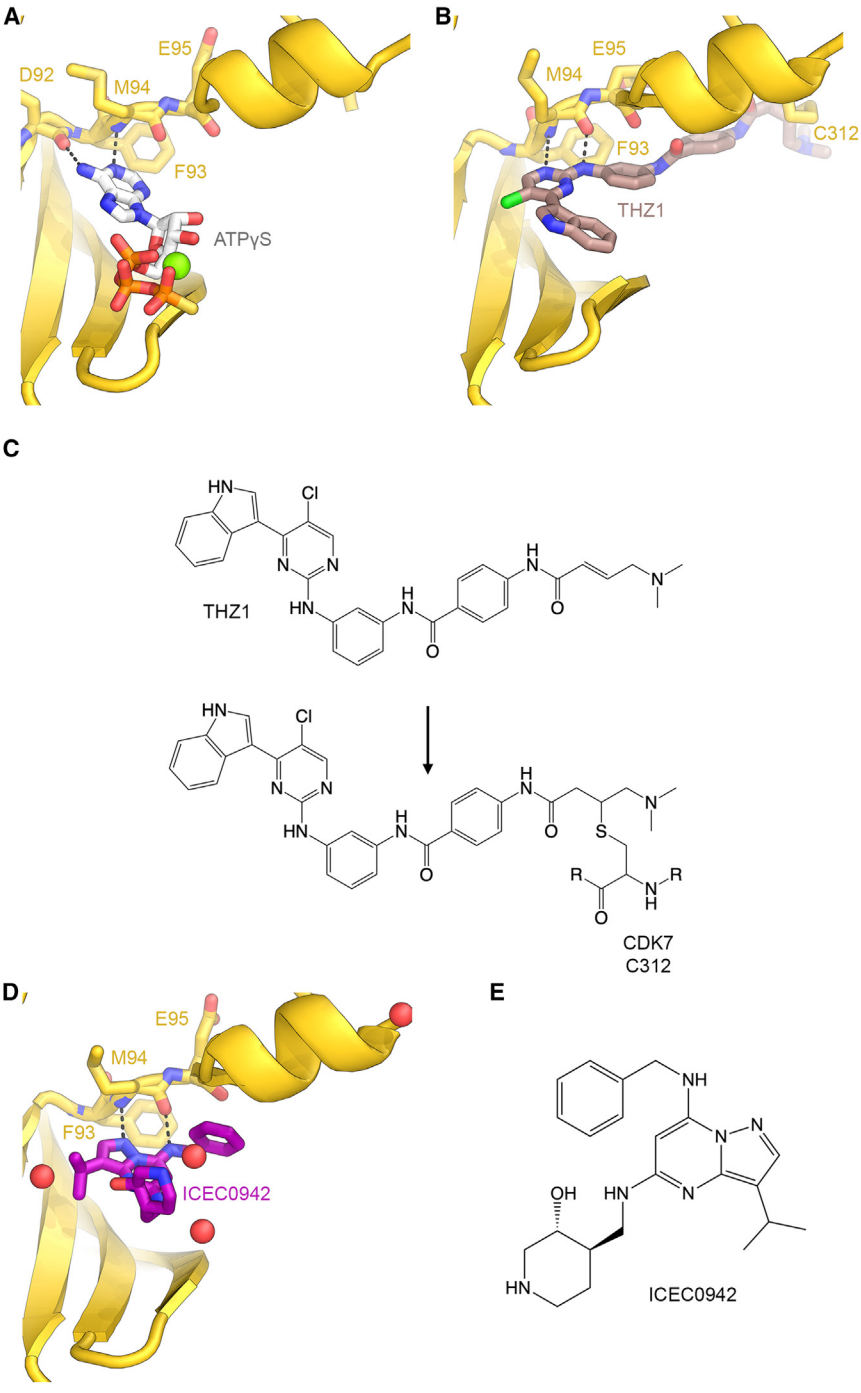

Figure 4. CAK-ligand complexes at 2.5–3.3 Å resolution

(A) Structure of CAK-ATPγS. Residues in the CDK7 hinge region are shown as sticks and numbered. Hydrogen bonds to the hinge region are indicated.

(B) Structure of the covalent CAK-THZ1 adduct. The ligand exploits hydrogen bonding to the hinge for initial binding and a covalent bond to CDK7 C312 for irreversible modification of the target kinase.

(C) Chemical structure of THZ1 and its action by formation of a covalent cysteine adduct.

(D) Structure of CAK-ICEC0942. Even though their chemical core is different from THZ1, pyrazolopyrimidines like ICEC0942 exploit similar hinge contacts with the backbone atoms of CDK7 M94.

(E) Chemical structure of ICEC0942.

a Talos Arctica and a Gatan K3 camera, resulting in a 2.5 Å-resolution map (Table 1, Figures 5A, and 5B). This structure provided us with important insight into inhibitor selectivity. Specifically, we found that the inhibitor pose between the CDK7-bound and the CDK2-bound structures (ICEC0942 target and off-target structures, respectively) is different (Figures 5C–5E). This conformational difference likely originates from differences in the molecular environment provided by the two CDK active sites, with slight differences in the conformation of the hinge region (green in Figures 5C and 5D) and amino acid substitutions in the β-strand adjacent to the inhibitor playing possible roles. These observations suggested possible avenues toward structure-based drug design by restricting the conformational space of the chemical groups that occupy different positions.[42] However, at the given resolution, there were also some ambiguities regarding the positioning of a hydroxypiperidine group at the other end of the inhibitor, suggesting that even higher resolution (see the following text and Figures 5F and 5G) would be beneficial to improve the accuracy of these drug-binding studies.

