Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Structural elucidation of modular allosteric regulation of the Swi2/Snf2 ATPases Mot1 and Ino80

Stephan Wilhelm Woike

aus

Münster (Westf.), Deutschland

2022

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 24.10.2022

.....

Stephan Wilhelm Woike

Dissertation eingereicht am 17.11.2022

1. Gutachter: Prof. Dr. Karl-Peter Hopfner
2. Gutachter: Prof. Dr. Roland Beckmann

Mündliche Prüfung am 12.12.2022

This thesis has been prepared from Dezember 2017 to Oktober 2022 in the laboratory of Professor Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-Universität München

This is a cumulative thesis based on the following publications:

Agata Butryn, Stephan Woike, Savera J Shetty, David T Auble, Karl-Peter Hopfner. Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state, *eLife* 7:e37774, 5th Oct 2018.

Stephan Woike, Sebastian Eustermann, James Jung, Simon Josef Wenzl, Götz Hagemann, Joseph Bartho, Katja Lammens, Agata Butryn, Franz Herzog, Karl-Peter Hopfner. Structural basis for TBP displacement from the TATA box by the Swi2/Snf2 ATPase Mot1, *Nature Structural and Molecular Biology*, accepted in principle, Sep 2022

Franziska Kunert, Felix J. Metzner, James Jung, Markus Höpfler, Stephan Woike, Kevin Schall, Dirk Kostrewa, Manuela Moldt, Jia-Xuan Chen, Susanne Bantele, Boris Pfander, Sebastian Eustermann, Karl-Peter Hopfner. Structural mechanism of extranucleosomal DNA readout by the INO80 complex, *Science Advances*, accepted in principle, Oct 2022.

One of the major lessons in all of biochemistry, cell biology and molecular medicine is that when proteins operate at the sub cellular level, they behave in a certain way as if they’re mechanical machinery.

James Rothman

Abstract

In eukaryotes, chromatin – the DNA packaged by nucleosomes and other bound proteins - is constantly reshaped by energy-dependent processes that facilitate accessibility of DNA for the replication, repair and transcription machinery. Swi2/Snf2 helicases translate energy derived from ATP-hydrolysis into DNA minor groove translocation resulting in either tracing DNA or pumping or pulling it to disrupt protein:DNA interactions, termed “chromatin remodeling”. The transcription regulator Mot1 is a single-subunit Swi2/Snf2 ATPase that removes TBP from the TATA box at the DNA promoter, thus recycling it and enabling a redistribution to other promoters. Despite a wealth of biochemistry, the chemo-mechanical details of the TBP removal were unknown. In the first publication, we present the crystal structure of near full-length Mot1 in an autoinhibited resting state. This allowed insight into the interaction between N-terminal HEAT-repeat arch and C-terminal ATPase in nucleotide-free state. In the second publication, we employed cryogenic electron microscopy (cryo-EM) to determine five structures of Mot1 bound to its TBP:DNA substrate with different ATP analogues. We could therefore dissect the stepwise dissociation of TBP from DNA in molecular detail and analyze the structure and function of the outermost C-terminal “bridge” element as an allosteric regulator of the remodeling activity of Mot1. Ultimately, we arrived at a model that involves a short-range, non-processive DNA translocation by Mot1, including bending and rotation of the DNA. This is in contrast to the processive DNA translocation of nucleosome remodelers, usually multi-subunit complex molecular machines that pump DNA around the histone octamer and thus slide nucleosomes and some even facilitate histone ejection and variant exchange. The resulting spaced nucleosome arrays and nucleosome-free regions are a prerequisite for DNA replication, repair and transcription. The INO80 complex is such a mega-Dalton multi-subunit nucleosome remodeler. In the third publication, we investigated the structural basis of INO80’s allosteric regulation by the so-called “A-module”. The A-module consists of nuclear actin in complex with actin-related proteins bound to a lever that feeds back to the motor ATPase. Although it is known that the A-module binds to extranucleosomal entry DNA, we present a model that explains INO80-specific monitoring of DNA shape by the A-module, the counter-grip subunit Arp5 and the motor ATPase itself. Consequently, mutual conformational feedback between the submodules yields a specific nucleosome positioning outcome.

Table of Contents

Introduction

1 The Swi2/Snf2 enzyme family and their mechanism of action

1.1 Superfamily 1 and 2 ATPases share a common architecture

Manipulation of nucleic acids is at the core of every life form, put into effect by six superfamilies (SF1-6) of helicases that facilitate replication, transcription or remodeling of the genetic material – whether it is DNA or RNA[1]. Helicases relay energy derived from hydrolysis of a nucleotide triphosphate (NTP, in most cases ATP) into mechanical movement between two ATPase lobes, which allows unidirectional movement along nucleic acid strands in an inchworm-like fashion[2–4]. *Bona fide* helicases processively track the nucleic acid leading strand with an additional "wedge" or "pin" element, separating double stranded DNA or RNA[5]. Superfamilies 3-6 comprise toroid helicases, which form hexameric assemblies and facilitate processive DNA strand separation[6]. Prominent members of these *bona fide* helicases are papilloma virus E1 helicase[7] (SF3), bacteriophage T7 gp4 protein[8] (SF4), bacterial Rho helicase[9] (SF5) or the eukaryotic MCM complex[10] (SF6).

Contrarily, SF1 and SF2 subfamilies commonly act as monomers and share a common architecture: two RecA-like lobes form a conserved core bordered by enzyme-specific regulatory domains[11] (see later). SF2 members are essentially part of all major cellular processes, like chromatin remodeling (Swi2/Snf2 ATPases[12,13]), splicing (DEAH helicases[14,15]), and as pattern-recognition receptors in the innate immune system (DExD/H helicases[16]), in repair of stalled replication forks (RecG helicases[17]) and as a tool against invading DNA (Type I/III restriction enzymes[18,19]) (Fig. 1.1.1 a).

Within the SF2 superfamily the so called "Swi2/Snf2" or "Snf2-like" ATPase family comprises a rich collection of subfamilies with a plethora of specialized enzymes. This manifests as evolutionary reinterpretations of the helicase core activity and is structurally based on conserved RecA-like building blocks[20]. The family name originates from a set of genes discovered in a genetic screen denoted "switch/sucrose non-fermenting", where mutations resulted in reduced activity of the sucrose fermenting enzyme invertase[21] and defective mating-type switching in yeast[22]. One of those was genes was identified as the Snf2 ATPase – the catalytic motor of the multi-subunit nucleosome SWI/SNF remodeling complex[23]. Early on, it became evident, that Swi2/Snf2

ATPases comprise a whole SF2 subfamily and are integral and abundant molecular motors in many cellular processes[24,25].

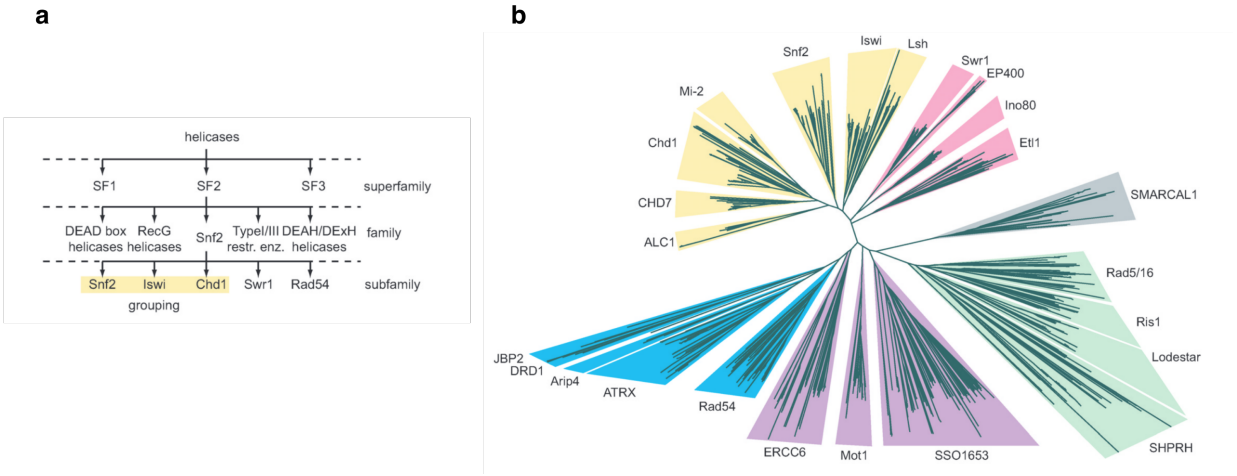


Fig. 1.1.1: Categorization of Swi2/Snf2 subfamilies (adapted from [26])
a) Hierarchical scheme of helicase families. b) Unrooted neighbor-joining tree based on multiple sequence alignments from the set of Snf2-like ATPases. Subfamilies are colored as follows: Snf2-like (yellow), Swr1-like (pink), distant (grey), Rad5/16-like (green), SSO1653-like (violet), Rad54-like (blue).

Flaus et al. made the latest attempt to catalogue the vast amount of genomic information on Swi2/Snf2 ATPases and grouped over 1300 family members into 24 subfamilies via multiple sequence alignment[26]. The multi-subunit INO80 complex as a prominent Swr1-like subfamily member will be described in more detail later, as well as the TBP:DNA dissociating Modifier of Transcription 1 (Mot1), which belongs to the SSO1653-like subfamily. More distantly related are the Rad54-like ATPases, that e.g., play a role in histone exchange (ATRX)[27,28], assist with branch migration in homologous recombination (Rad54)[28] or play roles in DNA lesion repair and replication stress tolerance (Rad5)[29] (Fig. 1.1.1 b).

1.2 Swi2/Snf2 ATPases remodel chromatin by DNA translocation

Instead of DNA or RNA strand separation as expected from *bona fide* helicases, Swi2/Snf2 ATPases employ the deeply conserved nucleic acid translocation activity to apply a rotational or protruding force to their nucleic acid substrate to remodel DNA-bound proteinaceous factors or nucleosomes[12]. Fairman-Williams et al.[11] identified 12 conserved sequence motifs for SF2 helicases, which founds the molecular underpinning of Swi2/Snf2 catalytic activity:

In an active conformation, structural motifs facilitating ATP binding and hydrolysis (motifs I, II and VI) align at the cleft of the lobes. Motif I (Walker A) binds the nucleotide phosphate moiety, whereas the Q motif coordinates the purine base. ATP hydrolysis is facilitated by motif II (Walker B), which harbors an aspartate followed by a glutamate. Aspartate coordinates the magnesium ion and glutamate conducts the subtraction of γ-phosphate in conjunction with motif VI at the opposite side. Engagement of all three motifs enables proper ATPase closure and hydrolytic activity (Fig. 1.2.1 a). To ensure coupling of ATP-hydrolysis and mechanical force towards the nucleic acid substrate, two motifs, III and Va, reside on opposite lobes. Placed between nucleotide binding pocket and nucleic acid binding site they are critical for the functional integrity of the enzyme (Fig. 1.2.1 b). Motifs Ia-c, IV, Vb are essential for proper accommodation of the nucleic acid substrate, contacting the phosphate backbone of the leading strand (Fig. 1.2.1 c)[11]. ATP-binding and hydrolysis induced opening and closing of the two RecA-like lobes relative to each other results in an inchworm or ratchet-like movement of the nucleic acid substrate, for most SF1 and 2 helicases in a 3' – 5' directionality[6].

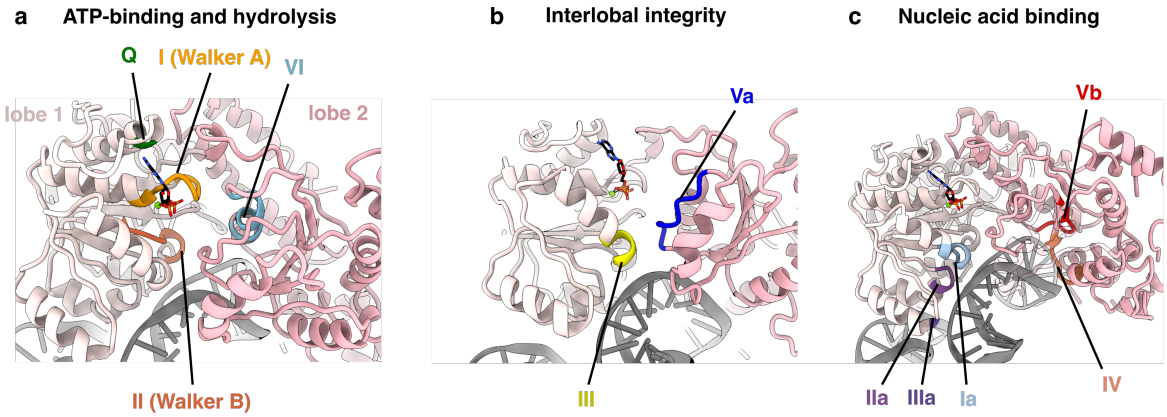


Fig. 1.2.1 Conserved sequence motifs in SF1 and SF2 helicases.
Motifs are indicated and colored in an ADP-bound yeast Snf2 ATPase structure (PDB 5Z3O).

Five more motifs were identified as essential for proper accommodation of the nucleic acid substrate: motifs Ia, IIa and IIIa on lobe 1 and IV and Vb on lobe 2 contact the phosphate-sugar backbone of RNA or DNA, thus binding it without sequence readout (Fig. 1.2.1 c).

Clearly, the conserved RecA-like core of Swi2/Snf2 ATPases determines their biochemical function as translocating DNA, but how do these enzymes differ structurally, allowing them to act on such a rich variety of cellular substrates?

1.3 Regulatory domains confer specificity

The primary sequence topology of a typical Swi2/Snf2 enzyme reveals an insertion between RecA1 and RecA2 harboring two helical domains denoted protrusion I and II, separated by a linker region, whereas C-terminal of motif VI, RecA2 extents into the so-called brace (Fig. 1.3.1). As we will see, even though these regulatory domains are relatively conserved in their basic architecture, differences between Swi2/Snf2 subfamilies cause diverse modes of action towards their specific substrates.

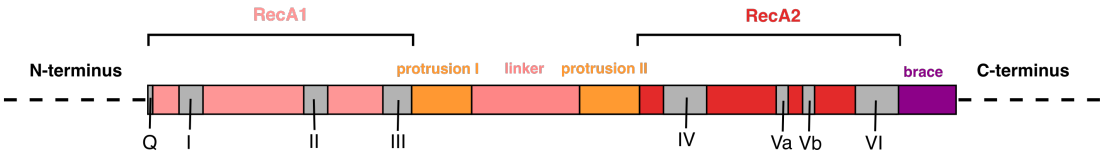


Fig. 1.3.1: Topology of a Swi2/Snf2 ATPase.
Regulatory domains and important conservered sequence motifs are indiciated.

As characterized by Flaus et al., the structural base of the two protrusions comprises highly conserved building blocks rich in aromatic residues, duplicated and assembled as variable helical extensions of differing length[26]. This is obvious when comparing the relatively short protrusion II of Snf2 and Chd1 ATPases with their longer counterparts in Swr1[30] and Mot1 (chapters 5.1 and 5.2) (Fig. 1.3.2). The brace folds like a tether from RecA2 back to RecA1 and is implicated in lobe-to-lobe communication during the process of ATP-binding and hydrolysis, modulating the conformational movements of the two lobes relative to each other[31,32]. Here again, the shaping of the brace noticeably differs between Swi2/Snf2 subfamilies. More recently, the brace was also assigned an extended role as being part of the so-called "integrative regulatory hub"[33] (see Chapter 3.5).

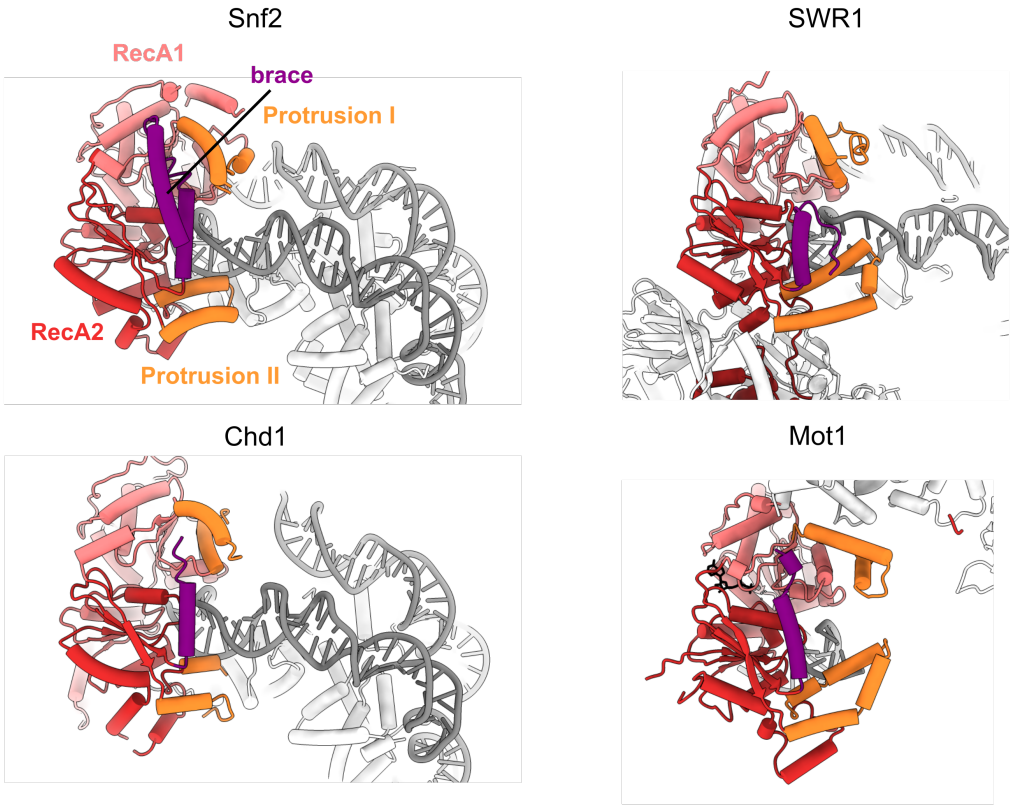


Fig. 1.3.2: Structural divergence of Swi2/Snf2 accessory domains.
Comparison of Swi2/Snf2 structures aligned via RecA1 (from top left to bottom right): *S. cerevisiae* Snf2 (PDB 5Z3V), *Homo sapiens* SWR1 (PDB 6GEJ), *S. cerevisiae* Chd1 (PDB 5O9G) and *Chaetomium thermophilum* Mot1 (PDB 7ZKE).

Besides functioning as a coupling domain tethering both RecA-like lobes and as part of a regulatory hub, the brace feeds back conformational input coming from the outermost C-terminus. Evidence is accumulating for many Swi2/Snf2 enzymes, that their outermost C-terminus acts as a molecular sensor conferring substrate specificity and relaying substrate binding towards the ATPase core via the brace. For example, in CHD remodelers substrate recognition is conferred by the DNA-binding domain (DBD)[34,35] and in ISWI remodelers by the HANT-SANT-SLIDE (HSS) domain[36,37], both interact with extranucleosomal DNA. An impressive example of evolutionary specialization is the distinct shape of the brace of retinoic acid-inducible gene I (RIG-I), a SF2 member and pattern-recognition receptor in the innate immune system's first line of defense[38]. RIG-I binds short duplex RNA via its RecA-like core ATPase. Hereby, specific recognition of a viral RNA fragment is ensured by recognition of the 5'-triphosphate moiety via a

C-terminal RD (regulatory/repressor domain)[39,40]. Proper placement of the RD domain at its 5'-RNA target induces a short-range translocation – most likely proofreading - step, accompanied by release of the N-terminal CARD (caspase activation and recruitment) domains from their autoinhibited state and downstream immune signaling[41].

The brace of RIG-I consists of two extensive α-helices that fence protrusion I. The N-terminal brace helix called pincer I spans both RecA-like lobes, whereas the C-terminal helix or pincer II extends along RecA1. A proline-rich linker connects the brace with the C-terminal RD domain (Fig. 1.3.3).

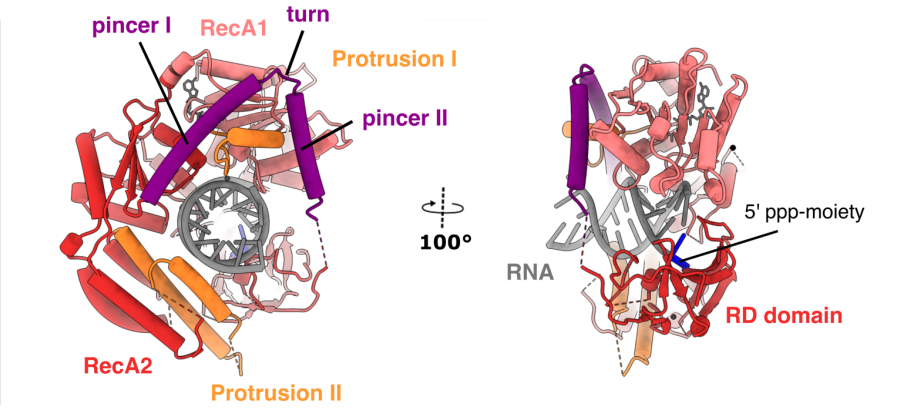


Fig. 1.3.3: Structure and regulatory domains of RIG-I.
Depicted are two rotational view of the *H. sapiens* RIG-I structure bound to dsRNA (PDB 5E3H).

Rawling et al. functionally dissected pincer:RecA1/2 binding, which led to insights into the role of the pincer helices as information hubs that integrate C-terminal substrate binding and ATPase activity. Mutational perturbations of the pincer II:RecA1 interaction resulted in decreased dsRNA binding and ATPase hydrolysis rate, ultimately leading to a dampened immune response[42]. In comparison with other SF2 members, the pincer2/RecA1 interaction has apparently undergone co-evolutionary structural adaptions towards higher substrate specificity.

Although the extensive brace in RIG-I is a special case in the SF2 helicase family, it reflects its role as an integrator of substrate binding and core ATPase motor activity also for other Swi2/Snf2 family members.

1.4 The ATPase terminal domains act as substrate sensors and allosteric regulators

Allosteric regulation of the ATPase by N and C-terminal substrate sensors is a common feature of Swi2/Snf2 enzymes. These auxiliary domains act as controlling instances to allow context-dependent regulation of the enzymatic activity[43]. For Swi2/Snf2 ATPases, the most basic mechanism of regulation is the control of RecA-like lobe closure. Apparently, some subfamily members are in a thermodynamic equilibrium between an open state with the lobes rotated relative to each other and a closed state[44]. The latter correlates with formation of the ATP-binding pocket at the cleft between properly aligned lobes and is often achieved through a swiveling motion of lobe 2 upon ATP-binding and/or substrate binding as exemplified for Rad54[45,46]. An interesting variety of this mechanism is Cockayne syndrome protein B (CSB), an Snf2 ATPase that acts as part of a multi-subunit complex in transcription-coupled repair. When elongating RNA polymerase II is inhibited by a bulky DNA lesion, CSB moves the transcription complex over the lesion by a short-range translocation step[47,48]. Effective ATPase closure depends on the ATPase C-terminus, which acts as complex assembly control. Binding of the CSB C-terminal CSA-interacting motif (CIM) suppresses RecA2 flexibility and stabilizes the ATPase in an active form with closed lobes[49] (Fig. 1.4 a, b).

Additional levels of regulatory complexity have emerged in nucleosome remodelers that react to a dynamic epigenetic environment. Some nucleosome remodelers make use of terminal auxiliary domains to contact histone epitopes. ISWI, to example, is released from autoinhibition by anchoring to the histone H2A/H2B acidic patch simultaneously via its N-terminal AutoN and C-terminal NegC domains[50] and RSC is activated by binding to the H2A/B acidic patch via the C-terminus of its motor ATPase Sth1[32]. But instead of a simple switch-on/off mechanism, an intricate interplay between the auxiliary domains fine-tunes the remodeling outcome in reaction to diverse substrate cues. In addition to histone epitopes, nucleosome remodelers dynamically sense extranucleosomal DNA to achieve directed nucleosome sliding towards the longer DNA overhang. As mentioned before, the chromodomain-helicase-DNA binding protein 1 (Chd1) harbors C-terminal SANT and SLIDE domains – collectively denoted DNA-binding domain (DBD) - that bind to extranucleosomal DNA, while the core ATPase resides at superhelical location 2 (SHL 2)[51] (Fig. 1.4 c). Chd1 has overlapping nucleosome positioning activity with other remodelers but has been also implicated as a specialized factor in facilitating polymerase II transcription through nucleosomes[52]. Truncating the C-terminal domains (Chd1$^{\Delta SANT/SLIDE}$) reduces remodeling activity and ATPase activity of Chd1 manifold compared to wildtype Chd1 (Chd1$^{\Delta N}$, aa118-939 expression construct) (Fig. 1.4 d). Clearly, the C-terminal DBD as an allosteric activator of nucleosome

positioning functions as an on-switch in response to extranucleosomal DNA, leading nucleosome sliding towards the DBD-bound DNA[35].

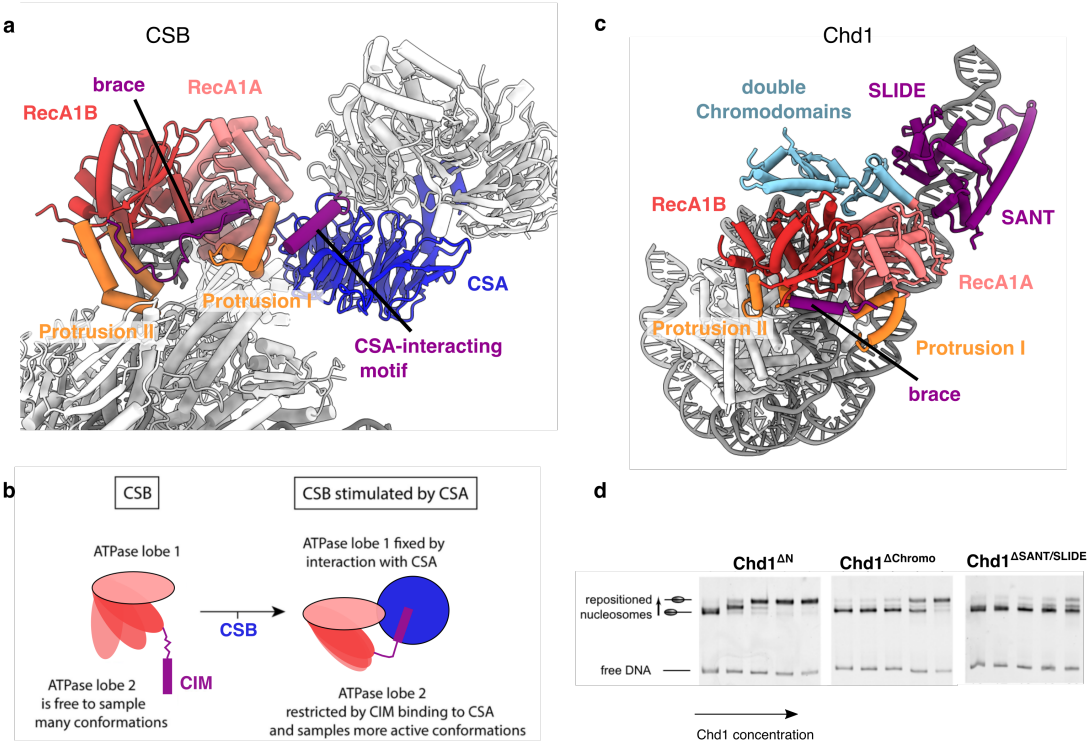


Fig. 1.4: Influence of C-terminal accessory domains on Swi2/Snf2 regulation.
a) Structure of the *H. sapiens* PolII-CSB-CSA-DDB1-UVSSA (PDB 7OOB) with CSB and CSA indicated and colored.
b) Model of CSB stabilization by restriction of the C-terminal CIM helix (modified from [47]).
c) *S. cerevisiae* Chd1:nucleosome structure (PDB 5O9G) with functional ATPase regions colored.
d) Nucleosome positioning gel shift assay for Chd1 truncations. Chd1 concentrations are (from left to right): 0, 0.1, 1, 10 and 100 nM (modified from [53])

In addition, Chd1 activity underlies a complex fine-tuning, obtained by the N-terminal chromodomains. A crystal structure of the substrate and nucleotide-free autoinhibited state of CHD1 revealed that the chromodomains occlude DNA access to the ATPase[53]. Consistently, a chromodomain deletion construct (Chd1$^{\Delta Chromo}$) showed a reduced sliding activity compared to wildtype Chd1 (Fig. 1.4 d), and a 200x higher binding affinity towards linear DNA over nucleosomes. By ATP-binding and subsequent ATPase closure the chromodomains are released, which are then able to bind nucleosomal DNA at SHL 1[53].

More specific, release of the double chromodomain is critical for engagement of the ATPase with its nucleosome target by enabling ATPase contacts to the H3 α1 helix and the histone H4 tail[54]. Additionally, the outermost Chd1 N-terminus competitively occupies the nucleosomal H2A/H2B acidic patch via the so called ChEX domain – suggestively conditional for Chd1 assisted transcription through nucleosomes[55]. It is evident that an intricate coordination of ATPase motor and coupled regulatory and accessory domains enable substrate discrimination and appropriate nucleosome sliding. As we will see, the concept of substrate sensing and feedback loops that regulate ATPase and remodeling activity is a common theme in Swi2/Snf2 ATPases. Intriguingly, this holds true as a unifying concept from single subunit enzymes like Chd1 and Mot1 up to multi-subunit, mega-Dalton complexes such as SWI/SNF or INO80.

1.5 Different modes of DNA translocation

A defining feature of SF2 helicases is an ATP-dependent nucleic acid translocation activity, which manifests either as directional minor groove tracking either in a processive and continuous manner or in short-range non-processive steps. *Bona fide* helicases translocate their nucleic acid substrate processively, meaning a repetitive flow of ATP-binding and hydrolysis whilst remaining bound to the nucleic acid strand and separating duplex DNA or RNA[56]. While different theories about the molecular mechanics underlying helicase progression circulated, structural studies of the PcrA helicase mimicking two different nucleotide-dependent conformations proposed the now widely accepted “inchworm” model of translocation: opening of the RecA-like lobes allows an extended grip on the phosphate backbone and by ATP-dependent closing, the nucleic acid is pulled in stepwise[2]. Three crystal structures of NS3 on singe-stranded DNA, mimicking ATP-binding (ADP-BeF_3), -hydrolysis (ADP-AlF_4^-) and nucleotide-free conformations, delineated the inchworm mechanism as conformational snapshots[5]. Translocation of DNA is also the core activity of nucleosome remodelers, but there are problems when applying the inchworm model to DNA sliding in the context of a nucleosome[57]. Fundamental building blocks of a nucleosome are histones, each comprising a three-helical histone fold domain. The disc-shaped nucleosome consists of 8 histones, where two H3-H4 dimers form the core onto which the two H2A/H2B dimers assembly on opposite sides[58]. Long N-terminal tails have emerged as a site of posttranslational modifications that serve as epigenetic markers and regulate the protein machinery interacting with chromatin[59] (Fig. 1.5.1 a). Typically, 147 bp of DNA is wrapped around the octameric nucleosome core. The core-DNA association is predominantly conferred via electrostatic interactions between positively charged histone amino acids residues and the

negatively charged DNA phosphate backbone and arginine insertions into the minor groove at 14 contact sites (via hydrogen bonds and non-polar interactions)[60] (Fig. 1.5.1 b).

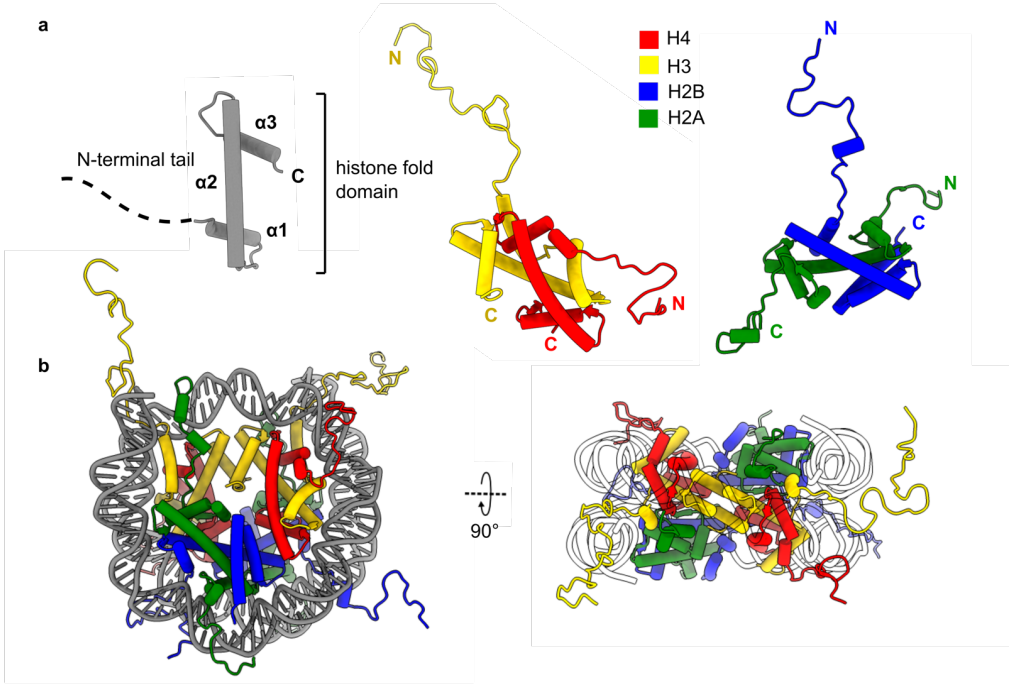


Fig. 1.5.1: Structure of a nucleosome (PDB 1KX5).
a) Visualization of a histone fold motif (left) and assembly of H3/H4 and H2A/H2B histones (right).
b) Two rotational views of a nucleosome (histones colored as in a)).

Obviously, when nucleosomes are repositioned by remodelers, the sum of histone-DNA interactions presents a formidable energy barrier with contributions from DNA intrinsic mechanistic properties such as bending and elastic energy depending on the sequence[61].

It is evident that stepwise DNA translocation with the energy derived from single ATP-hydrolysis events would not suffice to break all histone-DNA contacts all at once. An intriguing solution would be that energy might accumulate from many hydrolyzed ATPs, thus pumping DNA to create a bulge against a counter grip, as proposed for INO80[62]. When a certain size of the bulge would be reached, its release would push the DNA around the nucleosome in wave-like manner, which would explain the 10-20 bp step size observed for nucleosome sliding by INO80[63] (Fig. 1.5.2 a). As other remodelers do not feature a counter gripping domain or subunit, another model has gained popularity. With increasing availability of high-resolution remodeler structures, DNA binding modes could be compared between nucleotide states. First, three structures of ADP-BeF_x- and ADP-bound[64] and apo[31] states of the yeast Snf2 ATPase interacting with a nucleosome

were solved by cryo-EM. In the ATP-ground state mimic (ADP-BeF_x) the DNA was in B-form, whereas in the ADP or apo states an additional base pair was accommodated on the tracking strand. Apparently, the ADP and apo states induce a local twist-defect, which is resolved upon ATP rebinding. The twist defect therefore propagates around the entire nucleosome, ultimately stimulating stepwise DNA translocation around the histone octamer[65,66]. Recent structural work on Chd1 arrived at a similar model, with the additional insight that the twisted DNA adopts A-form geometry[55] (Fig. 1.5.2 b)

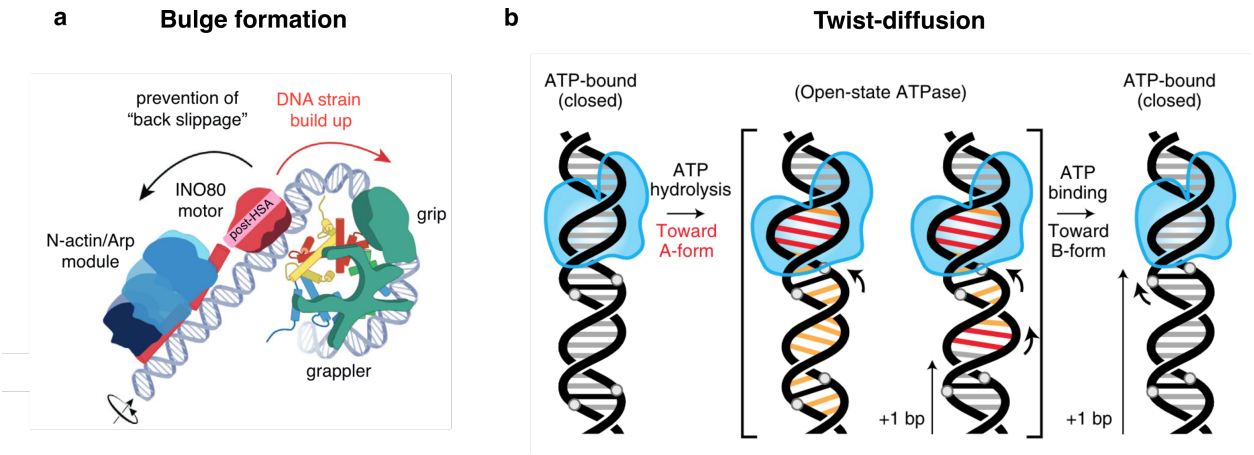


Fig. 1.5.2: DNA translocation models for nucleosome sliding.
a) Proposed bulge DNA bulge formation against by INO80 against a counter grip (adapted from [67]
b) Twist-diffusion model entailing A-form DNA as proposed for Chd1 (adapted from [55]).