Breaking the 2 Å-barrier: Technical considerations

Detailed analysis of inhibitor interactions and accurate placement of water molecules are facilitated by high resolution data, preferably extending to beyond 2 Å resolution. Such substantial resolution gain often necessitates the use of improved instrumentation. Considering that only very few structures of small protein complexes to better than 2 Å resolution had been determined at the time, use of a high-end instrument was likely required to obtain data of sufficient quality to meet this objective. The successful 2.1 Å-resolution structure determination of the SARS-CoV-2 ORF3a dimer (Figure 5H), a 64 kDa membrane protein complex,[96] using a 300 kV TFS Krios microscope equipped with a cold field emission gun (C-FEG), a Selectris X energy filter, and a Falcon 4i direct detector[97] suggested that this setup offered capabilities suitable for the study of ligand-bound CAK complexes at high resolution. This high-end 300 kV system promised several benefits over the 200 kV Talos Arctica used for the initial structure determinations of CAK, including the highly coherent C-FEG electron source that features a lower energy spread compared to conventional

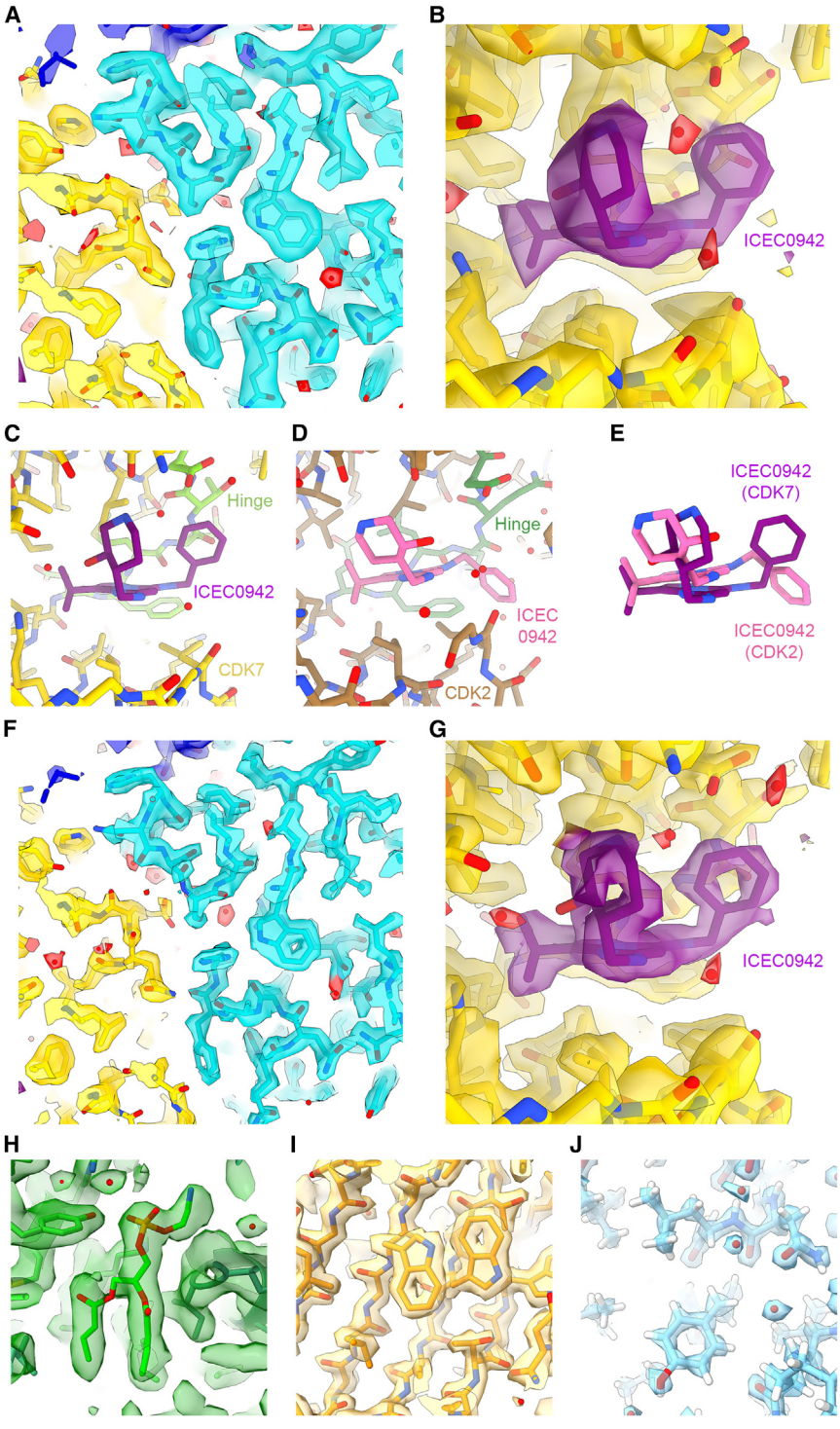

Figure 5. Comparisons of 2.5 Å and 1.9 Å-resolution CAK-ICEC0942 structures, comparison with CDK2-bound ICEC0942, and early high-resolution structures from an energy filtered 300 kV instrument with a cold-FEG

(A) Interface between CDK7 (gold) and cyclin H (cyan) at 2.5 Å resolution; water molecules are shown in red (TFS Talos Arctica at 200 kV, Gatan K3).[42]

(B) ICEC0942 ligand density at 2.5 Å resolution (TFS Talos Arctica at 200 kV, Gatan K3).

(C) Rendering of ICEC0942 bound to the CDK7 active site.[43] The hinge region connecting the N- and C-terminal kinase lobes is colored light green.

(D) Rendering of ICEC0942 bound to the CDK2 active site. The hinge region is colored green.[99]

(E) Superposition of the two ICEC0942 poses observed when bound to CDK7 (purple) and CDK2 (pink).

(F) Same view as (A), at 1.9 Å resolution (TFS Krios G4 at 300 kV, Selectris X, Falcon 4i).[43]

(G) Same view as (B), at 1.9 Å resolution (TFS Krios G4 at 300 kV, Selectris X, Falcon 4i).