But still, how does the twist-diffusion model solve the problem to overcome the energy barrier formed by the sum of histone-DNA contacts? This view is grounded in a misconception, which sees the nucleosome as "static", being in a stable thermodynamic minimum. In recent years, this view was repeatedly challenged e.g., by molecular dynamics simulations and the concept of "nucleosome breathing"; it appears that the nucleosome undergoes a series of low-barrier thermodynamic equilibria with dynamic changes in wrapping, turn and twist parameters[68]. Underlying this phenomenon is the repeated re-arrangement of histone-DNA contacts – energy consumed for breaking of one histone-DNA contact is almost simultaneously set free again by formation of the succeeding interaction, resulting in a neutral net energy balance. Therefore, a short-range DNA translocation via twist-defect diffusion would mean only a relatively limited energy expense – and as shown by structural data, a local twist defect in nucleosomal DNA is created by a Swi2/Snf2 ATPase in the ADP-bound or apo conformation[31,55,66]. Additionally, it was

assessed how such local twist defect could propagate around the whole nucleosome core via coarse-grain molecular dynamics simulations[69]. Now, also backed up by structural evidence[70], this elegant model of DNA propagation offers an intriguing explanation of DNA sliding around nucleosomes.

2 Modifier of Transcription 1 (Mot1)

2.1 Mot1 emerged as an essential transcriptional regulator

Mot1 was independently discovered by two biochemical approaches – on sequence level by genetic screens and as a byproduct in protein purification. In yeast genetic screens, it was identified as an *in trans* repressor of several target genes. When mutated, it led to an upregulation of the product of the alcohol dehydrogenase 2 (ADH2)[71]. Around the same time, another group found a suppressor mutation for pheromone-responsive genes in a gene locus they hence denoted "modifier of transcription 1 (MOT1). They were the first to clone and sequence the respective open reading frame (ORF) and identified it as an essential SF2 helicase, similar to Snf2 and Rad54[72]. But still, how this helical gene product represses transcription of other genes remained elusive. Independent of the genetic evidence, another group characterized Mot1's molecular activity *in vitro* on a protein level[73]. Their work founded on prior knowledge, that TBP – as part of TFIID - stably associates with the promoter TATA-box, establishing a transcriptionally competent complex[74]. They noticed that upon incubation with crude TBP purified from yeast, the typical TBP DNase footprint on adenovirus major late promoter DNA diminished after addition of ATP[73]. By identification, stepwise enrichment and purification of the targeted protein, they established a reductive *in vitro* reconstitution to narrow down the novel protein's activity.

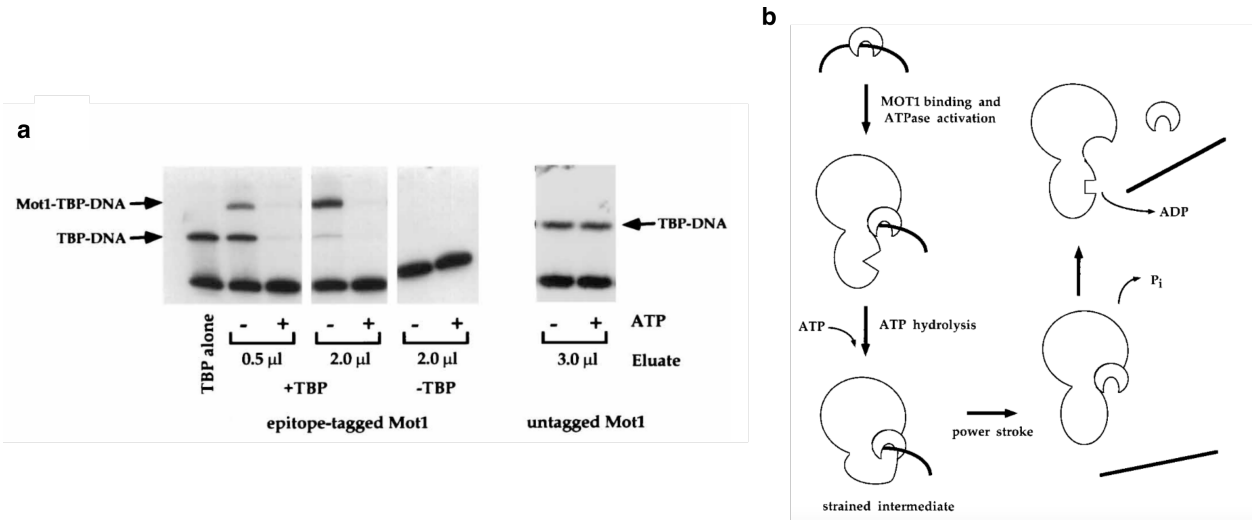


Fig. 2.1: *In vitro* analysis and early model of Mot1 activity (adapted from [75]).
a) Gel shift assay of ATP-dependent Mot1 activity acting on a TBP:DNA substrate.
b) Simple model of Mot1:TBP:DNA complex formation and ATP-dependent disruption.

In native gel shift experiments, TBP bound stably to the TATA-box of promoter DNA leading to higher retention in the gel, visible as an upwards shift compared to DNA alone. Upon addition of Mot1 a super-shifted Mot1:TBP:DNA complex was formed, which dissociated after addition of ATP. Mot1 alone, although bearing an ATPase competent for DNA-binding, did not show visible DNA-binding (Fig. 2.1 a)[75,76].

It was apparent, that Mot1's recruitment to the promoter site conveys through interaction with TBP and binding of DNA upregulates ATPase activity. By partial truncations the N-terminal region was identified as the site of TBP interaction, whereas the C-terminus moiety maps to upstream DNA. These findings were incorporated into a model that proposed a "power stroke" induced by ATP-hydrolysis resulting in a strained intermediate conformation, which would lead to dissociation of the Mot1:TBP complex from DNA[75] (Fig. 2.1 b).

In the light of these findings it was even more evident that the ATP-driven disruption of protein:DNA assemblies lies at the heart of Swi2/Snf2 function, as it had been discovered that the SWI/SNF complex disrupts histone:DNA contacts[77].

In the following years the group of David Auble continued to investigate the molecular intricacies of Mot1 biochemistry by refining the *in vitro* reconstitution system and using diverse DNA substrates. Transcriptional profiling revealed a counterintuitive finding: inhibiting Mot1 *in vivo*

led to repressed genes as expected, whereas other genes became transcriptionally activated[78]. Thus, it was repeatedly speculated that Mot1 could also be part of transcriptionally competent assemblies[79] or even form alternative pre-initiation complexes with TBP, TFIIB and polymerase II for stress-responsive genes[80]. Evidence now rather points to an indirect role of Mot1 in activating certain genes by redistributing TBP between different kinds of promoters. As TBP binds strongly to strong TATA-box promoters, Mot1 could bring it back into solution by disruption these high-affinity TBP:DNA complexes and transfers it to TATA-less promoters – either directly or via handover to TFIID or SAGA[81].

By using a natural promoter DNA sequence from the URA1 promoter, an additional cause of Mot1's transcriptional activation was demonstrated *in vitro*: the disruption of transcriptionally incompetent TBP:DNA complexes. At the URA1 promoter a strong TATA-sequence exists in two variants – one that leads to initiation of transcription upon TBP binding and the incorrect one on the opposite strand, resulting in TBP binding in the wrong orientation. The wrongly orientated TATA-box has a higher TBP affinity, so energy-dependent disruption by Mot1 is necessary to enable transfer of TBP and enhance productive binding[82]. All these findings widen the scope of Mot1's role in the cell that goes beyond the mere dislodgement of TBP. Rather, Mot1 is reactive to the genomic and cellular context of TBP and as we will see in the next chapter, fiercely competes with other TBP binding factors.

Therefore, before delving into the chemo-mechanical interplay between Mot1 and its TBP:DNA substrate, it is insightful to closer inspect the biological role of TBP in the cell.

2.2 TBP is the cornerstone of transcription initiation

On a molecular level, every biological organism relies on genetically encoded information to establish and maintain its cells. Complex multicellular organisms rely on a variety of specialized cell types, which emanate due to highly regulated differential gene expression. In eukaryotic transcription, a subset of genes is copied into RNA dependent on the specific cellular context. Within the course of evolution, the interplay between regulatory factors influencing transcription increased in numbers and complexity[83]. This is impressively reflected by the up to 44 polypeptides that constitute the polymerase II initiation complex[84]. As we will see, regulation of such a biological system achieves complexity not only by the number of participating factors, but already in its first basic step – in this case the decisive recognition of DNA by the first protein arriving at the promoter site[85].

For all archaea and eukaryotes, TATA-box binding protein (TBP) is the first protein that establishes contact to the promoter. Its saddle-shaped 2-fold structural symmetry forms a hydrophobic concave site, which has an affinity to AT-rich DNA. Although TBP preferably binds AT-rich sequences, it also resides at CpG-rich and other promoter sequences[86]. A rate-limiting step towards DNA binding is the propensity of TBP to self-dimerize via its C-terminus[87]. However, TBP shows the highest affinity to the so-called strong TATA-box, with a 1000-fold higher affinity compared to unspecific DNA *in vitro*[88]. The TBP:DNA interaction is kinetically remarkably stable, with *in vitro* residency half times of up to 60 minutes[89].Binding to DNA is mainly facilitated by insertion of opposing pairs of two phenylalanines respectively into the minor groove, which leads to a 90° kink and unwinding of TATA-box DNA[90,91]. In an extensive analysis it was shown that the sequence surrounding the TATA-motif is a determinant of TBP turnover rates. A “strong” TATA-box correlates with dynamic “noisy” gene expression as a stress response to external cues, whereas “TATA-less” sequences are predominantly abundant in constitutive gene promoters with long residence times of TFIID and constitutive “household” transcription[92].

The convex site of TBP evolved as a binding platform for other general transcription factors. As examined in an evolutionary analysis, the N-terminal convex half developed into a binding hub exhibiting a conserved patch of positively charged amino acids, where general transcription factors compete for binding[93] (Fig. 2.2 a). Exclusive occupation by different proteins confers another layer of complexity in transcriptional regulation manifesting as a competition between activators (e.g., TFIIA, TFIID, Brf1) and local suppressors (Mot1, NC2).

This exemplifies in crystal structures of TFIID component TAF1 in polymerase II or Brf1 in polymerase III transcriptional complexes: both anchor at the TBP regulatory region through electrostatic interactions[94,95] (Fig. 2.2 b). TFIIB interacts with the adaptor region on the opposite half of TBP, which explains the steric allowance of simultaneous binding of TFIID and TFIIB in the polymerase II initiation complex. On the other hand, Mot1 competes with TFIID – and also for TFIIA (not shown) - for the regulatory region of TBP. Together with its cofactor NC2, binding at the adaptor region, Mot1 would occupy TBP completely, increasing the likelihood of successful TBP disruption[96] (Fig. 2.2 b).

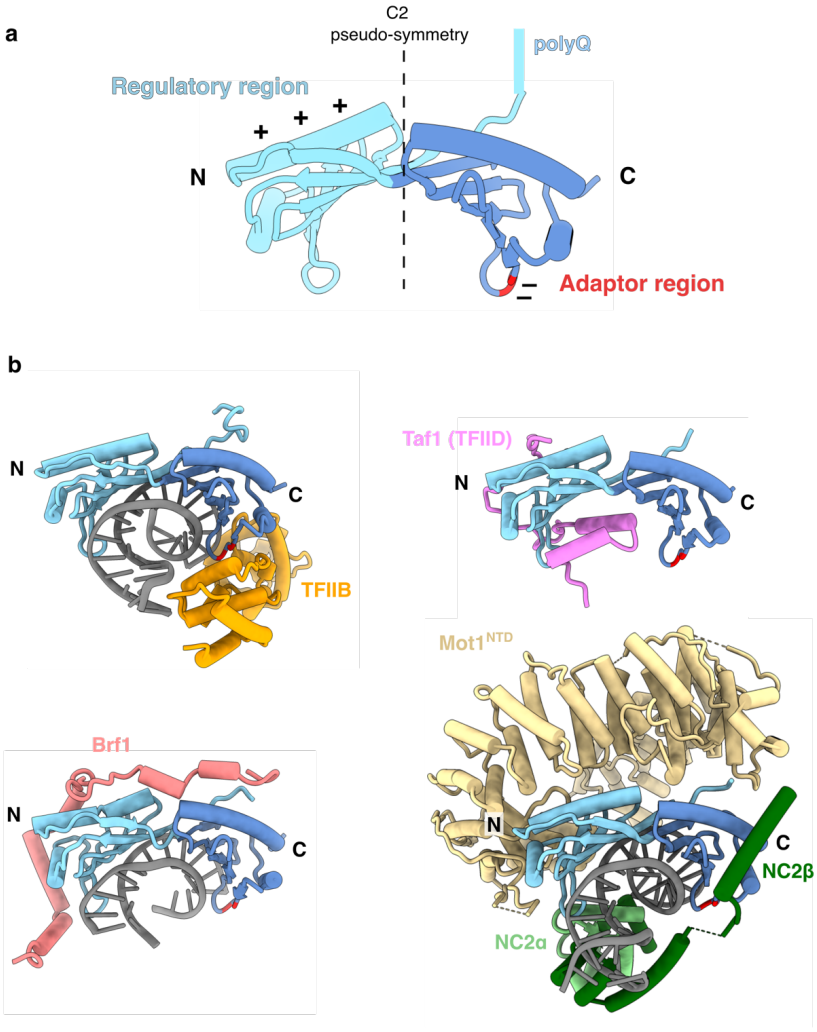


Fig. 2.2: Competitive occupation of TBP Regulatory and adaptor regions by transcriptional co-factors.

a) Schematic view of the TBP structure. Structural regions acting positively (+) and negatively (–) charged binding hubs are designated as well as N- (N) and C-terminal (C) halves.
b) Exemplary structures of TBP bound by transcriptional co-factors (from top left to bottom right): TBP:DNA:TFIIB (PDB 1VOL), Taf1:TBP (PDB 4B0A), Brf1:TBP:DNA (PDB 1NGM) and Mot1:TBP:DNA:NC2 (PDB 4WZS).

2.3 Structural information elucidates Mot1, TBP and DNA interplay on a molecular level

Although Mot1 was biochemically well characterized, there was a lack of structural information, especially regarding the interaction between Mot1^{NTD} and TBP and the working mode of the ATPase. Moreover, how the ATP-driven force translates into disruption of TBP and DNA was unknown and would have benefited from structural insights. The first crystal structure of Mot1 in complex with TBP was published in 2011[97]. The complex from the microsporidium *Encephalitozoon cuniculi* (*Ec*) shows the ring-shaped architecture of Mot1^{NTD} comprising 16 HEAT (Huntingtin, elongation factor 3, protein phosphatase 2 A and lipid kinase TOR) repeats. This spring-like HEAT repeat array approaches TBP at its convex site via the so-called N-terminal regulatory domain on helix 2 (H2) – a conserved mode of TBP binding was discussed in chapter 2.2. A second anchor point is formed by a protruding loop connecting the two helices of HR 4 that binds to the turns connecting TBP's beta sheet (Fig. 2.3 a).

As the Mot1:TBP complex was crystallized without DNA it represents the so called "product state" after successful dissociation of TBP from DNA. Although Mot1^{CTD} could not be resolved by X-ray crystallography, assumptions could be made about the mechanism of TBP disruption. Emerging from between HEAT repeats 2 and 3, a partially structured "latch" occupies the concave DNA-binding site of TBP, which could compete with DNA binding and seems to be a key element of successful TBP dislodgement (Fig. 2.3 a). The latch might act like a bottle-opener by inserting at the TBP concave site and inhibit rebinding to the TATA-box. Indeed, when truncating the latch, the disruption was inhibited showing its essential participation in the TBP:DNA dissociation process (Fig. 2.3 c). In the light of this work, but aiming at a different confirmation, another crystal structure of *Ec*Mot1^{NTD} bound to DNA was published in complex with negative cofactor 2 (NC2) assembled at TBP:DNA opposite of Mot1[96]. NC2 comprises two subunits, NC2α and NCβ, that each bear a histone-fold domain, which allow them to align to the DNA curvature imposed by TBP (Fig. 2.3 b). *In vivo*, Mot1 and NC2 mostly co-localizes at promoter sites and it has been speculated that NC2 marks TBP for effective Mot1 recruitment[98]. Consistently, NC2 occupies functional binding sites on TBP and hinders binding of other transcription factors such as TFIIA[99]. Also, it was shown that TBP:NC2 complexes can slide along DNA, probably as a search mode ensuring proper TBP binding at strong TATA-boxes[100]. Gel shift assays confirmed that NC2 stabilizes the Mot1:TBP:DNA assembly, but does not detectably affect ATP-dependent TBP:DNA dissociation rate (Fig. 2.3 c).