(H) Head group of a bound phospholipid in the 2.1 Å-resolution structure of the SARS-CoV-2 ORF3a membrane protein complex (EMD-22898, PDB ID 7KJR).[96]

(I) Structural detail of streptavidin at 1.7 Å resolution (EMD-31083, PDB ID 7EFC).

(J) Atomic-resolution (~1.2 Å) cryo-EM structure of apo-ferritin (EMD-11638, PDB ID 7A4M).[97]

FEGs,[97] the highly stable energy filter allowing energy filtration with a slit width of only 5–10 eV,[97] and the DQE advantage of contemporary direct electron detectors (including the Gatan K3 and TFS Falcon 4i) when used at 300 kV as compared to 200 kV. Notably, in addition to the SARS-CoV-2 ORF3a structure mentioned previously, this instrument had been used to achieve 1.7 Å resolution for streptavidin (EMD-31083, Figure 5I) and true atomic resolution (~1.2 Å) for the highly symmetrical test specimen apo-ferritin (Figure 5J), representing a technical milestone for single-particle cryo-EM.[97]

The use of energy filters to remove inelastically scattered electrons is essential in cryogenic electron tomography (cryo-ET) due to the thickness of tomographic specimens. However, energy filters are also commonly employed in single particle cryo-EM work, where specimens are typically much thinner than in cryo-ET and inelastic scattering is less severe. Nevertheless, several studies found improved resolution and reduced B-factors when using Gatan K3/BioQuantum and TFS Falcon 4i/Selectris X camera/energy filter combinations for single-particle cryo-EM with several GPCR specimens,[61] the human β3 GABAA receptor homopentamer in complex with a megabody,[97] and apo-ferritin, though in the latter case only for ice thickness >1000 Å.[98] Apo-ferritin at 440 kDa and the GABAA receptor-megabody complex at almost 500 kDa are substantially larger than CAK while the GPCR samples at approx. 150 kDa are in a roughly similar size range. A direct comparison between zero-loss filtered and unfiltered CAK datasets to investigate the possible benefits of energy filtration for structure determination of sub-100 kDa specimens showed that the energy filter improved data quality and enabled higher resolution.[43] This reinforces the notion that energy filters can have

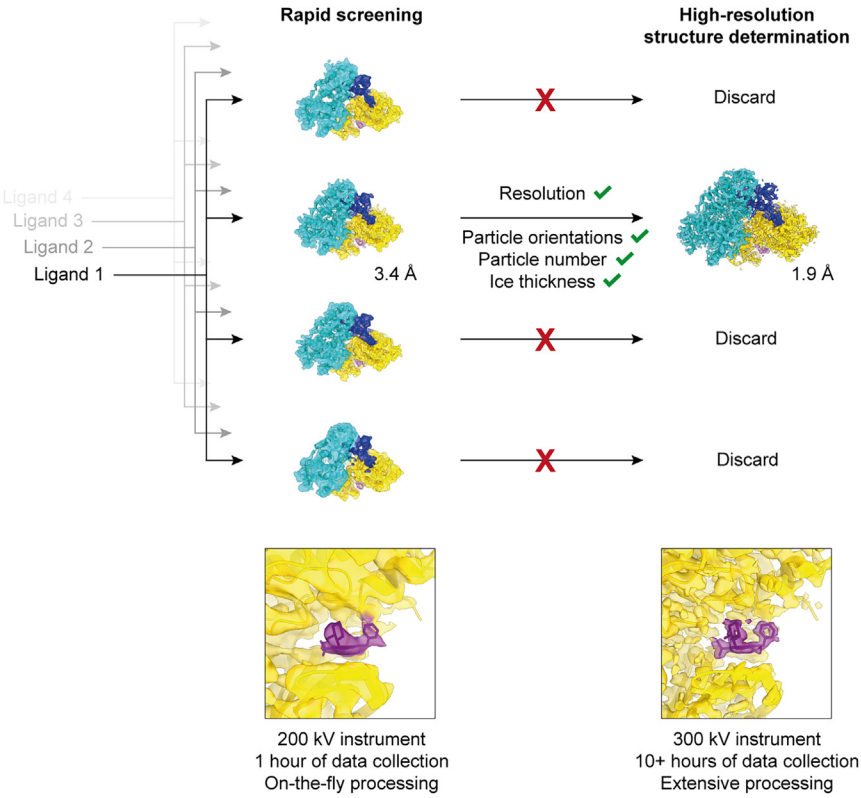

Figure 6. High-throughput screening
High-throughput screening workflow with initial rapid grid screening followed by high-resolution structure determination on selected specimens. This approach can be scaled to larger numbers of compounds.

appreciable benefits for high-resolution single particle work, in addition to their well-established application to comparably thick samples in cryo-ET, and that these benefits can be readily observed and exploited with small specimens in comparably thin ice.

Sub-2 Å-resolution structures of the human CAK

A pilot experiment using the microscope setup described previously and the CAK-THZ1 preparation that had resulted in a 3.3 Å-structure previously (see aforementioned text) yielded 1.9 Å resolution from approximately 10,000 micrographs.[43] Notably, this experiment used CAK harboring a truncated form of MAT1, resulting in an approx. 85 kDa CAK core module[41] and thus represents a bona fide sub-100 kDa structure.

Subsequent structure determinations of the CAK in complex with ICEC0942 and related pyrazolopyrimidine-type compounds from roughly 5,000 micrographs each routinely yielded structures at 2 Å resolution or better (Table 1, Figures 5F and 5G), resolving side chain conformations, water molecules, and post-translational modifications of the proteins. These results obtained using cryo-EM are on par with the 2.15 Å resolution of a recent X-ray crystal structure of the human CAK in complex with a nanobody,[28] in which a part of the CDK7 active site is engaged in a crystal contact. Notably, the 1.9 Å-resolution cryo-EM structure of CAK-ICEC0942 resolved the ambiguity in the conformation of the hydroxypiperidine group and revealed the involvement of this group in hydrogen bonding networks with bound water molecules, thereby rationalizing the known contribution of this group to inhibitor selectivity, which is essential for the possible application of the compound as a cancer therapeutic. The precise positioning of this hydroxy group appears to be critical, considering that ICEC0943, the enantiomer of ICEC0942, is far less selective, binds CDK7 far less tightly, and shows substantially increased conformational variability when bound to the CDK7 active site.[43,99]

Increased throughput for ligand screening

The aforementioned experiments demonstrated the feasibility of routine near-atomic resolution structure determination of the human CAK with approximately 10 h of data collection time on a high-end instrument. For structure-based drug design workflows, such high resolution is important, but so is high throughput in order to enable rapid structure determination of large numbers of candidate compounds in complex with their targets and off-targets. Notably, there is a trade-off between throughput and resolution, and workflow design requires optimization of time allocation to balance these two requirements. Furthermore, specimens on grids that are ill-suited to provide high-resolution data, e.g., due to poor ice quality, low particle numbers, or preferred orientation (all of which are properties that may vary from grid to grid) need to be eliminated from the data collection pipeline as early as possible to avoid wasting high-end instrument time. Ideally, these initial screening steps are performed on more readily accessible, less costly instruments.

To enable efficient high-throughput compound screening workflows, we therefore used a 200 kV TFS Glacios-type instrument with an energy filter to rapidly screen multiple specimens of individual inhibitor complexes to identify promising grids and exploit high-end 300 kV instrument time optimally (Figure 6). Using on-the-fly processing,[100] these workflows allowed us to obtain reconstructions of ligand-bound CAK at 3–4 Å resolution within 1 h of the start of the data collection with a high degree of automation using multi-grid functionality in the data collection software.[43] Notably, these reconstructions not only provided information on ice thickness, particle orientation distribution, and particle coverage, but additionally were of sufficient quality to identify the presence of bound ligands. Grids identified in such rapid screening workflows were then progressed to high-end data collection for high-resolution structure determination.

We anticipate that these high-resolution/high-throughput cryo-EM applications and the instrumentation that enables them will become more widely used in the future, and that they will enable structure-based drug design for previously intractable targets. Other structural biology methods are routinely used for

structure determination of low-affinity binders, such as initial lead compounds or during fragment screening in drug discovery pipelines, and methods that allow the exploitation of cryo-EM for fragment screening are beginning to emerge.[101] Such developments will be further facilitated by higher throughput and higher resolution achieved using the latest cryo-EM instrumentation.

Future perspectives

I hope that this review has illustrated how methodological developments and biological discovery in the field of small-complex cryo-EM are tightly intertwined. Now that the molecular structure of the human CAK has been described in great detail, further biological insights will come from the analysis of structures of CAK caught in the act of phosphorylating its substrates to elucidate the molecular basis of the substrate specificity of CDK7. Future technical development will need to be directed toward increasing reproducibility of sample preparation, which will be critical for the screening of larger numbers of compounds. Additionally, increasing high-resolution data collection and data processing throughput will mitigate existing bottlenecks on the way to high-resolution, high-throughput cryo-EM structure determination of small, ligand-bound complexes. This will facilitate the use of cryo-EM in structure-based drug design workflows.

ACKNOWLEDGMENTS

The author is indebted to Eva Nogales for her generous support of the early stages of the CAK structure determination project, to Abhay Kotecha and his team for their generous support of the latest high-resolution work, and to Simak Ali for his contributions to the study of ligand-bound CAK complexes. The author thanks Claudio Alfieri and Victoria Cushing for comments on the manuscript. Figures were created using UCSF Chimera X[102] and PyMol (The PyMol Molecular Graphics System, Version 2.5 Schrodinger LLC). B.J.G. was supported by a career development award from the Medical Research Council United Kingdom (grant number MR/V009354/1).

AUTHOR CONTRIBUTIONS

B.J.G. wrote the manuscript and prepared the figures.

DECLARATION OF INTERESTS

The author declares no competing interests.