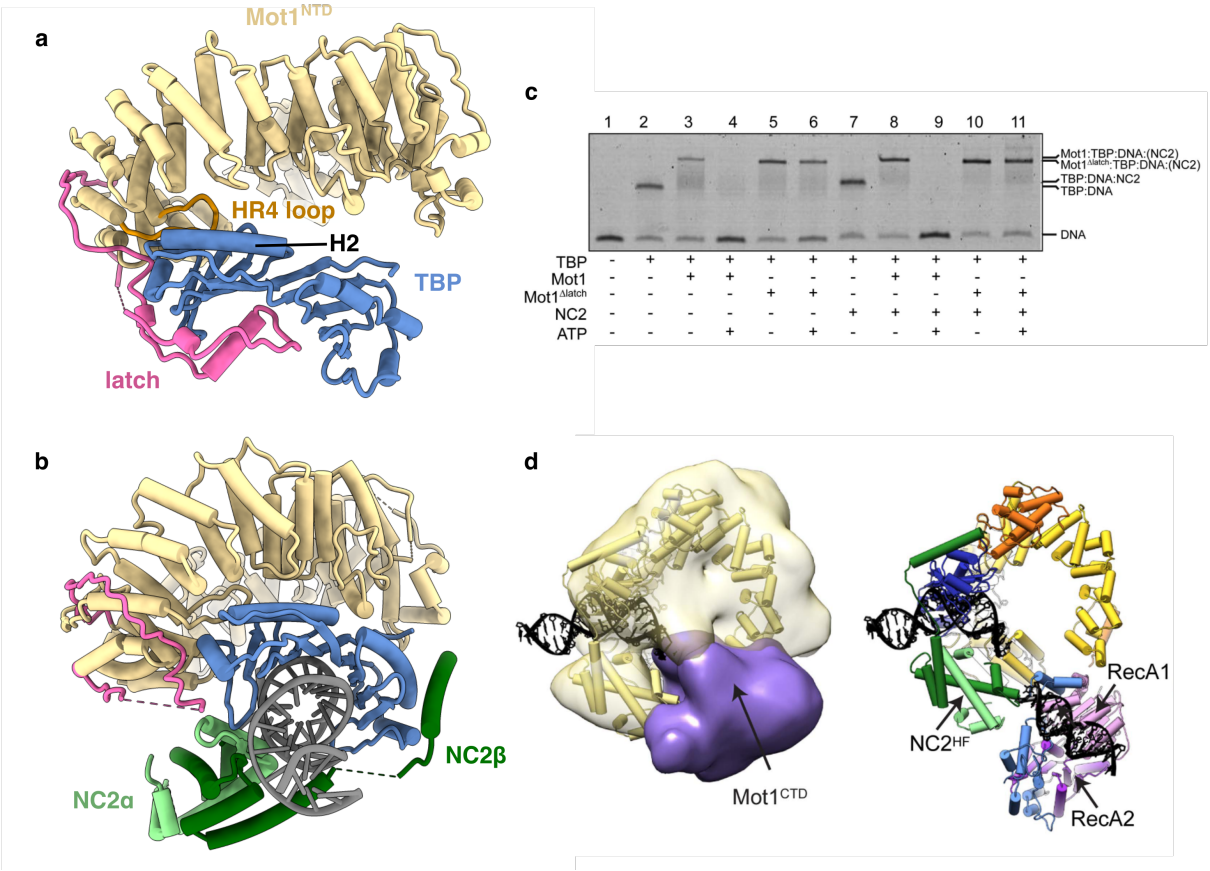


Fig. 2.3: Structural studies of *Ec*Mot1:TBP:DNA(:NC2).
a) Structure of Mot1^{NTD}:TBP (PDB 3OC3) showing TBP recognition by the HEAT repeats (HR) and HR4 loop and latch insertion at the TBP concave site in the so called "product state".
b) Structure of Mot1^{NTD} bound to TBP:DNA clamped by NC2 (PDB 4WZS).
c) Gel shift assays displaying the stepwise assembly of an Mot1:TBP:DNA complex and examining the effects of a Mot1$^{\Delta latch}$ mutant, NC2 binding and ATP addition (adapted from [96]).
d) Low resolution negative stain EM map of complex from b) but with full-length Mot1. Overall complex topology can be identified by rigid-body docked Mot1^{NTD}:TBP:DNA:NC2 complex (yellow) and colored density for Mot1^{CTD} (violet) (left). An *Sso*Rad54-like:DNA crystal structure (PDB 1Z63) could be docked into the C-terminal density resulting in a provisional model (right) (adapted from [96]).

To overcome the limits of crystallization and elucidate the topology of Mot1^{CTD}, negative stain EM was employed, which revealed extra density located at a site, which would correspond to upstream DNA. A *Sso*Rad54-like:DNA crystal structure[45], with one of the closest known sequence similarities to yeast Mot1, fitted reliably into the density (Fig. 2.2.1 d). But still, due to the low

resolution of the microscopy data, the detailed setup of the ATPase and its role in the TBP:DNA disruption remained a matter of speculation.

A step closer to a better understanding of the C-terminal ATPase of Mot1 was the crystal structure of the full-length Mot1, which for the first time yielded a high-resolved detailed structure of Mot1^{CTD101} (Chapter 5.1).

2.4 The molecular mechanism of Mot1 remodeling activity remained elusive

Despite the progress in the structural characterization of Mot1, a lack of understanding remained of how the C-terminal ATPase translates an energy-dependent conformational change into dislodgement of TBP from DNA. Due to its relationship to the Swi2/Snf2 family and thus similarity to nucleosome remodeler ATPases, there was an early interest in Mot1 as a simple model system for its often multi-subunit and more complex relatives. Based on the characteristic DNA translocation activity of Swi2/Snf2 ATPases, one early assumption was that Mot1 could act analogously. This would mean either a translocation of Mot1 from upstream DNA towards or away from TBP, resulting in a pushing or pulling force. Another variant of the same scenario would be a dislocation of the TATA-box by Mot1, which, engaged with TBP, could pull or pump DNA to break the TBP:DNA interaction[102]. Opposing to this idea, a DNA footprinting experiment could not show detectable DNA translocation of Mot1 during TBP disruption. However, this did not exclude the possibility of a short-range of just few base pairs[103].

2.5 Mechanistic models for TBP disruption by Mot1

Insights into Mot1's catalytic activity have been mostly relying on so-called "ensemble measurements". These are based on observations of a system in a thermodynamic equilibrium with nano- up to micromolar amounts of protein. It allows to follow the catalytic event from educts to products and analyze the effect of nucleotide addition. Although this experimental setup allows to characterize the overall enzymatic activity, the structural rearrangements and chemo-mechanical activity of the protein complex remain unobservable. Besides structural methods such as X-ray crystallography and cryo-EM, which capture stabilized conformations and allow mere identification of dynamic regions, progress in biophysical methods such as Förster Resonance Electron Transfer (FRET) enabled a time-resolved dynamic perspective. Driven by the still open question of how the Mot1 ATPase acts on DNA to achieve TBP disruption, the Auble group extended their usual gel-based assays by a dual-label FRET setup for the Mot1:TBP:DNA

complex. The fluorophore donor and receptor were located either up- and downstream of the TATA-box or on TBP and upstream of the TATA-box. This allowed to monitor changes in DNA bending and TBP:DNA proximity dependent on Mot1 and eventual nucleotide addition by FRET. It was observed that Mot1-binding alone did lead to an unbending of the 90° kinked TBP-bound DNA and that ATP-binding alone was insufficient for complete TBP disruption. Instead, TBP remained in close proximity to DNA. Only ATP-hydrolysis led to complete dissociation. This was indirectly confirmed by adding ATPγS, which was expected to function as an unhydrolyzable ATP analogue but still resulted in partial TBP disruption[104] (Fig 2.5.1 a). Apparently, ATPγS was slowly hydrolyzed, which strengthened the proposal that ATP-hydrolysis is needed for complete dissociation. By measuring the binding affinity of the mere ATPase domain towards DNA, a nucleotide dependence was revealed. The Mot1 C-terminal ATPase supplemented with ATP and analogues mimicking the ATP ground state – ATPγS and ADP-AlF$_x$ – showed a reduced affinity to DNA compared to apo state (Fig. 2.5.1 b). This implicates alternating cycles of gripping and releasing the DNA during ATP-binding and -hydrolysis in an inchworm-like fashion, as also seen for other SF2 members[2,4,5], in favor of a small-range DNA translocation.

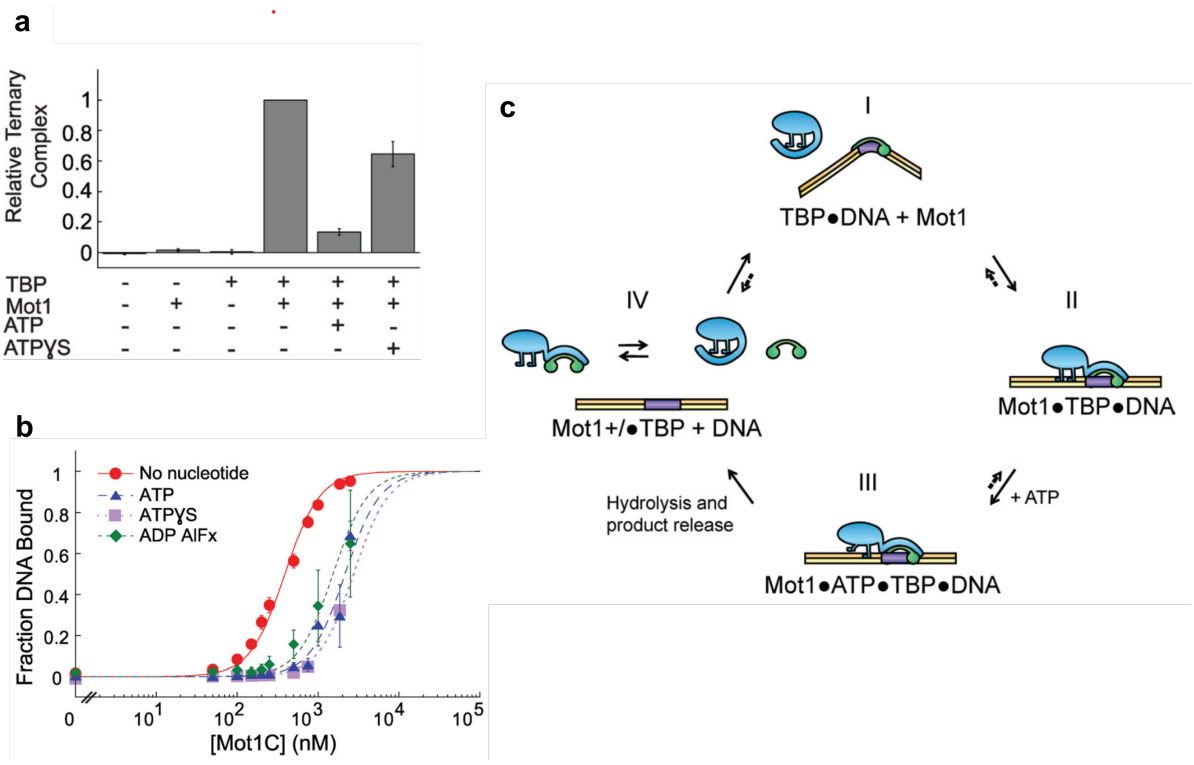


Fig 2.5.1: Affinity of Mot1 to TATA-box DNA (modified from [104]).

a) Quantitative analysis of nucleotide-dependent relative Mot1:TBP:DNA complex stability based on native gel shifts. Error bars represent the standard error.

b) Quantitative analysis of gel-based nucleotide-dependent Mot1:DNA binding affinities.
c) "Two-step mechanism" of Mot1-induced TBP:DNA disruption. (I) TBP binds DNA at the TATA-box and severely kinks it. (II) Binding of nucleotide-free Mot1 to TBP:DNA leads to unbending of the DNA. (III) A small conformational change takes place upon binding of ATP, but with downstream DNA still close to TBP (step 1). (IV) ATP-hydrolysis is necessary for complete disruption of TBP by Mot1 (step 2). Redistribution of TBP to another TATA-box or hand-off to other TBP-binding factors could reinitiate the circle.

Based on their findings, Moyle-Heyrman et al. proposed a "two-step mechanism" of Mot1-facilitated TBP remodeling. As a critical step, they monitored an additional conformational change upon ATP-binding and the necessity of subsequent ATP hydrolysis for complete complex dissociation (Fig. 2.5.1 c). Hence, they employed an in-solution FRET design, solely enabling to detect conformational equilibria derived from ensemble measurements[104]. Extending the technique, the Lamb group relied on single-pair FRET (spFRET) in a flow-cell, which allowed for measuring distances between the fluorophores and retrieve dynamic information[105].

The group confirmed an ATP-binding induced substantial conformational change and narrowed it down to an actual partial removal of Mot1:TBP from DNA. Because of the use of a flow cell, excess Mot1 that had not assembled as a Mot1:TBP:DNA complex could be washed out. Remarkably, Mot1 bound in the complex was not able to efficiently displace TBP from DNA, which could be reverted by washing Mot1 molecules into the solution. Even without extra ATP addition, the presence of nucleotide-free Mot1 accelerated the otherwise inefficient TBP displacement[105] (Fig. 2.5.2). Whether a molecular interaction between two Mot1 molecules drove this increase in remodeling efficiency or consecutive binding and rebinding events of monomeric Mot1 merely enhanced the likelihood of disruption could not be answered with their methods. To summarize, the Auble and Lamb groups arrived at a model of Mot1-induced TBP remodeling that comprises an unbending of DNA upon Mot1-binding at TBP:DNA, followed by a conformational change upon binding of ATP. But still, only ATP-hydrolysis leads to complete disruption of the ternary complex, with the requirement of other Mot1 molecules in solution for efficient TBP removal from DNA. In the light of our structural data, where we observed an inclination towards dimerization for all nucleotide-bound Mot1:TBP:DNA complexes (Chapter 5.2), the role of a potential oligomerization of Mot1 remains a conundrum, which will be discussed in Chapter 6.4.

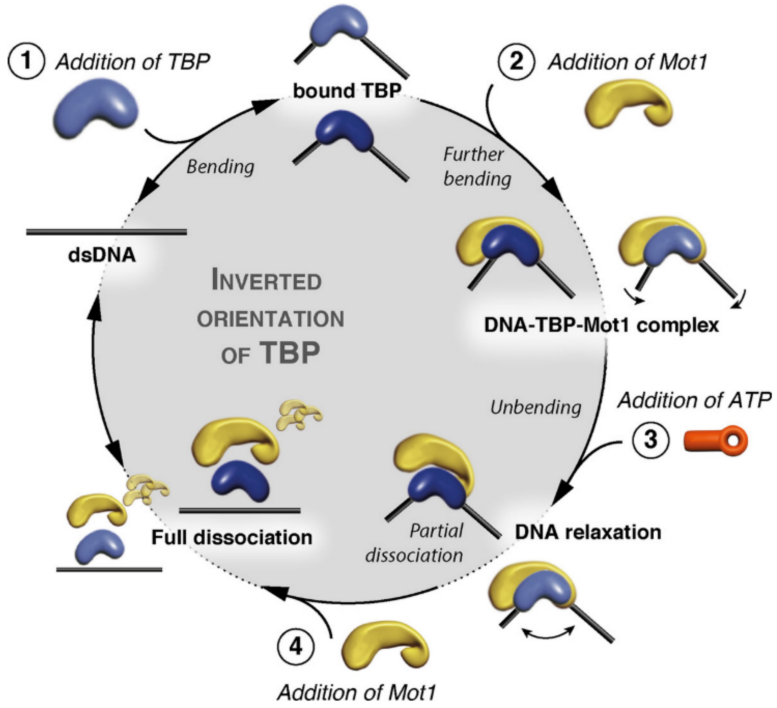


Fig. 2.5.2: Model of Mot1-induced TBP remodeling based on spFRET measurements (adapated from [105]).:
Two scenarios are illustrated: TBP bound in the correct orientation to DNA (beyond the circle) and TBP bound in an inverted orientation to the TATA motif at a different promoter (inner grey circle). (1) TBP binding creates a bending of DNA, (2) which is eventually increases upon Mot1 engagement. (3) Addition of ATP leads to DNA relaxation or partial release of Mot1:TBP from DNA respectively. (4) For an efficient complex dissociation, one or more Mot1 molecules in solution are needed without the necessity of extra ATP.

3 The INO80 A-module

3.1 Nucleosome positioning is facilitated by four remodeler families

The arrangement of eukaryotic chromatin as repetitive nucleosome arrays not only serves to reduce the spatial extension of DNA packed in the cell nucleus, but also is an essential feature of gene regulation[106]. The way nucleosomes are topologically organized influences the accessibility of the occupied DNA. For a long time, a hierarchical model of chromatin packaging was advocated, with a 10 nm beads-on-string arrangement of nucleosomes that reorganize into a denser 30 nm fiber[107] up to further degree of condensation ultimately arriving at condensed

mitotic chromosomes. This rather static model has been challenged especially by techniques that allow *in situ* visualization of chromatin in the nucleus[108,109]. Hence, the hierarchical model has been more and more replaced by a dynamic view of nucleosome organization, characterizing it as a fluid-like state where higher order compaction still occurs, but not manifested as static fibers of defined sizes[110]. The most compacted forms are constitutive heterochromatin, which comprises repetitive gene loci like pericentromeres and telomeres and facultative heterochromatin, which is closed in some cell types and variably opened for gene transcription[111]. Actively transcribed genes lie in euchromatin regions that allow dynamic nucleosome spacing, as the recruitment and assembly of the transcription machinery depends on accessible DNA, especially at the promoter region flanking the transcription start site (TSS)[112].

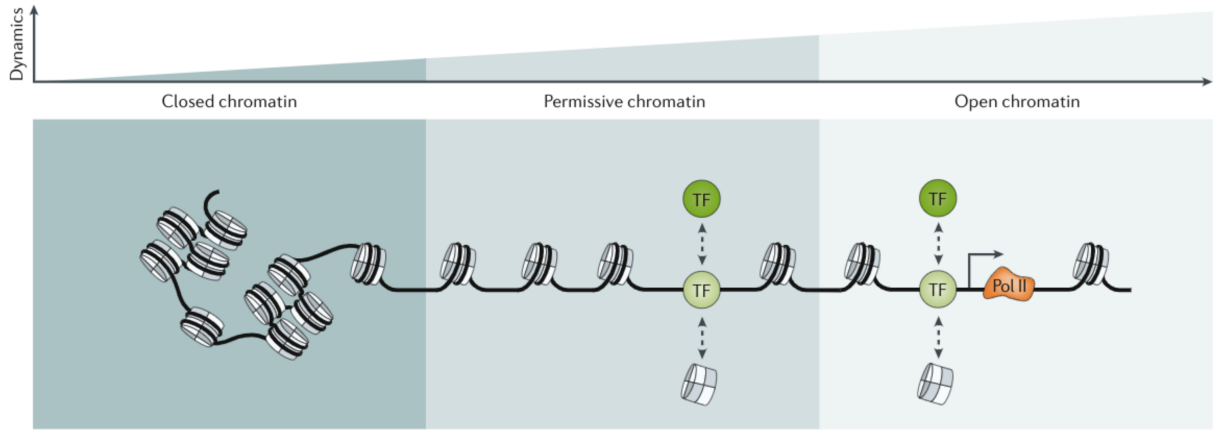


Fig. 3.1.1: Transient chromatin dynamics from closed to open chromatin (adapted from [112]).

The dynamic nature of chromatin reflects the multifactorial influence on the establishment of nucleosome-free regions. Establishment of such a region is initiated at permissive chromatin, where pioneer transcription factors bind to nucleosomes and general transcription factors with sequence affinity to naked DNA compete with nucleosome occupation[113] (Fig. 3.1.1). Thus, access to DNA results from a complex interplay of transcription factors, histone readers and writers[114,115] and nucleosome remodeling complexes.

Four families of nucleosome remodelers organize the positioning and composition of nucleosomes. The latter are classified into four families playing a decisive role in shaping the overall chromatin landscape and the establishment of nucleosome variants and spacing and phasing (Fig. 3.1.2). The SWI/SNF family is – in addition to the capability to create spaced nucleosome arrays – able to eject nucleosome core complexes and thus plays an important role

in regulating the overall nucleosome density. The INO80/SWR1 family is essential for transcriptional regulation, but also for replication and DNA repair[116,117]. SWR1 replaces the H2A histone with the H2A.Z histone variant that amongst other signaling functions marks transcriptionally poised genes[118]. Reports that propose a role for INO80 in reversing this reaction remain a matter of debate[119,120]. The INO80 complex establishes a nucleosome-free region around the TSS by placing the bordering nucleosomes up- (-1) and downstream (+1) of the TSS and creating a regular spacing of the proximal nucleosome arrays. Intriguingly, INO80 is the only remodeler capable to form nucleosome arrays *in vitro* that closely resemble native nucleosome arrays, as exemplified in yeast[121].

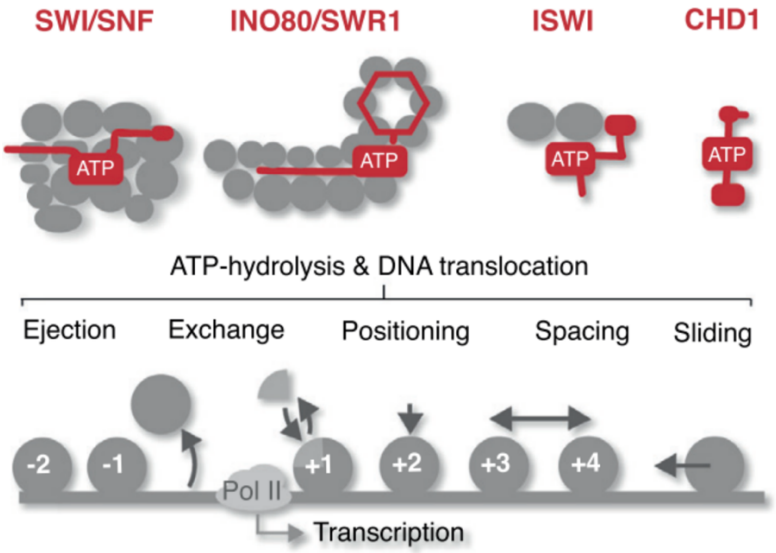


Fig. 3.1.2: Overview and functions of the four remodeler families (adapted from [67]). Top: Schemes indicating the architecture of the four remodeler families with the central ATPase colored red and additional subunits colored grey. Bottom: Schematic nucleosome array around the TSS. Arrows indicate diverse remodeling activities.

The ISWI complexes feature a HAND-SANT-SLIDE (HSS) domain that acts as molecular ruler able to detect a neighboring nucleosome as a barrier and instantiate arrays with defined inter-nucleosome distances[122]. CHD1 family remodelers detect extranucleosomal DNA with their DNA-binding domain (DBD)[51] (see Chapter 1.4), where Chd1 acts in concert with polymerase II and the FACT histone chaperone to facilitate transcription through nucleosomes[52].

In contrast, SWI/SNF and INO80/SWR1 remodelers are mega-Dalton assemblies of over 15 subunits. In these families, the ATPase functions as the motor of DNA translocation that connects to scaffolding proteins, which form a binding hub for diverse regulatory subunits. This modular architecture of the human SWI/SNF homologues, the BAF complexes, gives rise to a myriad of different combinations dependent on the cellular and developmental context[123] – another example of the increasing complexity of transcriptional regulation during evolution.

3.2 Actin and actin-related proteins form versatile building blocks

Both SWI/SNF and INO80/SWR1 feature a so-called A-module: nuclear actin (N-actin) and/or actin-related proteins (Arps) forming a stacked formation that residues on an N-terminal extension of the ATPase, a helical scaffold denoted HSA (helicase-SANT-associated) domain[67,124]. In the case of INO80, the HSA engages with longer (around 40 bp) stretches of extranucleosomal entry DNA via an array of basic residues in addition to eventual DNA binding moieties in the Arps[125]. Affinity enrichment mass spectrometry with a nanobody directed against nuclear N-actin from yeast whole-cell extract confirmed the A-module composition of INO80, SWR1 and the histone acetyltransferase complex NuA4[125] (Fig. 3.2.1 a). The combination of an Arp4/N-actin pair is common in all three complexes, but only INO80 features trimeric A-module assembly with an additional Arp8 occupying the free side of N-actin opposite of Arp4[125] (Fig. 3.2.1 b). Another Arp, Arp5, acts as part of the INO80 core forming a counter grip opposite of the motor ATPase[62,126].

Divergently, SWI/SNF complexes comprise the basic A-module subunits Arp7:Arp9 instead. It appears that during evolution, different combinations of Arps and actin developed as versatile building blocks, but how do they differ structurally? First, Actin and Arps comprise a conserved basic core made of two similar subdomains (SD) 1 and 3. Because of their similarity – a five-stranded beta sheet with an intermittent alpha helix - SD 1 and 3 most probably developed by gene duplication. An ATP-binding site is located at the interface between the SDs[127,128]. Two smaller subdomains (SD 2 and 4), which are stacked onto SD 1 and 3 respectively exhibit structural variability, constituting the main moiety of structural diversity between Actin and Arps[129].

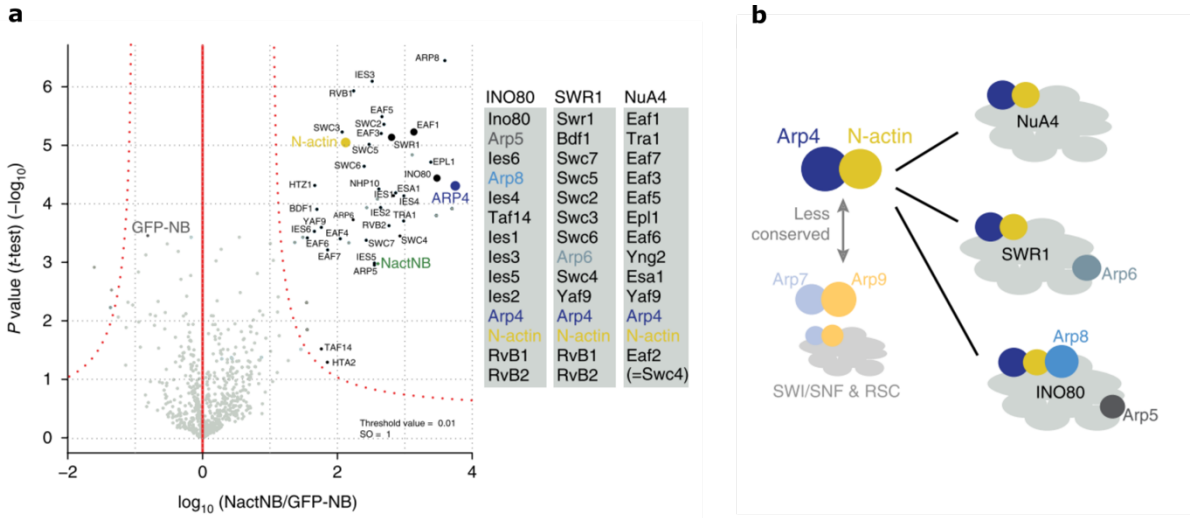


Fig. 3.2.1: A-module composition of remodeling complexes (adapted from [125]).
a) Variety of A-modules in four remodelers.
b) Affinity enrichment mass spectrometry analysis of a yeast whole-cell extract pulldown with n-actin nanobody (NactNB) as bait and GFP-nanobody (GFP-NB) as control. The volcano plot shows enrichment of all known subunits of INO80, SWR1 and NuA4. The assay was performed in triplicates and analyzed via t-test (p-value: 0.01).

The asymmetry between SD 3/4, forming a barbed end and SD 1/2, forming a pointed end results in a polarity that directs polymerization of globular actin (G-actin) (Fig. 3.2.2 a) to filamentous actin (F-actin) in the cytoplasm[130]. Arps differ in their subdomain structure by additional extensions in the barbed-end subdomains, which changes their surface and therefore create versatile possibilities for protein or nucleic acid interactions (Fig 3.2.2 b).

A structure of the yeast Swr1 HSA with Arp4-N-Actin revealed first insights into their structural arrangement as a regulatory module[131]. Instead of a “front-to-front” alignment as seen for F-actin filaments in the cytoplasm[132] (Fig. 3.2.2 c), Arp4 and nuclear actin (N-actin) interact in “front-to-back” and the cavity formed by their barbed ends forms a hydrophobic interaction with the HSA helix[131,133] (Fig. 3.2.2 d). When combined with Arps in a so-called A-module, polymerization of N-actin is sterically hindered, which allows its utilization as a building block and binding hub[134].

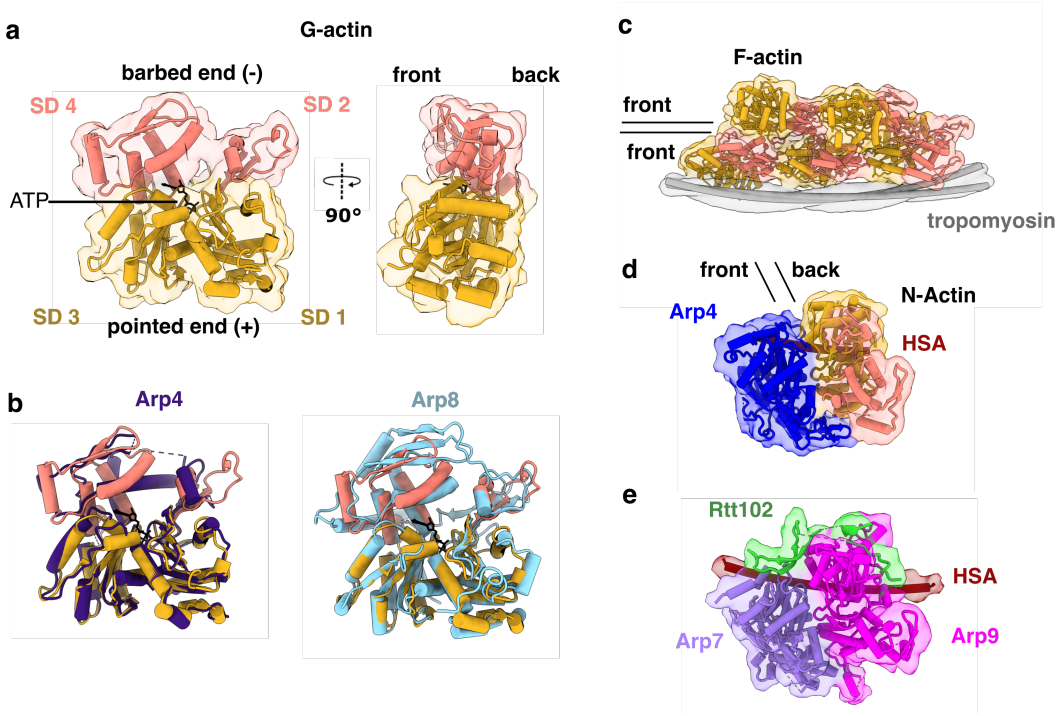


Fig. 3.2.2: Structural comparison of actin and actin-related protein assemblies.
a) Crystal structure of globular actin (G-actin) isolated from rabbit skeletal muscles (PDB 1ATN) with subdomains (SD) and topology indicated.
b) Alignments of human Arp4 (PDB 6LTJ) and Arp8 (PDB 4FO0) structures with the G-actin structure from a).
c) F-actin:tropomyosin structure (PDB 3J8A) showing the front-to-front arrangement in muscular actin filaments.
d) Arp4:N-actin bound to the HSA (PDB 5I9E) from yeast SWR1 complex showing a front-to-back arrangement.
e) Arp7:Arp9:Rtt102 bound to the HSA (PDB 4I6M) from yeast RSC complex.

Basically, actin with its conserved subdomain structure served as an evolutionary starting point that culminated in the variety of surface structure seen in the nuclear Arps 4 – 9. In SWI/SNF complexes, instead of N-actin and Arp4, Arp7 and Arp9 arrange analogously[135] (Fig. 3.2.2 e). This reflects the evolutionary stability of overall A-module architecture in remodeler complexes, despite the divergence in subunit composition. The Arp7:Arp9:HSA structure is appended by Rtt102, tethering the two Arps proximal to their HSA underpinning[135]. Although mostly unstructured, Rtt102 is anchored via a β-hairpin spanning the cleft between Arp7 and Arp9. Its biochemical relevance remains obscure, but it proposedly serves the role of a stabilizing subunit, because its deletion impaired complex expression and led to a reduced-growth yeast phenotype[135]. Early on, the A-module has been identified as a critical component of SWI/SNF and

INO80 family chromatin remodelers, as well as for the related NuA4 histone acetyltransferase[136]. In remodelers, its deletion has severe diminishing effects on the ATPase activity and ATP-dependent DNA translocation, hence on the overall remodeling activity and – in the RSC complex – nucleosome ejection[137]. Again, it was the combination of structural and biochemical data that brought light into the molecular intricacies of the interplay between ATPase and A-module in chromatin remodelers, especially regarding the characterization of the INO80 complex[62,125,138].

3.3 The INO80 A-module allosterically regulates remodeling activity

In contrast to the heterodimeric A-modules in SWI/SNF remodelers, the INO80 A-module features an additional Arp8 that unusually orientates "side-to-front" towards N-actin, yielding an extended binding interface and a stabilized three-subunit complex[125]. Even more, Arp8 binding proved to be indispensable to A-module integrity. Deleting it or the HSA-domain led to suppression of remodeling activity[139]. By solving the crystal structure of *S. cerevisiae* Arp8:N-actin:Arp4 complex bound to the HSA-domain, Knoll et al. showed that the trimeric complex assembles via their barbed ends on a segmented α-helical HSA domain – a "two-plug scaffold" instead of a continuous helix[125].

Previous structures elucidated the architecture of the INO80^{core} (C-module) complex[62,126]: the Swi2/Snf2 motor ATPase harbors an INO80 family specific insertion between RecA1 and RecA2, which is clamped between two RuvB-like hexamers. The counter-grip subunits Arp5/Ies6 and the histone acidic patch-sensing subunit Ies2 assemble at the platform established by the RuvB1/2 scaffold and ATPase. The Ino80 subunit can be interpreted as a central scaffold, where all other INO80 subunits associate with (Fig. 3.3.1 a, top). N-terminally the post-HSA domain is followed by the HSA-domain with the A-module subunits and – in the case of yeast INO80 – Ies4 and Taf14[139]. In human INO80, the latter two subunits are missing and instead, the transcription factor YY1 is an integral part[140]. At the outermost N-terminus a set of subunits exerts species-specific functions, e.g., recruitment of INO80 to genomic sites and - as shown for yeast N-terminal module (N-module) – acts as part of a "molecular ruler" that senses distances to general transcription factors or between nucleosomes, thus directing nucleosome positioning[63,141].

Arp4 and N-actin bind to the N-terminal HSA-helix (HSA α1), whereas Arp8 resides on the C-terminal HSA-helix (HSA α2) (Fig. 3.3.1 a, bottom). When analyzing the biochemical activity of a INO80 C- and A-module complex, it became evident that both HSA helices are necessary for efficient nucleosome sliding *in vitro* and that complete HSA deletion abrogated the sliding reaction completely (Fig. 3.3.1 b). The same truncations yielded analogous results in yeast

genome-wide nucleosome positioning assays, with much weaker nucleosome occupancies around the TSS [125](Fig. 3.3.1 c).

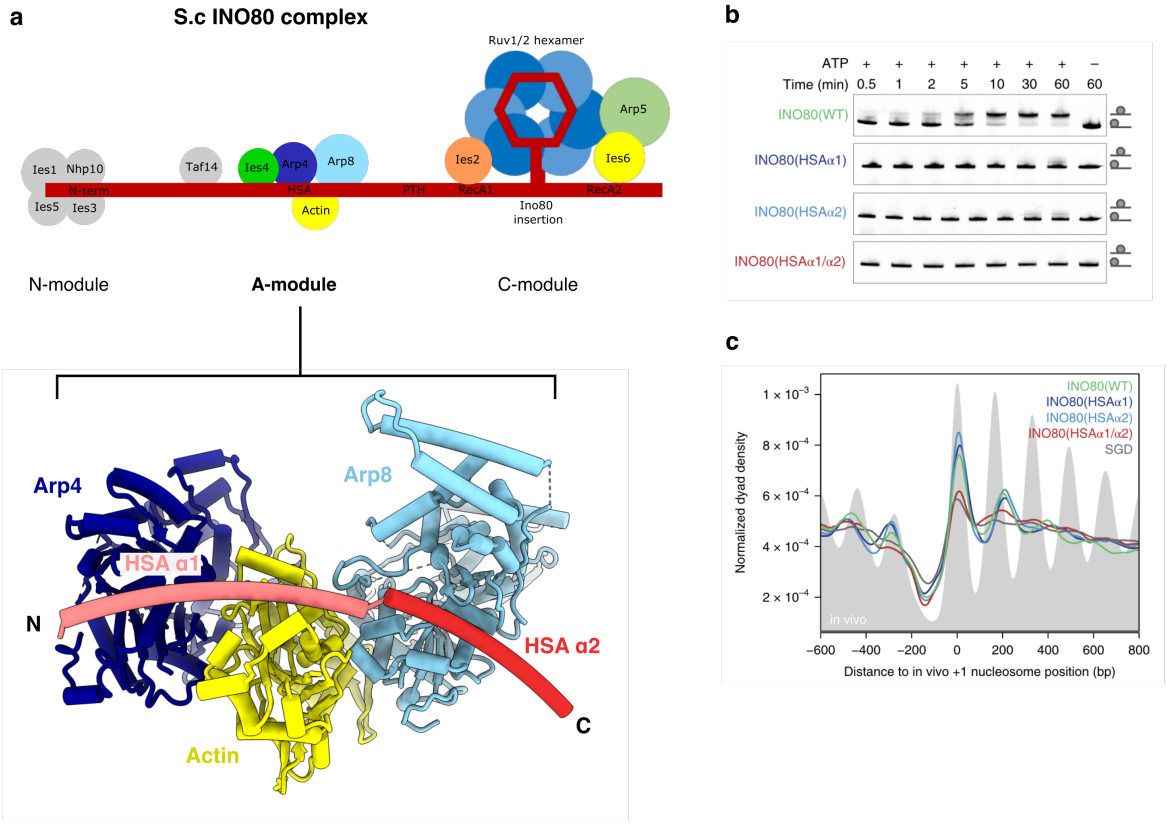


Fig. 3.3.1: Composition of INO80 and structure and biochemical activity of the A-module.
a) Top: Topology the *Saccharomyces cerevisiae* INO80. Bottom: crystal structure of the *Sc*A-module (PDB 5NBN).
b) Gel-based sliding assay showing the effect of HSA truncations on nucleosome remodeling activity (adapted from [125]).
c) Genome-wide nucleosome positioning assay showing the effect of HSA truncations of INO80 (adapted from [125]).

But how is the A-module placed in the context of a nucleosome-bound INO80? In a previous low-resolution cryo-EM reconstruction, the A-module was located aligning with extranucleosomal entry side DNA[62] (Fig. 3.3.2).