REFERENCES

1. Morgan, D.O. (1995). Principles of CDK regulation. Nature *374*, 131–134.
2. Malumbres, M. (2014). Cyclin-dependent kinases. Genome Biol. *15*, 122. https://doi.org/10.1186/gb4184.
3. Schulze-Gahmen, U., De Bondt, H.L., and Kim, S.H. (1996). High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. J. Med. Chem. *39*, 4540–4546. https://doi.org/10.1021/jm960402a.
4. Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J., and Pavletich, N.P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature *376*, 313–320.
5. Russo, A.A., Jeffrey, P.D., and Pavletich, N.P. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. Nat. Struct. Biol. *3*, 696–700.
6. Bloom, J., and Cross, F.R. (2007). Multiple levels of cyclin specificity in cell-cycle control. Nat. Rev. Mol. Cell Biol. *8*, 149–160. https://doi.org/10.1038/nrm2105.
7. Fisher, R.P. (2005). Secrets of a double agent: CDK7 in cell-cycle control and transcription. J. Cell Sci. *118*, 5171–5180.
8. Bösken, C.A., Farnung, L., Hintermair, C., Merzel Schachter, M., Vogel-Bachmayr, K., Blazek, D., Anand, K., Fisher, R.P., Eick, D., and Geyer, M. (2014). The structure and substrate specificity of human Cdk12/Cyclin K. Nat. Commun. *5*, 3505.
9. Larochelle, S., Amat, R., Glover-Cutter, K., Sansó, M., Zhang, C., Allen, J.J., Shokat, K.M., Bentley, D.L., and Fisher, R.P. (2012). Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. Nat. Struct. Mol. Biol. *19*, 1108–1115.
10. Baumli, S., Lolli, G., Lowe, E.D., Troiani, S., Rusconi, L., Bullock, A.N., Debreczeni, J.E., Knapp, S., and Johnson, L.N. (2008). The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. EMBO J. *27*, 1907–1918.
11. Rimel, J.K., Poss, Z.C., Erickson, B., Maas, Z.L., Ebmeier, C.C., Johnson, J.L., Decker, T.M., Yaron, T.M., Bradley, M.J., Hamman, K.B., et al. (2020). Selective inhibition of CDK7 reveals high-confidence targets and new models for TFIIH function in transcription. Genes Dev. *34*, 1452–1473. https://doi.org/10.1101/gad.341545.120.
12. Li, Y.-C., Chao, T.-C., Kim, H.J., Cholko, T., Chen, S.-F., Li, G., Snyder, L., Nakanishi, K., Chang, C.-E., Murakami, K., et al. (2021). Structure and noncanonical Cdk8 activation mechanism within an Argonaute-containing Mediator kinase module. Sci. Adv. *7*, eabd4484. https://doi.org/10.1126/sciadv.abd4484.
13. Fisher, R.P., and Morgan, D.O. (1994). A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. Cell *78*, 713–724.
14. Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O., and Reinberg, D. (1995). Cdk-activating kinase complex is a component of human transcription factor TFIIH. Nature *374*, 283–287.
15. Fisher, R.P., Jin, P., Chamberlin, H.M., and Morgan, D.O. (1995). Alternative mechanisms of CAK assembly require an assembly factor or an activating kinase. Cell *83*, 47–57.
16. Devault, A., Martinez, A.M., Fesquet, D., Labbé, J.C., Morin, N., Tassan, J.P., Nigg, E.A., Cavadore, J.C., and Dorée, M. (1995). MAT1 ('menage à trois') a new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and Xenopus CAK. EMBO J. *14*, 5027–5036.
17. Serizawa, H., Mäkelä, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A., and Young, R.A. (1995). Association of Cdk-activating kinase subunits with transcription factor TFIIH. Nature *374*, 280–282.
18. Larochelle, S., Pandur, J., Fisher, R.P., Salz, H.K., and Suter, B. (1998). Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. Genes Dev. *12*, 370–381.
19. Larochelle, S., Batliner, J., Gamble, M.J., Barboza, N.M., Kraybill, B.C., Blethrow, J.D., Shokat, K.M., and Fisher, R.P. (2006). Dichotomous but stringent substrate selection by the dual-function Cdk7 complex revealed by chemical genetics. Nat. Struct. Mol. Biol. *13*, 55–62.
20. Schachter, M.M., Merrick, K.A., Larochelle, S., Hirschi, A., Zhang, C., Shokat, K.M., Rubin, S.M., and Fisher, R.P. (2013). A Cdk7-Cdk4 T-loop phosphorylation cascade promotes G1 progression. Mol. Cell *50*, 250–260. https://doi.org/10.1016/j.molcel.2013.04.003.
21. Larochelle, S., Merrick, K.A., Terret, M.-E., Wohlbold, L., Barboza, N.M., Zhang, C., Shokat, K.M., Jallepalli, P.V., and Fisher, R.P. (2007). Requirements for Cdk7 in the assembly of Cdk1/cyclin B and activation of Cdk2 revealed by chemical genetics in human cells. Mol. Cell *25*, 839–850.
22. O'Gorman, W., Thomas, B., Kwek, K.Y., Furger, A., and Akoulitchev, A. (2005). Analysis of U1 small nuclear RNA interaction with cyclin H. J. Biol. Chem. *280*, 36920–36925. https://doi.org/10.1074/jbc.m505791200.
23. Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.M., and Chambon, P. (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. Cell *90*, 97–107.
24. Inamoto, S., Segil, N., Pan, Z.Q., Kimura, M., and Roeder, R.G. (1997). The cyclin-dependent kinase-activating kinase (CAK) assembly factor, MAT1, targets and enhances CAK activity on the POU domains of octamer transcription factors. J. Biol. Chem. *272*, 29852–29858.

25. Chen, D., Riedl, T., Washbrook, E., Pace, P.E., Coombes, R.C., Egly, J.M., and Ali, S. (2000). Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7. Mol. Cell *6*, 127–137.

26. Labbé, J.C., Martinez, A.M., Fesquet, D., Capony, J.P., Darbon, J.M., Derancourt, J., Devault, A., Morin, N., Cavadore, J.C., and Dorée, M. (1994). p40MO15 associates with a p36 subunit and requires both nuclear translocation and Thr176 phosphorylation to generate cdk-activating kinase activity in Xenopus oocytes. EMBO J. *13*, 5155–5164.

27. Larochelle, S., Chen, J., Knights, R., Pandur, J., Morcillo, P., Erdjument-Bromage, H., Tempst, P., Suter, B., and Fisher, R.P. (2001). T-loop phosphorylation stabilizes the CDK7-cyclin H-MAT1 complex in vivo and regulates its CTD kinase activity. EMBO J. *20*, 3749–3759.

28. Düster, R., Anand, K., Binder, S.C., Schmitz, M., Gatterdam, K., Fisher, R.P., and Geyer, M. (2024). Structural basis of Cdk7 activation by dual T-loop phosphorylation. Preprint at bioRxiv *1*. https://doi.org/10.1101/2024.1102.1114.580246.

29. Akoulitchev, S., and Reinberg, D. (1998). The molecular mechanism of mitotic inhibition of TFIIH is mediated by phosphorylation of CDK7. Genes Dev. *12*, 3541–3550. https://doi.org/10.1101/gad.12.22.3541.

30. Sava, G.P., Fan, H., Coombes, R.C., Buluwela, L., and Ali, S. (2020). CDK7 inhibitors as anticancer drugs. Cancer Metastasis Rev. *39*, 805–823. https://doi.org/10.1007/s10555-020-09885-8.

31. Patel, H., Periyasamy, M., Sava, G.P., Bondke, A., Slafer, B.W., Kroll, S.H.B., Barbazanges, M., Starkey, R., Ottaviani, S., Harrod, A., et al. (2018). ICEC0942, an Orally Bioavailable Selective Inhibitor of CDK7 for Cancer Treatment. Mol. Cancer Therapeut. *17*, 1156–1166.