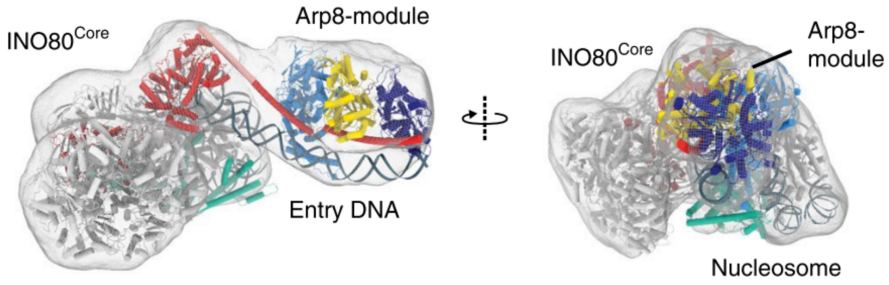


Fig. 3.3.2: Rotated views of a model of INO80^{core} with A-module (Arp8 module) docked into a low-resolution cryo-EM density (adapted from [125]).

Binding of the entry DNA overhang was confirmed by competition EMSAs of A-module towards 0N0 and 0N80 (80 bp DNA overhang at the entry side), which showed a clear preference for the latter. Additionally, almost over the entire length the HSA bears a sequence of basic residues to potentially contact the DNA phosphate backbone. Mutating the lysine to glutamine residues abolished DNA binding completely, showing that the HSA indeed engages extranucleosomal DNA[125].

Summarizing, the A-module emerged as a decisive allosteric regulator of INO80 remodeling activity and with recent structural and functional data we are now beginning to understand the intricacies of its regulatory influence. While it was observed that the INO80 N-module regulates nucleosome positioning in a DNA linker length dependent way[63,141], the A-module couples ATPase activity and remodeling events. But how is this coupling implemented on a structural level and is DNA itself influencing ATP-dependent nucleosome positioning?

3.4 The A-module is a DNA shape sensor

The generation of yeast whole-genome libraries, their reconstitution with nucleosomes via salt-gradient dialysis (SGD) and subsequent MNase digest combined with paired-end sequencing allowed a genome-wide analysis of nucleosome positioning preferences by INO80. One insight was the beforementioned “molecular ruler” feature conferred by N- and A-module, that yields formation of evenly spaced arrays, aligned to barrier factors such as GRFs or DNA ends[141]. However, DNA-length-dependent nucleosome alignment does not explain the strong positioning of the +1 nucleosome downstream of the TSS with its long nucleosome-depleted region, even in the absence of barrier factors[121]. Beyond a proposed influence of sequence motifs,

Oberbeckmann et al. followed the intriguing idea, that DNA shape features (e.g., rigidity, torsion, twist) determined by certain DNA sequences might conduct placement of nucleosomes by INO80. Combining structure guided INO80 mutations, biochemistry and statistical analysis, they correlated nucleosome occupancies conducted by INO80 with local DNA shape features[142]. SGD nucleosomes with their intrinsic histone octamer preferences do not facilitate formation of NDRs dependent on DNA shape *in vitro*, but instead are naturally guided by nucleosome-positioning sequences that exhibit a propensity for bending around the nucleosome core. In contrast, DNA shape apparently directs INO80-dependent strong NDR and +1 nucleosome positioning. When aligning DNA at the INO80 positioned linker-+1-nucleosome with a nucleosome-bound INO80 model, two contrasting DNA shape regions were identified. First, the DNA bound to the A-module exhibits a large propeller twist, which is also indicative of high torsion. Second, the DNA between motor ATPase and Arp5 exhibits low propeller twist[142] (Fig. 3.4).

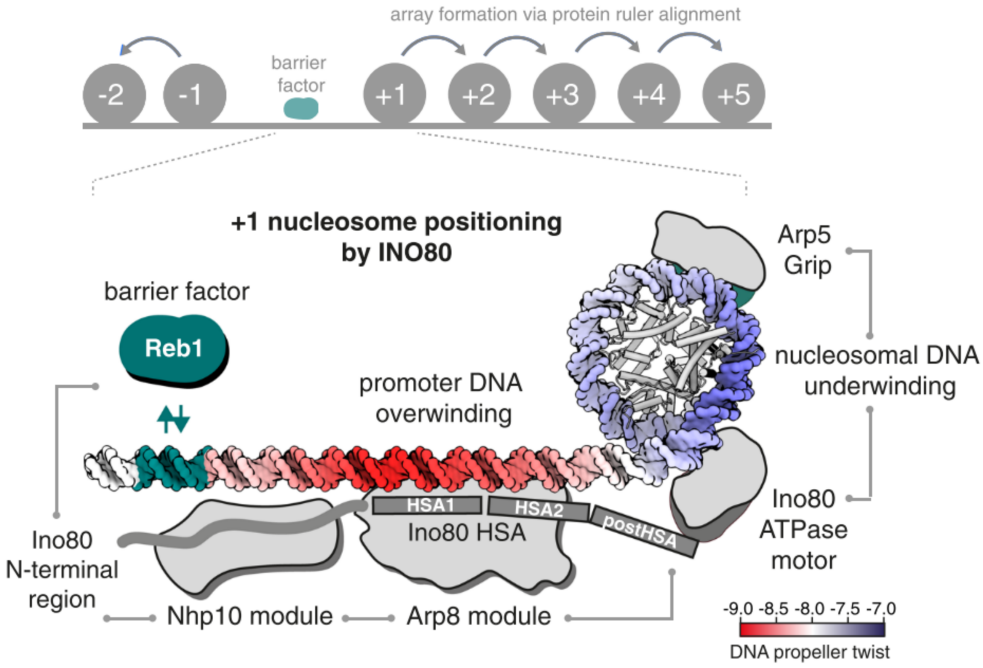


Fig. 3.4: Scheme of barrier factor and DNA shape influence on nucleosome positioning downstream of the TSS by INO80 (adapted from [142]).

Nucleosome arrays can be aligned to a barrier factor e.g., GRFs like Reb1. The distance to the barrier is read-out by the extended N- and A-modules (in yeast: Nhp10 and Arp8 module). The A-module could have a binding preference for overwound extranucleosomal DNA, while the DNA between ATPase motor and Arp5 is underwound, which would favor accumulated torsion by bulge formation due to DNA pumping.

In the light of these findings, it can be proposed that INO80 regulation by DNA shape acts via two allosteric axes: first, by an increased binding affinity of the A-module to twisted overwound DNA and second, by underwound DNA between Arp5 and ATPase motor that is inclined for bulge formation upon DNA pumping. Integration of distance information towards neighboring nucleosomes, barrier factors or DNA ends via the N- and A- module molecular ruler propensity results in a nucleosome positioning outcome that is highly dependent on the (epi)genomic context.

3.5 The A-module delivers feed back to the ATPase via a regulatory hub

Beyond the insight that extranucleosomal DNA-binding by the INO80 A-module positively couples ATPase activity and remodeling remains the question how this coupling manifests on a structural level. It was observed for the RSC remodeler ATPase, Sth1, that a functional interdependence between the post-HSA domain and protrusion I exists[143]. This was not surprising, as the A-module conformation relays to the post-HSA, which engages in a regulatory interaction with the ATPase[144]. Because binding of Arp7:Arp9:Rtt102 to the HSA destabilized the post-HSA:protrusion I interaction, it was proposed that the post-HSA, in a productive interaction with protrusion I, normally restrains the remodeling activity by decoupling effective DNA translocation from ATPase activity. Disruption of the protrusion I:post-HSA interaction by binding of the A-module to extranucleosomal DNA might allow the ATPase to obtain a fully closed conformation competent for DNA translocation. In an attempt to narrow down the involved protein domains, three point mutations were identified in a Sth1 ΔARP suppressor mutant, of which two localize at protrusion I. Strikingly, these mutants (E676Q and L681F) led to a gain-of-function phenotype – an increase in nucleosome sliding and ejection[143] (Fig. 3.5 a). As observed before[144], the N-terminal protrusion I helix contacts the N-terminal post-HSA domain, which regulates remodeling efficiency meaning a change in the remodeling reaction with no significant effect on ATPase activity. The other functional half of this regulatory hub consists of the C-terminal protrusion I helix contacting the brace, which implements the remodeling activity *per se*, as mutations perturbing the interaction eliminated nucleosome sliding[33]. Because these findings were based on mutational interaction studies, a definite structure-function causality could not be drawn, which would have required to compare high-resolved remodeler structures in different ATPase and A-module conformations. Nevertheless, structural support for the functional results came from Baker et al.: a four-helix-bundle interaction of the two protrusion I helices, as well as HSA and post-HSA domains was confirmed by cryo-EM studies investigating dynamics and conformational space of the A-module of the RSC complex[32] (Fig. 1.3.3 b, c). Here – by

comparison with other Swi2/Snf2 remodeler structures - they noticed a conformational shift of protrusion I dependent on the nucleotide state and stabilization of the second brace helix only in ADP or apo bound states, but not in ADP-BeF_x bound states. Apparently, this so-called regulatory hub turned out to be a culmination point of two-sided conformational feedback: substrate-binding induced conformational changes by the A-module delivered via HSA/post-HSA and, *vice versa*, nucleotide-state induced changes relayed towards the A-module.

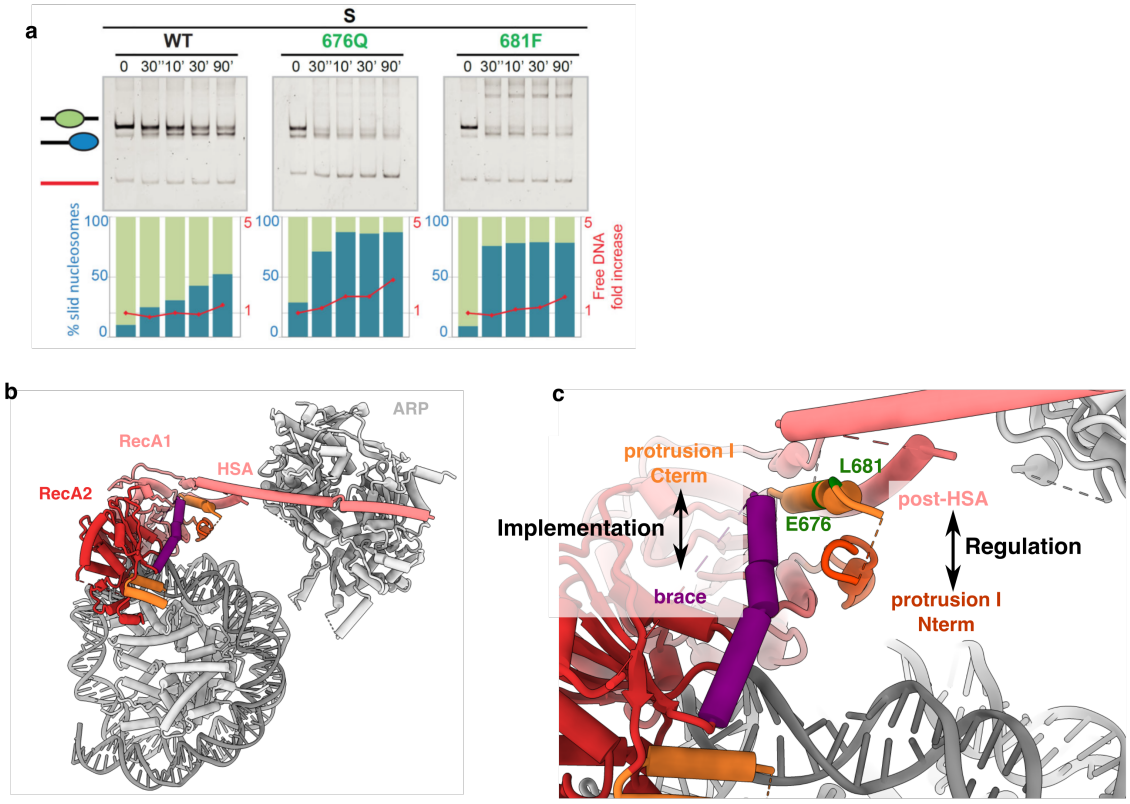


Fig. 3.5 Regulatory hub of the Sth1 ATPase.
a) Nucleosome sliding activity as monitored by gel shifts with the DNA fluorescently labelled. Two gain-of-function mutations (E676Q and L681F) on the C-terminal protrusion increase sliding (adapted from [143])
b) Structure of the partial RSC complex comprising the Sth1 ATPase (colored) bound to a nucleosome including the N-terminal HSA domain with ARPs (PDB 6VZ4)
c) Close up on the Sth1 four-helix regulatory hub in b) and proposed functional attributes influencing remodeling according to [33].

4 Objectives

Although acting on different chromatin targets, the multi-subunit INO80 nucleosome remodeler and the single-subunit transcriptional regulator Mot1 both belong to the Swi2/Snf2 ATPase family. Their ATPase domains share a conserved core comprising two RecA lobes with an ATP-hydrolysis dependent DNA minor groove tracking activity. Both enzymes are regulators of chromatin components, thereby interfering with information processing at the DNA level and shaping the chromatin landscape. While the multi-subunit INO80 complex positions nucleosomes as equally spaced arrays, Mot1 dissociates TBP from DNA enabling TBP recycling and distribution between promoter sites.

One aim of this thesis was to structurally and biochemically characterize Mot1 acting on its TBP:DNA substrate. Five cryo-EM reconstructions shed light on the stepwise conformational trajectory of the TBP removal from DNA and allowed to identify critical structural elements. Importantly, we could define the function of all subdomains of the C-terminal bridge element and concisely lay out its role as an essential allosteric regulator of the remodeling reaction. The derived conformational trajectory reveals a concerted interplay of non-processive DNA translocation, accompanied by rotational dislocation of the TATA-box, a pivot movement of the Mot1 N-terminal domain and competitive occupation of the TBP DNA-binding site by a hook element.

Although the Ino80 ATPase is allosterically regulated as well, this is implemented in a more complex manner by the so-called A-module, comprising two actin-related proteins (Arp8 and Arp4) sandwiching an actin monomer – assembled on a helical HSA-domain. The A-module recognizes extranucleosomal DNA dependent on its shape parameters determined by the sequence and acts as a molecular ruler element against barriers such as general regulatory factors (GRFs) or DNA ends. We reconstructed cryo-EM structures of the A-module in the context of INO80 bound to the nucleosomes and extranucleosomal DNA from three species (*Saccharomyces cerevisiae*, *Chaetomium thermophilum* and *Homo sapiens*). This allowed an interspecies comparison, and the structural identification of novel DNA-binding domains. Further, we discovered a conserved inter-molecular anchor motif, a two-tryptophan bearing β-hairpin structure, that facilitates the interaction between Arps/N-actin and so-called client proteins (Ies4 and YY1), which are involved in the recruitment of INO80 to its genomic target sites.

In summary, by applying state-of-the art cryo-EM combined with functional analyses of the small one-subunit remodeler Mot1 and the mega-Dalton multi-subunit complex INO80, we could

derive details of Swi2/Snf2 enzyme regulation by characterizing regulatory domains and elucidating their conformational space.

5 Publications

5.1 Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state

Agata Butryn, Stephan Woike, Savera J Shetty, David T Auble, Karl-Peter Hopfner. Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state, *eLife* 7:e37774, 5th Oct 2018.

Summary

This publication reveals the crystal structure of the near full-length Mot1 from *Chaetomium thermophilum*. So far, only the N-terminal domain of Mot1 was structurally characterized. Here, because no nucleotide was added prior to crystallization, the C-terminal Swi2/Snf2 ATPase domain is in an open, autoinhibited resting state. The two RecA-like lobes of the ATPase are disarranged by a 180° rotation of RecA2 relative to RecA1, which prevents formation of a nucleotide binding pocket. The structure shows, how RecA1 incorporates at the gap of the N-terminal HEAT-repeat arch between HEAT repeat 1 and 16 and supports a spiral-like geometry of the Mot1 N-terminus. Besides the crystal structure, the publication comprises NADH-coupled ATPase assays and electrophoretic mobility shift assays (EMSAs) to analyze the TBP disruption from the TATA-motif on DNA (denoted “remodeling”) and to investigate the coupling between ATPase rate and remodeling activity. A series of point mutations on RecA1 to disrupt the resting state showed an increase in basal ATPase rate, but binding of the TBP:DNA substrate had a larger effect. Deletion of the Mot1 outermost C-terminal bridge domain resulted in enhanced ATPase activity, but with diminished Mot1-induced TBP dissociation from DNA. Thus, the bridge domain was identified to play a decisive role in the allosteric coupling between ATPase activity and remodeling.

Author contribution

I established and conducted the EMSAs under native conditions for a gel-based readout of the TBP:DNA dissociation by wildtype Mot1 and diverse mutants and quantified the gel bands.

eLIFE elifesciences.org

RESEARCH ADVANCE

Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state

Agata Butryn[1,2†], Stephan Woike[1,2], Savera J Shetty[3], David T Auble[3], Karl-Peter Hopfner[1,2,4*]

[1]Department of Biochemistry, Ludwig-Maximilians-Universität München, Munich, Germany; [2]Gene Center, Ludwig-Maximilians-Universität München, Munich, Germany; [3]Department of Biochemistry and Molecular Genetics, University of Virginia Health System, Charlottesville, United States; [4]Center for Integrated Protein Sciences Munich, Munich, Germany

Abstract Swi2/Snf2 ATPases remodel protein:DNA complexes in all of the fundamental chromosome-associated processes. The single-subunit remodeler Mot1 dissociates TATA box-binding protein (TBP):DNA complexes and provides a simple model for obtaining structural insights into the action of Swi2/Snf2 ATPases. Previously we reported how the N-terminal domain of Mot1 binds TBP, NC2 and DNA, but the location of the C-terminal ATPase domain remained unclear (***Butryn et al., 2015***). Here, we report the crystal structure of the near full-length Mot1 from *Chaetomium thermophilum*. Our data show that Mot1 adopts a ring like structure with a catalytically inactive resting state of the ATPase. Biochemical analysis suggests that TBP binding switches Mot1 into an ATP hydrolysis-competent conformation. Combined with our previous results, these data significantly improve the structural model for the complete Mot1:TBP:DNA complex and suggest a general mechanism for Mot1 action.

DOI: https://doi.org/10.7554/eLife.37774.001

***For correspondence:** hopfner@genzentrum.lmu.de

Present address: †Diamond Light Source Limited, Harwell Science and Innovation Campus, Didcot, United Kingdom

Competing interests: The authors declare that no competing interests exist.