32. Hu, S., Marineau, J.J., Rajagopal, N., Hamman, K.B., Choi, Y.J., Schmidt, D.R., Ke, N., Johannessen, L., Bradley, M.J., Orlando, D.A., et al. (2019). Discovery and Characterization of SY-1365, a Selective, Covalent Inhibitor of CDK7. Cancer Res. *79*, 3479–3491.

33. Marineau, J.J., Hamman, K.B., Hu, S., Alnemy, S., Mihalich, J., Kabro, A., Whitmore, K.M., Winter, D.K., Roy, S., Ciblat, S., et al. (2022). Discovery of SY-5609: A Selective, Noncovalent Inhibitor of CDK7. J. Med. Chem. *65*, 1458–1480. https://doi.org/10.1021/acs.jmedchem.1c01171.

34. Yu, D., Jeon, Y., Lee, S.-J., Kim, J., and Nam, K. (2022). Q901; a highly selective covalent cdk7 inhibitor inducing substantial anti-tumor effect in a broad spectrum of solid tumor lineages. Cancer Res. *82* (*Supplement*). Abstract nr 2574.

35. Martin, M.P., Endicott, J.A., and Noble, M.E.M. (2017). Structure-based discovery of cyclin-dependent protein kinase inhibitors. Essays Biochem. *61*, 439–452. https://doi.org/10.1042/EBC20170040.

36. Davies, T.G., Tunnah, P., Meijer, L., Marko, D., Eisenbrand, G., Endicott, J.A., and Noble, M.E. (2001). Inhibitor Binding to Active and Inactive CDK2 The Crystal Structure of CDK2-Cyclin A/Indirubin-5-Sulphonate. Structure *9*, 389–397. https://doi.org/10.1016/s0969-2126(01)00598-6.

37. Lolli, G., Lowe, E.D., Brown, N.R., and Johnson, L.N. (2004). The crystal structure of human CDK7 and its protein recognition properties. Structure *12*, 2067–2079.

38. Andersen, G., Poterszman, A., Egly, J.M., Moras, D., and Thierry, J.C. (1996). The crystal structure of human cyclin H. FEBS Lett. *397*, 65–69.

39. Kim, K.K., Chamberlin, H.M., Morgan, D.O., and Kim, S.H. (1996). Three-dimensional structure of human cyclin H, a positive regulator of the CDK-activating kinase. Nat. Struct. Biol. *3*, 849–855.

40. Peissert, S., Schlosser, A., Kendel, R., Kuper, J., and Kisker, C. (2020). Structural basis for CDK7 activation by MAT1 and Cyclin H. Proc. Natl. Acad. Sci. USA *117*, 26739–26748.

41. Greber, B.J., Perez-Bertoldi, J.M., Lim, K., Iavarone, A.T., Toso, D.B., and Nogales, E. (2020). The cryoelectron microscopy structure of the human CDK-activating kinase. Proc. Natl. Acad. Sci. USA *117*, 22849–22857. https://doi.org/10.1073/pnas.2009627117.

42. Greber, B.J., Remis, J., Ali, S., and Nogales, E. (2021). 2.5Å-resolution structure of the human CDK-activating kinase bound to the clinical inhibitor ICEC0942. Biophys. J. *120*, 677–686.

43. Cushing, V.I., Koh, A.F., Feng, J., Jurgaityte, K.T., Bondke, A., Kroll, S.H.B., Barbazanges, M., Scheiper, B., Bahl, A.K., Barrett, A.G.M., et al. (2024). High-resolution cryo-EM of the human CDK-activating kinase for structure-based drug design. Nat. Commun. *15*, 2265.

44. van Eeuwen, T., Li, T., Kim, H.J., Gorbea Colón, J.J., Parker, M.I., Dunbrack, R.L., Garcia, B.A., Tsai, K.-L., and Murakami, K. (2021). Structure of TFIIK for phosphorylation of CTD of RNA polymerase II. Sci. Adv. *7*, eabd4420. https://doi.org/10.1126/sciadv.abd4420.

45. Kühlbrandt, W. (2014). Biochemistry. The resolution revolution. Science *343*, 1443–1444.

46. Bai, X.-C., McMullan, G., and Scheres, S.H.W. (2015). How cryo-EM is revolutionizing structural biology. Trends Biochem. Sci. *40*, 49–57.

47. Lander, G.C., and Glaeser, R.M. (2021). Conquer by cryo-EM without physically dividing. Biochem. Soc. Trans. *49*, 2287–2298.

48. Greber, B.J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D., and Ban, N. (2015). The complete structure of the 55S mammalian mitochondrial ribosome. Science *348*, 303–308.

49. Liao, M., Cao, E., Julius, D., and Cheng, Y. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. Nature *504*, 107–112.

50. Amunts, A., Brown, A., Toots, J., Scheres, S.H.W., and Ramakrishnan, V. (2015). The structure of the human mitochondrial ribosome. Science *348*, 95–98.

51. He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. Nature *533*, 359–365.

52. Amunts, A., Brown, A., Bai, X.-C., Llácer, J.L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S.H.W., and Ramakrishnan, V. (2014). Structure of the yeast mitochondrial large ribosomal subunit. Science *343*, 1485–1489.

53. Fernández, I.S., Bai, X.-C., Murshudov, G., Scheres, S.H.W., and Ramakrishnan, V. (2014). Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. Cell *157*, 823–831.

54. Nguyen, T.H.D., Galej, W.P., Bai, X.-C., Oubridge, C., Newman, A.J., Scheres, S.H.W., and Nagai, K. (2016). Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. Nature *530*, 298–302.

55. McMullan, G., Faruqi, A.R., Clare, D., and Henderson, R. (2014). Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. Ultramicroscopy *147*, 156–163.

56. Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M.I., Pragani, R., Boxer, M.B., Earl, L.A., Milne, J.L.S., and Subramaniam, S. (2016). Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. Cell *165*, 1698–1707.

57. Danev, R., and Baumeister, W. (2017). Expanding the boundaries of cryo-EM with phase plates. Curr. Opin. Struct. Biol. *46*, 87–94.