Funding: See page 9

Received: 01 May 2018
Accepted: 04 October 2018
Published: 05 October 2018

Reviewing editor: Geeta J Narlikar, University of California, San Francisco, United States

Introduction

Swi2/Snf2 ATPases are members of the NTP-dependent helicase/translocase superfamily 2 (SF2) and are well known as the principal ATP hydrolyzing 'engines' of chromatin remodelers that govern processes such as transcription, replication, and DNA repair (***Flaus et al., 2006***; ***Narlikar et al., 2013***; ***Hopfner et al., 2012***; ***Becker and Workman, 2013***). It is generally assumed, that the Swi2/Snf2 ATPase motor translocates on the minor groove of double-stranded DNA and that this universal core activity generates the force for the large diversity of remodeling reactions catalyzed by Swi2/Snf2 proteins (***Saha et al., 2002***; ***Whitehouse et al., 2003***; ***Zofall et al., 2006***; ***Dürr et al., 2005***). However, very little is known about how groove tracking activity is converted into the diverse chemo-mechanical remodeling reactions (***Hauk and Bowman, 2011***; ***Narlikar et al., 2013***; ***Blossey and Schiessel, 2018***). In the absence of substrates, remodelers have been observed in catalytically inactive resting states (***Hauk et al., 2010***; ***Xia et al., 2016***; ***Yan et al., 2016***), but it is unclear how universal auto-inhibited resting states are. Recent work provides some insight into how Swi2/Snf2 chromatin remodelers interact with and reconfigure nucleosomal substrates (***Liu et al., 2017***; ***Farnung et al., 2017***; ***Ayala et al., 2018***; ***Eustermann et al., 2018***; ***Sundaramoorthy et al., 2018***). However, the architecture and chemo-mechanical mechanisms of the diverse types of remodeling reactions are not well understood for the great majority of enzymes in this class.

The single subunit remodeler Mot1 (Modifier of transcription 1) is a Swi2/Snf2 enzyme that either activates or represses transcription in a context-dependent manner by dissociating TATA box-binding protein (TBP) and Negative Cofactor 2 (NC2) from promoter DNA (***Dasgupta et al., 2002***; ***Zentner and Henikoff, 2013***). Mot1 is an essential Swi2/Snf2 enzyme in yeast and the Mot1-TBP-

NC2 regulatory axis is highly conserved in eukaryotes. The 140 – 210 kDa Mot1 protein has two functional domains. The N-terminal domain (Mot1^{NTD}) recognizes TBP while the C-terminal domain (Mot1^{CTD}) contains the catalytic Swi2/Snf2 ATPase that binds the DNA upstream from the TATA box (**Moyle-Heyrman et al., 2012**). The structure of the Mot1^{NTD} has been determined by X-ray crystallography in complex with TBP (**Wollmann et al., 2011**) and in complex with TBP:NC2:DNA (**Butryn et al., 2015**). Low-resolution negative stain electron microscopy and chemical crosslinking coupled to mass spectrometry indicated the approximate location of the Mot1^{CTD} near the opening of the Mot1^{NTD} horseshoe (**Butryn et al., 2015**). However, the orientation of the Swi2/Snf2 domain and consequently the path of DNA remained elusive. As a result, the mechanism of Mot1-mediated dissociation of TBP complexes is still not well understood.

Here, we report the crystal structure of the near full-length Mot1 protein from *Chaetomium thermophilium*. Our structure reveals the location and orientation of the Swi2/Snf2 domain and, supported by mutagenesis studies, suggests a new type of resting state. Our data allow us also to derive a model for the Mot1 remodeler in complex with TBP and DNA.

Results and discussion

Architecture of *Ct*Mot1

We crystallized the near full-length Mot1 protein from *Chaetomium thermophilum*. The construct covers the entire Mot1^{NTD} and Mot1^{CTD} domains but lacks 50 amino acids from the C-terminus (**Figure 1A**). We determined the structure of this construct (residues 1–1836, Mot1$^{\Delta C}$) harboring a point mutation in the Walker B motif (E1434Q) by Se-SAD to 3.2 Å (**Table 1**).

The *Ct*Mot1 enzyme is a ring-shaped protein (**Figure 1B**). The *Ct*Mot1^{NTD} consists of 16 HEAT repeats (HR) with insertions at four sites and is similar to the much smaller *Encephalitozoon cuniculi* orthologue (*Ec*Mot1^{NTD}) with some notable differences. The helices forming the HEAT repeats are not extended in number but in length and the insertion elements into the HEAT repeats are longer. Thus, genome compaction in *E. cuniculi* did not alter the overall architecture of Mot1, which appears to be highly conserved in evolution, consistent with its critical function.

The structure reported here is the first to visualize the position and orientation of the Swi2/Snf2 ATPase domain in Mot1. The ATPase domain contains two characteristic lobes connected by a short hinge helix. Each lobe consists of a RecA-like subdomain (1A or 2A) that harbors the SF2-specific sequence motifs responsible for ATP and DNA binding as well as Swi2/Snf2-specific helical subdomains 1B, 2B, and 'brace' that emanate from 1A, 2A and the C-terminus of 2A, respectively. Lobe 1 of the *Ct*Mot1^{CTD} contacts the C-terminus of *Ct*Mot1^{NTD} via HR16, a small insertion within HR12, and a ~ 45 amino acids linker. This highly hydrophobic surface has a total area of 2500 Å^2. The tip of subdomain 1B interacts with HR1, thus lobe 1 effectively closes the ring structure of *Ct*Mot1. The architectural constraints imposed by a ring explain the conservation of the number of HEAT elements among Mot1 proteins. Lobe 2 binds the cleft between lobe 1 and HR1/2. The ~1900 Å^2 interface between lobe 2 and the remainder of *Ct*Mot1 is dominated by hydrogen bonds and salt bridges.

In some remodelers, the brace is directly followed by a 'bridge' element (**Hauk et al., 2010**), also referred to as NegC (**Clapier and Cairns, 2012**) or SnAC (**Sen et al., 2011**; **Xia et al., 2016**). While *Ec*Mot1 does not possess this element, in *Ct*Mot1 it is 64 amino acids long (residues 1822 – 1886). The bridge can act as a positive or negative auto-regulatory element via mechanisms that are not understood (**Wang et al., 2014**; **Xia et al., 2016**; **Yan et al., 2016**; **Clapier and Cairns, 2012**; **Carroll et al., 2014**; **Sen et al., 2011**). The bridge was almost entirely omitted from our crystallization construct and the only included residues (1822 –1836) together with the C-terminal expression tag are not visible in the electron density maps.

In summary, the structure reveals the architecture of the *Ct*Mot1 protein. It forms a ring-like structure in which the substrate-interacting HEAT repeat 'arch' binds both lobes of the ATPase domain from opposing sites.

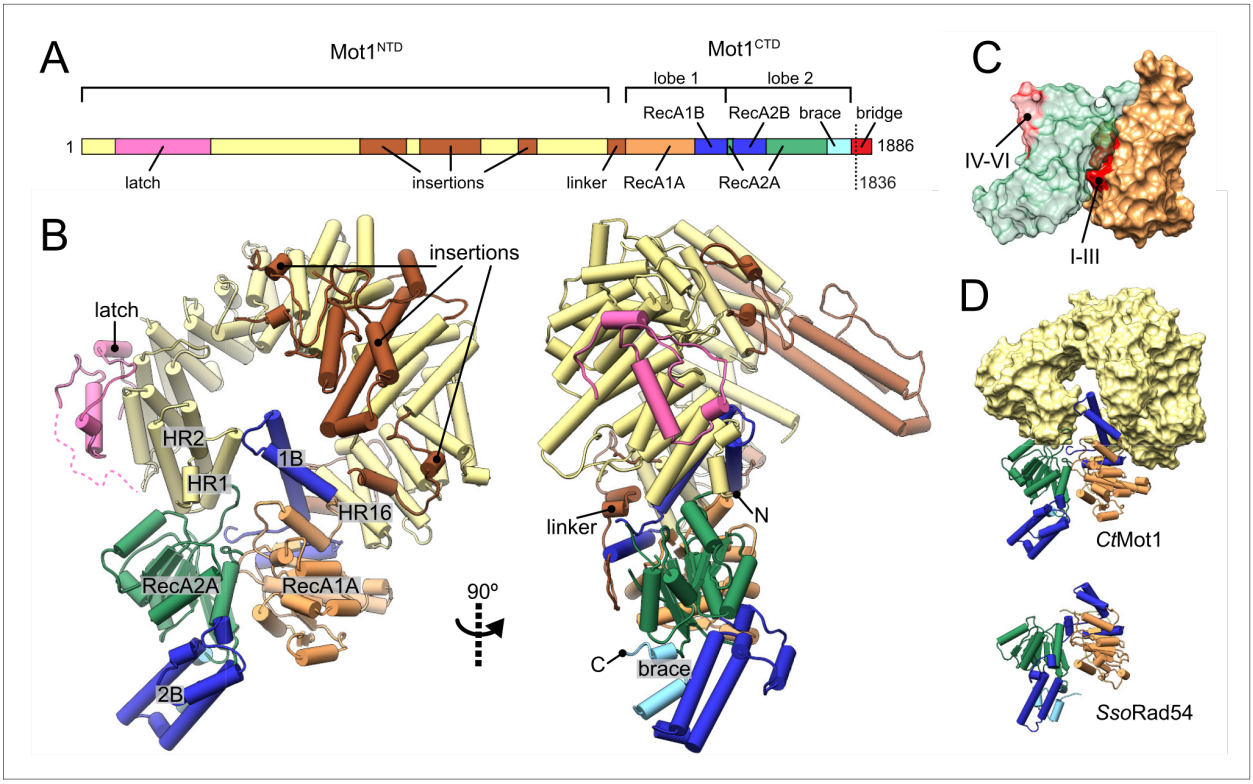


Figure 1. Structure of the *Chaetomium thermophilum* Mot1. (**A**) Domain organization of *Ct*Mot1. The HEAT repeats are in yellow. The latch is in pink, other insertions of *Ct*Mot1^{NTD} and the linker are in brown. RecA-like subdomains of *Ct*Mot1^{CTD} are in orange (1A) and green (2A). Swi2/Snf2-specific insertions 1B and 2B are in dark blue. Brace and bridge elements are in light blue and red, respectively. The boundary of the crystallization construct (residue 1836) is marked with the dotted line. (**B**) Cartoon representation of the structure. N- and C-termini are labelled N and C, respectively. HEAT repeats 1, 2, and 16 are labelled HR1, HR2, and HR16, respectively. Missing residues of the latch are represented by the dotted line. (**C**) Surface representation of *Ct*Mot1^{CTD} lobe 1 (orange) and 2 (green). Regions where helicase motifs are located on each lobe are colored in red. (**D**) Side-by-side comparison of *Ct*Mot1^{CTD} (top panel) and *Sso*Rad54 (***Dürr et al., 2005***) (bottom panel). *Ct*Mot1^{NTD} is represented as yellow surface. If not stated otherwise, all panels have color coding as in A.

DOI: https://doi.org/10.7554/eLife.37774.002

The following figure supplement is available for figure 1:

Figure supplement 1. Auto-inhibited conformations of Swi2/Snf2 domains.

DOI: https://doi.org/10.7554/eLife.37774.003

Apo *Ct*Mot1 adopts an auto-inhibited resting state

In all species tested (*H. sapiens, S. cerevisiae, C. thermophilum, E. cuniculi*), Mot1's ATPase is robustly activated by TBP:DNA complexes, but very little if at all by DNA alone (***Auble et al., 1997***; ***Adamkewicz et al., 2000***; ***Wollmann et al., 2011***; ***Chicca et al., 1998***). Interestingly, some Mot1 species are activated by TBP alone and do not require DNA, although a more robust activation is generally observed in the presence of both DNA and TBP. This suggests that the conformation of the Mot1^{CTD} is structurally coupled to TBP binding to the Mot1^{NTD} and that Mot1 alone is in an inactive state (***Adamkewicz et al., 2000***; ***Moyle-Heyrman et al., 2012***). Indeed, comparison of *Ct*Mot1^{CTD} to other SF2 enzymes shows that lobe 2 is flipped ~180° from an 'active' conformation in which the ATPase and DNA-binding motifs would be properly aligned, that is lobe 1's motifs I-III are properly situated in the ATP-binding cleft, while lobe 2's motifs IV-VI are situated on the outside and are fully solvent-exposed (***Figure 1C***). As more Swi2/Snf2 domain structures have become available, it has become evident that many show an auto-inhibited conformation with misaligned lobes 1 and 2

Table 1. Data collection and refinement statistics for the *Ct*Mot1 structure.

Data collection	
Space group	$P2_1$
Unit cell	
a, b, c (Å)	93.2, 96.9, 129.7
α, β, γ (°)	90.0, 97.6, 90.0
Resolution (Å)	48.7 (3.3–3.2)*
Total reflections	239071 (10913)
Unique reflections	36422 (1888)
R_{meas} [%]	14.4 (88.7)
$I/\sigma I$	11.8 (2.6)
$CC_{1/2}$	0.99 (0.79)
Completeness (%)	97.1 (68.8)
Redundancy	6.6 (5.8)
Refinement	
Resolution (Å)	48.7 (3.3–3.2)
No. reflections	36410 (2930)
R_{work}	0.19 (0.42)
R_{free}	0.24 (0.42)
No. atoms	12390
Protein	12390
Ligand/ion	0
Water	0
B factors (Å^2)	
Protein	75
Ligand/ion	
Water	
R.M.S deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.463
Ramachandran plot	
Favored [%]	97
Allowed [%]	3
Outliers [%]	0

* Values in parentheses are for highest-resolution shell.

DOI: https://doi.org/10.7554/eLife.37774.004

(***Figure 1—figure supplement 1***) (***Dürr et al., 2005***; ***Hauk et al., 2010***; ***Xia et al., 2016***; ***Yan et al., 2016***). For example, the DNA binding site of the *Saccharomyces cerevisiae* Chd1 Swi2/Snf2 domain is directly occluded by the chromodomain, providing a means of specific activation of the enzyme by interaction with a nucleosomal substrate (***Hauk et al., 2010***; ***Farnung et al., 2017***) (***Figure 1—figure supplement 1A***). Intriguingly, the 'resting' conformation of *Ct*Mot1^{CTD} is very similar to the crystallographic conformation of the *Sulpholobus solfataricus* Rad54 Swi2/Snf2 domain (***Dürr et al., 2005***) (***Figure 1D*** and ***Figure 1—figure supplement 1B***). The functional relevance of the *Sso*Rad54 Swi2/Snf2 domain conformation remained unclear because the crystallized and functionally analyzed fragment of *Sso*Rad54 comprised only the isolated Swi2/Snf2 domain.

Although the precise orientation of lobe 2 might be additionally determined by the bridge element that is missing in the structure, our structural data suggest that the auto-inhibited resting state of *Ct*Mot1 is stabilized by the interactions between subdomain 2A and HR1/2. To test this, we mutated ion pairs (R4-D1720, R45-D1716) and a hydrophobic loop (L1658A/Y1659A) to destabilize the resting state (***Figure 2A and B***). Basal ATPase activity of the point mutants was greatly increased compared to wild-type *Ct*Mot1 (WT) and was not further stimulated by DNA and TBP (***Figure 2C*** and ***Figure 2—figure supplement 1A***). *Ct*Mot1$^{\Delta C}$ did not show increased basal ATPase rates and was not activated by TBP alone. However, its ATPase activity in the presence of DNA-containing complexes exceeded that of the WT enzyme. To find out whether this elevated ATPase activity of the mutants translates into productive disruption of the substrate complexes, we performed remodeling assays. Notably, despite an increase in the ATP hydrolysis rate, the ability of *Ct*Mot1$^{\Delta C}$ to dissociate TBP:DNA complexes was impaired (***Figure 2D***). Assays performed under less efficient dissociation conditions that allowed the TBP:DNA complexes to persist confirmed that all other tested mutants (L1658A/Y1659A, R4D and D1720R) indeed behaved as the WT (***Figure 2—figure supplement 1B and C***). This shows that the bridge element acts in response to TBP binding and, similarly to SnAC in Snf2 (***Xia et al., 2016***), ensures productive coupling of ATP hydrolysis to the remodeling reaction.

Taken together, our data show that Mot1 adopts a resting state with low catalytic activity by stabilizing lobe 2 of the Swi2/Snf2 domain in an inactive conformation relative to lobe 1. Mobilization of lobe 2 from its auto-inhibited state explains the activation of Mot1's ATPase by TBP and TBP:DNA complexes.

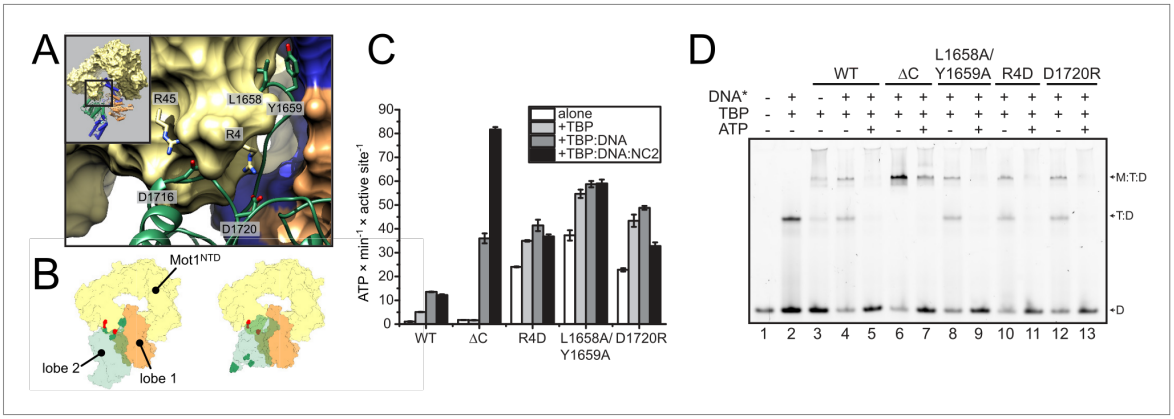


Figure 2. Analysis of *Ct*Mot1 mutants. (**A**) View at the interface between RecA2A (green cartoon), HR1/2 (yellow surface), and lobe 1 (orange/blue surface). Residues analyzed in this study are shown as sticks and labelled accordingly. (**B**) Cartoon model showing the positions of mutations in the Mot1^{NTD} (red spheres on yellow surface) and in lobe 2 (green spheres on green surface). Left: orientation as in the *Ct*Mot1 crystal structure. Right: *Ct*Mot1 with ATPase modeled as in the *S. cerevisiae* Chd1:nucleosome complex, that is the ATP hydrolysis-competent conformation (***Farnung et al., 2017***). (**C**) ATPase activity of the mutants. Error bars represent standard deviations from three technical replicates. *Ct*Mot1^{WT} is labelled as WT, *Ct*Mot1$^{\Delta C}$ as ΔC. (**D**) Electrophoretic mobility shift assay showing ATP-dependent dissociation of Mot1:TBP:DNA and TBP:DNA complexes. All *Ct*Mot1 constructs form ternary complexes with labelled DNA and TBP (M:T:D). In the presence of ATP and unlabeled competitor DNA (DNA*), wild-type *Ct*Mot1 (WT), L1658/Y1659, R4D, and D1720R mutants fully disrupt M:T:D and T:D complexes (lanes 5, 9, 11, and 13, respectively), whereas *Ct*Mot1$^{\Delta C}$ (ΔC) is less efficient (lane 7).