58. Khoshouei, M., Radjainia, M., Baumeister, W., and Danev, R. (2017). Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. Nat. Commun. *8*, 16099.

59. Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J.M., and Baumeister, W. (2014). Volta potential phase plate for in-focus phase contrast transmission electron microscopy. Proc. Natl. Acad. Sci. USA *111*, 15635–15640.

60. Buijsse, B., Trompenaars, P., Altin, V., Danev, R., and Glaeser, R.M. (2020). Spectral DQE of the Volta phase plate. Ultramicroscopy *218*, 113079. https://doi.org/10.1016/j.ultramic.2020.113079.

61. Danev, R., Belousoff, M., Liang, Y.-L., Zhang, X., Eisenstein, F., Wootten, D., and Sexton, P.M. (2021). Routine sub-2.5 Å cryo-EM structure determination of GPCRs. Nat. Commun. *12*, 4333.

62. Schwartz, O., Axelrod, J.J., Campbell, S.L., Turnbaugh, C., Glaeser, R.M., and Müller, H. (2019). Laser phase plate for transmission electron microscopy. Nat. Methods *16*, 1016–1020.

63. Herzik, M.A., Wu, M., and Lander, G.C. (2019). High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. Nat. Commun. *10*, 1032. https://doi.org/10.1038/s41467-019-08991-8.

64. Herzik, M.A., Wu, M., and Lander, G.C. (2017). Achieving better-than-3-Å resolution by single-particle cryo-EM at 200 keV. Nat. Methods *14*, 1075–1078.

65. Cheng, Y., Grigorieff, N., Penczek, P.A., and Walz, T. (2015). A primer to single-particle cryo-electron microscopy. Cell *161*, 438–449.

66. Wu, M., Lander, G.C., and Herzik, M.A. (2020). Sub-2 Angstrom resolution structure determination using single-particle cryo-EM at 200 keV. J. Struct. Biol. X *4*, 100020. https://doi.org/10.1016/j.yjsbx.2020.100020.

67. Merk, A., Fukumura, T., Zhu, X., Darling, J.E., Grisshammer, R., Ognjenovic, J., and Subramaniam, S. (2020). 1.8 Å resolution structure of β-galactosidase with a 200 kV CRYO ARM electron microscope. IUCrJ *7*, 639–643. https://doi.org/10.1107/S2052252520006855.

68. Han, Y., Fan, X., Wang, H., Zhao, F., Tully, C.G., Kong, J., Yao, N., and Yan, N. (2020). High-yield monolayer graphene grids for near-atomic resolution cryoelectron microscopy. Proc. Natl. Acad. Sci. USA *117*, 1009–1014. https://doi.org/10.1073/pnas.1919114117.

69. Russo, C.J., and Passmore, L.A. (2014). Electron microscopy. Ultrastable gold substrates for electron cryomicroscopy. Science *346*, 1377–1380.

70. Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530.

71. Kuper, J., Braun, C., Elias, A., Michels, G., Sauer, F., Schmitt, D.R., Poterszman, A., Egly, J.-M., and Kisker, C. (2014). In TFIIH, XPD helicase is exclusively devoted to DNA repair. PLoS Biol. *12*, e1001954.

72. Abdulrahman, W., Iltis, I., Radu, L., Braun, C., Maglott-Roth, A., Giraudon, C., Egly, J.-M., and Poterszman, A. (2013). ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. Proc. Natl. Acad. Sci. USA *110*, E633–E642.

73. Busso, D., Keriel, A., Sandrock, B., Poterszman, A., Gileadi, O., and Egly, J.M. (2000). Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities. J. Biol. Chem. *275*, 22815–22823.

74. Kokic, G., Chernev, A., Tegunov, D., Dienemann, C., Urlaub, H., and Cramer, P. (2019). Structural basis of TFIIH activation for nucleotide excision repair. Nat. Commun. *10*, 2885.

75. Garrett, S., Barton, W.A., Knights, R., Jin, P., Morgan, D.O., and Fisher, R.P. (2001). Reciprocal activation by cyclin-dependent kinases 2 and 7 is directed by substrate specificity determinants outside the T loop. Mol. Cell Biol. *21*, 88–99. https://doi.org/10.1128/MCB.21.1.88-99.2001.

76. Merrick, K.A., Larochelle, S., Zhang, C., Allen, J.J., Shokat, K.M., and Fisher, R.P. (2008). Distinct activation pathways confer cyclin-binding specificity on Cdk1 and Cdk2 in human cells. Mol. Cell *32*, 662–672.

77. Yan, C., Dodd, T., He, Y., Tainer, J.A., Tsutakawa, S.E., and Ivanov, I. (2019). Transcription preinitiation complex structure and dynamics provide insight into genetic diseases. Nat. Struct. Mol. Biol. *26*, 397–406.

78. He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. Nature *495*, 481–486.

79. Schilbach, S., Hantsche, M., Tegunov, D., Dienemann, C., Wigge, C., Urlaub, H., and Cramer, P. (2017). Structures of transcription pre-initiation complex with TFIIH and Mediator. Nature *551*, 204–209.

80. Robinson, P.J., Trnka, M.J., Bushnell, D.A., Davis, R.E., Mattei, P.-J., Burlingame, A.L., and Kornberg, R.D. (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. Cell *166*, 1411–1422.e16.

81. Abdella, R., Talyzina, A., Chen, S., Inouye, C.J., Tjian, R., and He, Y. (2021). Structure of the human Mediator-bound transcription preinitiation complex. Science *372*, 52–56. https://doi.org/10.1126/science.abg3074.

82. Chen, X., Wang, X., Liu, W., Ren, Y., Qu, X., Li, J., Yin, X., and Xu, Y. (2022). Structures of +1 nucleosome–bound PIC-Mediator complex. Science *378*, 62–68. https://doi.org/10.1126/science.abn8131.

83. Chen, X., Qi, Y., Wu, Z., Wang, X., Li, J., Zhao, D., Hou, H., Li, Y., Yu, Z., Liu, W., et al. (2021). Structural insights into preinitiation complex assembly on core promoters. Science *372*, eaba8490. https://doi.org/10.1126/science.aba8490.

84. Chen, X., Yin, X., Li, J., Wu, Z., Qi, Y., Wang, X., Liu, W., and Xu, Y. (2021). Structures of the human Mediator and Mediator-bound preinitiation complex. Science *372*, eabg0635. https://doi.org/10.1126/science.abg0635.

85. Abril-Garrido, J., Dienemann, C., Grabbe, F., Velychko, T., Lidschreiber, M., Wang, H., and Cramer, P. (2023). Structural basis of transcription reduction by a promoter-proximal +1 nucleosome. Mol. Cell *83*, 1798–1809.e7. https://doi.org/10.1016/j.molcel.2023.04.011.

86. Espinoza, F.H., Farrell, A., Erdjument-Bromage, H., Tempst, P., and Morgan, D.O. (1996). A Cyclin-Dependent Kinase-Activating Kinase (CAK) in Budding Yeast Unrelated to Vertebrate CAK. Science *273*, 1714–1717.

87. Kaldis, P., Sutton, A., and Solomon, M.J. (1996). The Cdk-activating kinase (CAK) from budding yeast. Cell *86*, 553–564.

88. Thuret, J.-Y., Valay, J.-G., Faye, G., and Mann, C. (1996). Civ1 (CAK In Vivo), a Novel Cdk-Activating Kinase. Cell *86*, 565–576.

89. Keogh, M.-C., Cho, E.-J., Podolny, V., and Buratowski, S. (2002). Kin28 is found within TFIIH and a Kin28-Ccl1-Tfb3 trimer complex with differential sensitivities to T-loop phosphorylation. Mol. Cell Biol. *22*, 1288–1297.

90. Feaver, W.J., Svejstrup, J.Q., Henry, N.L., and Kornberg, R.D. (1994). Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. Cell *79*, 1103–1109.

91. Lee, K.M., Saiz, J.E., Barton, W.A., and Fisher, R.P. (1999). Cdc2 activation in fission yeast depends on Mcs6 and Csk1, two partially redundant Cdk-activating kinases (CAKs). Curr. Biol. *9*, 441–444.

92. Hermand, D., Pihlak, A., Westerling, T., Damagnez, V., Vandenhaute, J., Cottarel, G., and Mäkelä, T.P. (1998). Fission yeast Csk1 is a CAK-activating kinase (CAKAK). EMBO J. *17*, 7230–7238.

93. Kwiatkowski, N., Zhang, T., Rahl, P.B., Abraham, B.J., Reddy, J., Ficarro, S.B., Dastur, A., Amzallag, A., Ramaswamy, S., Tesar, B., et al. (2014). Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. Nature *511*, 616–620.

94. Olson, C.M., Liang, Y., Leggett, A., Park, W.D., Li, L., Mills, C.E., Elsarrag, S.Z., Ficarro, S.B., Zhang, T., Düster, R., et al. (2019). Development of a Selective CDK7 Covalent Inhibitor Reveals Predominant Cell-Cycle Phenotype. Cell Chem. Biol. *26*, 792–803.e10. https://doi.org/10.1016/j.chembiol.2019.02.012.

95. Zhang, T., Kwiatkowski, N., Olson, C.M., Dixon-Clarke, S.E., Abraham, B.J., Greifenberg, A.K., Ficarro, S.B., Elkins, J.M., Liang, Y., Hannett, N.M., et al. (2016). Covalent targeting of remote cysteine residues to develop CDK12 and CDK13 inhibitors. Nat. Chem. Biol. *12*, 876–884. https://doi.org/10.1038/nchembio.2166.

96. Kern, D.M., Sorum, B., Mali, S.S., Hoel, C.M., Sridharan, S., Remis, J.P., Toso, D.B., Kotecha, A., Bautista, D.M., and Brohawn, S.G. (2021). Cryo-EM structure of SARS-CoV-2 ORF3a in lipid nanodiscs. Nat. Struct. Mol. Biol. *28*, 573–582. https://doi.org/10.1038/s41594-021-00619-0.

97. Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P.M.G.E., Grigoras, I.T., Malinauskaite, L., Malinauskas, T., Miehling, J., et al. (2020). Single-particle cryo-EM at atomic resolution. Nature *587*, 152–156.

98. Neselu, K., Wang, B., Rice, W.J., Potter, C.S., Carragher, B., and Chua, E.Y.D. (2023). Measuring the effects of ice thickness on resolution in single particle cryo-EM. J. Struct. Biol. X *7*, 100085. https://doi.org/10.1016/j.yjsbx.2023.100085.

99. Hazel, P., Kroll, S.H.B., Bondke, A., Barbazanges, M., Patel, H., Fuchter, M.J., Coombes, R.C., Ali, S., Barrett, A.G.M., and Freemont, P.S. (2017). Inhibitor Selectivity for Cyclin-Dependent Kinase 7: A Structural, Thermodynamic, and Modelling Study. ChemMedChem *12*, 372–380. https://doi.org/10.1002/cmdc.201600535.

100. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

101. Saur, M., Hartshorn, M.J., Dong, J., Reeks, J., Bunkoczi, G., Jhoti, H., and Williams, P.A. (2020). Fragment-based drug discovery using cryo-EM.

Drug Discov. Today *25*, 485–490. https://doi.org/10.1016/j.drudis.2019.12.006.

102. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25. https://doi.org/10.1002/pro.3235.

103. Greber, B.J., Boehringer, D., Leibundgut, M., Bieri, P., Leitner, A., Schmitz, N., Aebersold, R., and Ban, N. (2014). The complete structure of the large subunit of the mammalian mitochondrial ribosome. Nature *515*, 283–286.

104. Greber, B.J., Toso, D.B., Fang, J., and Nogales, E. (2019). The complete structure of the human TFIIH core complex. Elife *8*, e44771.