DOI: https://doi.org/10.7554/eLife.37774.005

The following source data and figure supplement are available for figure 2:

Source data 1. Raw data from the ATPase activity assay used for ***Figure 2C*** and ***Figure 2—figure supplement 1A***.

DOI: https://doi.org/10.7554/eLife.37774.007

Source data 2. Raw data from quantification of electrophoretic mobility shift assay used for ***Figure 2—figure supplement 1B***.

DOI: https://doi.org/10.7554/eLife.37774.008

Figure supplement 1. Analysis of *Ct*Mot1 mutants.

DOI: https://doi.org/10.7554/eLife.37774.006

Model of the Mot1:TBP:NC2:DNA complex

The new structure of the near full-length *Ct*Mot1 protein together with prior structures enables us to provide a model for the DNA path in the Mot1-bound protein:DNA complex (***Figure 3***). The *Ec*Mot1^{NTD}:TBP:NC2:DNA complex can be readily superimposed with *Ct*Mot1 through the conserved structure of the HEAT repeats. Likewise, superimposing *Sso*Rad54:DNA with *Ct*Mot1$^{\Delta C}$ via lobe 1 visualizes how the *Ct*Mot1 ATPase could initially contact duplex DNA since in all Snf2/Swi2 protein:substrate structures contacts between nucleic acid and the RecA1 subdomain are preserved. In the resulting model, localization and orientation of the Swi2/Snf2 domain is in full agreement with prior EM and CX-MS analyses (***Butryn et al., 2015***) and with biochemical studies showing that Mot1 covers two helical turns upstream from the TATA box (***Darst et al., 2001***; ***Sprouse et al., 2006***; ***Moyle-Heyrman et al., 2012***). Notably, the superimposed DNA segment bound by the ATPase is an almost direct continuation of the promoter DNA fragment from the *Ec*Mot1^{NTD}:TBP:NC2:DNA crystal structure. Assuming the generally proposed directionality of ATP dependent translocation of Swi2/Snf2 motor domains on dsDNA (***Zofall et al., 2006***; ***Saha et al., 2002***; ***Whitehouse et al., 2003***), the structure of *Ct*Mot1 and the specific orientation of lobe 1 now suggests that the Swi2/Snf2 motor translocates 'towards' the TATA box and TBP along the nucleic acid scaffold.

Our model of the Mot1:TBP:NC2:DNA complex suggests where the Swi2/Snf2 domain of Mot1 might engage with upstream DNA and provides new insight into how ATP hydrolysis-associated events are coupled to dissociation of protein:DNA substrates. Since processive ATP-dependent translocase activity has not been observed in biochemical studies, Mot1 could exploit short-range tracking toward TBP. Given the immediate vicinity of the Swi2/Snf2 domain to TBP, very few

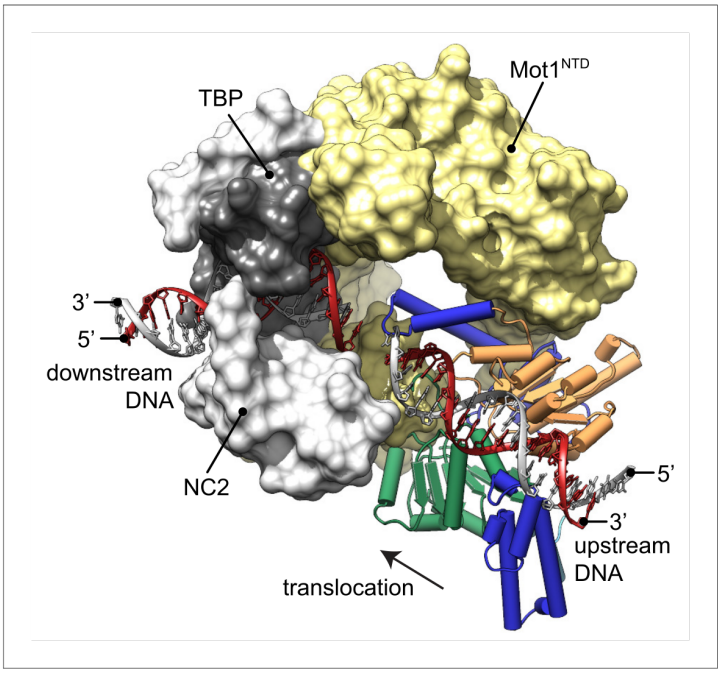


Figure 3. Model of the Mot1:TBP:NC2:DNA complex. *Ct*Mot1$^{\Delta C}$ was superimposed onto *Ec*Mot1^{NTD}:TBP:NC2:DNA via the HEAT repeats (yellow surface). The path of the upstream DNA was determined by superimposing *Sso*Rad54:DNA onto *Ct*Mot1$^{\Delta C}$ via lobe 1. The TATA box strand from *Ec*Mot1^{NTD}:TBP:NC2:DNA as well as the corresponding strand from *Sso*Rad54:DNA are marked in gray. The non-TATA box strands are in red. Substrate proteins TBP and NC2 are represented as dark and light gray surfaces, respectively. The black arrow represents the direction in which Swi2/Snf2 domain is proposed to translocate along the DNA scaffold. *Ct*Mot1^{CTD} is color-coded as in ***Figure 1***.

DOI: https://doi.org/10.7554/eLife.37774.009

translocation steps could lead to the displacement of TBP by steric collision (**Darst et al., 2001**; **Auble and Steggerda, 1999**; **Butryn et al., 2015**). In addition, Mot1 could simply displace TBP from DNA by overwinding or introducing other small distortions into upstream DNA (**Moyle-Heyrman et al., 2012**; **Butryn et al., 2015**). Similar effects have been observed for other transcription factors, for which not only binding but also dissociation rates depend on the structure of their recognition sites affected by the presence of other factors bound nearby (**Luo et al., 2014**; **Kim et al., 2013**). This allosteric effect can be accounted for by local changes to the major and minor groove width (**Kim et al., 2013**). Such a scenario is plausible since changes two helical turns upstream from the TBP binding site could have an immediate allosteric effect on severely bent and widened TATA box (**Tora and Timmers, 2010**).

Interestingly, while Mot1's ATPase orientation suggests that it 'pulls' DNA from TBP and overwinds DNA at the substrate, the reverse architecture is seen for the multisubunit INO80 remodeler: here the motor appears to pump DNA into the nucleosome and to underwind DNA at the substrate (**Eustermann et al., 2018**). Thus, our results suggest that Swi2/Snf2 proteins can use DNA translocation in different ways to disrupts protein:DNA interfaces.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Gene (*Chaetomium thermophilum*)	CtMot1	this paper	UniProtKB: G0S6C0 _CHATD	gene cloned from a cDNA library
Cell line (*Escherichia coli*)	Rosetta(DE3)	Novagen	Merck: 70954	
Cell line (*Escherichia coli*)	B843(DE3)	Novagen	Merck: 69041	
Recombinant DNA reagent	pETDuet-1	Novagen	Merck: 71146	used to express full-length CtMot1 (1–1886) and its point mutants
Recombinant DNA reagent	pET21b	Novagen	Merck: 69741	used to express *Ct*Mot1 (1–1836) and *Ct*Mot1 (1–1836, E1434Q)
Chemical compound, drug	L(+)-Selenomethionine	Acros Organics	Acros Organics: 259960010	42 mg/mL final concentration
Chemical compound, drug	SelenoMethionine Medium Base plus Nutrient Mix	Molecular Dimensions	Molecular Dimensions: MD12-501	
Chemical compound, drug	Adenosine 5′-triphosphate disodium salt hydrate (ATP)	Sigma-Aldrich	Sigma: A2383-10G	
Chemical compound, drug	β-Nicotinamide adenine dinucleotide reduced disodium salt hydrate (NADH)	Sigma-Aldrich	Sigma: 10107735001	
Chemical compound, drug	Phospho(enol)pyruvic acid monopotassium salt (PEP)	PanReac AppliChem	AppliChem: A2271	
Chemical compound, drug	Pyruvate kinase/lactic dehydrogenase enzymes from rabbit muscle	Sigma-Aldrich	Sigma: P0294	
Software, algorithm	XDS	**Kabsch, 2010**, doi: 10.1107/S0907444909047374		
Software, algorithm	*PHENIX*	**Adams et al., 2010**, doi:10.1107/S0907444909052925		

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Software, algorithm	*Coot*	***Emsley et al., 2010***, doi: 10.1107/S0907444910007493		
Software, algorithm	UCSF Chimera	***Pettersen et al., 2004*** , doi: 10.1002/jcc.20084		http://www.rbvi.ucsf.edu/chimera/
Software, algorithm	ImageJ 1.51 k	***Schneider et al., 2012***, doi: 10.1038/nmeth.2089		quantification of electrophoretic shift assay
Software, algorithm	OriginPro 2015	OriginLab, Northampton, MA		
Sequence-based reagent	48 bp dsDNA	Biomers		5′–CAGTACGGCCG GGCGCCCGGCA TGGCGGCCTATAAAA GGGGGTGGAAT–3′
Sequence-based reagent	48 bp 6-FAM labelled dsDNA	Biomers		5′–CAGTACGGCCGGGCGCCCG GCATGGCGGCCTATAAAA GGGGGTGGAAT–3′
Sequence-based reagent	36 bp dsDNA	Biomers		5′–CGGCCGGGCGCCCGG CATGGCGGCCTAT AAAAGGGC–3′

Protein purification

The sequence of the full-length Mot1 (1 – 1886) was isolated from the *Chaetomium thermophilum* cDNA library and cloned into pETDuet-1 vector (Novagen, Germany) harboring N-terminal His_6 tag followed by TEV cleavage site. *Ct*Mot1$^{\Delta C}$(1 – 1836) was cloned into pET21 vector containing PreScission protease cleavage site and C-terminal His_6 tag. Both constructs were expressed in *Escherichia coli* Rosetta(DE3) cells (Novagen) and purified using Ni^{2+}-NTA agarose (QIAGEN, Germany). After proteolytic cleavage of the expression tags, the proteins were further purified using ion-exchange chromatography (HiTrap Q HP, GE Healthcare, Germany) and size exclusion chromatography (HiLoad 16/60 200 pg, GE Healthcare). Proteins were concentrated to ~15 mg/ml in 20 mM Tris pH 7.5, 50 mM NaCl and 15% glycerol and stored at −80°C. Selenomethionine labelling of *Ct*Mot1$^{\Delta C}$ was performed in *E. coli* B843 (Novagen) using SelenoMethionine Medium Base and Nutrient Mix (Molecular Dimensions, UK) supplemented with L(+)-Selenomethionine (Acros Organics, Germany) at 42 mg/L. Purification of selenium-derivatized protein was performed according to the same protocol as for the native protein.

Crystallization and structure determination

Crystals of selenomethionine-derivatized *Ct*Mot1$^{\Delta C}$ were grown at 20°C by streak seeding in 0.1 M Tris pH 8.9, 0.2 M ammonium acetate and 13% (w/v) PEG 3350. Plate-like crystals with average dimensions of 700 × 150 × 30 µm appeared after three days and were cryocooled in liquid nitrogen using mother liquor supplemented with butanediol at 25% final concentration.

The data were collected at the European Synchrotron Radiation Facility in Grenoble, France at the peak of Se K-edge at 100K. Images were indexed, integrated, scaled, and merged in space group $P2_1$ to 3.2 Å using XDS package (***Kabsch, 2010***). The initial model was built manually to the experimental electron-density derived from SAD phasing using *PHENIX* AutoSol wizard (***Adams et al., 2010***). Alternating cycles of manual building using *Coot* (***Emsley et al., 2010***) and refinement with *PHENIX* yielded the final model (R_{work}/R_{free} of 19.0/23.8%) covering 87% of all residues.

ATPase assay

The assays were performed using an NADH-coupled assay as described (***Kiianitsa et al., 2003***). Reactions were performed using 2 mM phosphoenolpuryvate, 25 U/mL of pyruvate kinase/lactic dehydrogenase mix, 1 mM ATP, and 1 mM NADH at final concentrations. Test samples contained 100 nM dsDNA (5′–CAGTACGGCCGGGCGCCCGGCATGGCGGCCTATAAAAGGGGGTGGAAT–3′ top strand), 100 nM TBP, 100 nM NC2 and 250 nM *Ct*Mot1.

Electrophoretic mobility shift assays

Electrophoretic mobility shifts were essentially performed as described (***Darst et al., 2001***) with some modifications. In the assay shown in ***Figure 2D***, fluorescently labelled dsDNA (40 nM, 5′–CAG TACGGCCGGGCGCCCGGCATGGCGGCCTATAAAAGGGGGTGGAAT–3′ top strand with 6-FAM label on the 5′ end of the reverse strand) was incubated with TBP (10 nM) and Mot1 (25 nM) for 10 min. Unlabeled dsDNA competitor (800 nM, 5′–CGGCCGGGCGCCCGGCATGGCGGCCTA TAAAAGGGC–3′ top strand) was then added directly followed by ATP addition (50 µM) for 10 min. Samples were loaded onto 6% polyacrylamide gels and run at 160 V and 4° for 60 min and imaged using Typhoon FLA 9000 imager. The assays shown in ***Figure 2—figure supplement 1B*** and quantified in ***Figure 2—figure supplement 1C*** were prepared analogously, but TBP was added at a concentration of 15 nM and ATP was added for 6 min before loading the reactions on the gel.

Accession numbers

The coordinates and structure factors were deposited in the Protein Data Bank under accession code 6G7E.

Acknowledgements

We thank the Max-Planck Crystallization Facility (Martinsried) for crystallization trials and the European Synchrotron Radiation Facility (Grenoble) and the Deutsches Elektronen-Synchrotron (PETRA III, Hamburg) for beamtime and excellent support. AB acknowledges support from the Integrated Analysis of Macromolecular Complexes and Hybrid Methods in Genome Biology (DFG GRK1721) training program.

Additional information

Funding

Funder	Grant reference number	Author
National Institutes of Health	GM055763	David T. Auble
European Commission	ERC Advanced Grant ATMMACHINE	Karl-Peter Hopfner
Deutsche Forschungsgemeinschaft	Gottfried Wilhelm Leibniz-Prize	Karl-Peter Hopfner

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Agata Butryn, Formal analysis, Investigation, Writing—original draft; Stephan Woike, Savera J Shetty, Investigation; David T Auble, Formal analysis, Writing—review and editing; Karl-Peter Hopfner, Conceptualization, Formal analysis, Supervision, Writing—original draft

Author ORCIDs

Agata Butryn http://orcid.org/0000-0002-5227-4770
Karl-Peter Hopfner https://orcid.org/0000-0002-4528-8357

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.37774.015
Author response https://doi.org/10.7554/eLife.37774.016

Additional files

Supplementary files

- Transparent reporting form

DOI: https://doi.org/10.7554/eLife.37774.010

Data availability

The coordinates and structure factors are deposited in the Protein Data Bank under accession code 6G7E. All data generated or analysed during this study are included in the manuscript and supporting files. Source data files have been provided for Figures 2 and Figure 2-figure supplement 1.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Butryn A, Woike S, Shetty SJ, Auble DT, Hopfner K	2018	Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state	https://www.rcsb.org/structure/6G7E	Publicly available at the RCSB Protein Data Bank (accession no. 6G7E)

References

Adamkewicz JI, Mueller CG, Hansen KE, Prud'homme WA, Thorner J. 2000. Purification and enzymic properties of Mot1 ATPase, a regulator of basal transcription in the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry* **275**:21158–21168 . DOI: https://doi.org/10.1074/jbc.M002639200

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. *PHENIX* : a comprehensive Python-based system for macromolecular structure solution . *Acta Crystallographica Section D Biological Crystallography* **66**:213–221. DOI: https://doi.org/10.1107/S0907444909052925

Auble DT, Wang D, Post KW, Hahn S. 1997. Molecular analysis of the SNF2/SWI2 protein family member MOT1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. *Molecular and Cellular Biology* **17**:4842–4851. DOI: https://doi.org/10.1128/MCB.17.8.4842, PMID: 9234740

Auble DT, Steggerda SM. 1999. Testing for DNA tracking by MOT1, a SNF2/SWI2 protein family member. *Molecular and Cellular Biology* **19**:412–423. DOI: https://doi.org/10.1128/MCB.19.1.412, PMID: 9858565

Ayala R, Willhoft O, Aramayo RJ, Wilkinson M, McCormack EA, Ocloo L, Wigley DB, Zhang X. 2018. Structure and regulation of the human INO80-nucleosome complex. *Nature* **556**:391–395. DOI: https://doi.org/10.1038/s41586-018-0021-6, PMID: 29643506

Becker PB, Workman JL. 2013. Nucleosome remodeling and epigenetics. *Cold Spring Harbor Perspectives in Biology* **9**:a017905 . DOI: https://doi.org/10.1101/cshperspect.a017905

Blossey R, Schiessel H. 2018. The latest twists in chromatin remodeling. *Biophysical Journal* **114**:2255–2261. DOI: https://doi.org/10.1016/j.bpj.2017.12.008, PMID: 29310890

Butryn A, Schuller JM, Stoehr G, Runge-Wollmann P, Förster F, Auble DT, Hopfner K-P. 2015. Structural basis for recognition and remodeling of the TBP:dna:nc2 complex by Mot1. *eLife* **4**:e07432. DOI: https://doi.org/10.7554/eLife.07432

Carroll C, Bansbach CE, Zhao R, Jung SY, Qin J, Cortez D. 2014. Phosphorylation of a C-terminal auto-inhibitory domain increases SMARCAL1 activity. *Nucleic Acids Research* **42**:918–925. DOI: https://doi.org/10.1093/nar/gkt929, PMID: 24150942

Chicca JJ, Auble DT, Pugh BF. 1998. Cloning and biochemical characterization of TAF-172, a human homolog of yeast Mot1. *Molecular and Cellular Biology* **18**:1701–1710. DOI: https://doi.org/10.1128/MCB.18.3.1701, PMID: 9488487

Clapier CR, Cairns BR. 2012. Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature* **492**:280. DOI: https://doi.org/10.1038/nature11625, PMID: 23143334

Darst RP, Wang D, Auble DT. 2001. MOT1-catalyzed TBP-DNA disruption: uncoupling DNA conformational change and role of upstream DNA. *The EMBO Journal* **20**:2028–2040. DOI: https://doi.org/10.1093/emboj/20.8.2028, PMID: 11296235

Darst RP, Dasgupta A, Zhu C, Hsu JY, Vroom A, Muldrow T, Auble DT. 2003. Mot1 regulates the DNA binding activity of free TATA-binding protein in an ATP-dependent manner. *Journal of Biological Chemistry* **278**:13216–13226. DOI: https://doi.org/10.1074/jbc.M211445200, PMID: 12571241

Dasgupta A, Darst RP, Martin KJ, Afshari CA, Auble DT. 2002. Mot1 activates and represses transcription by direct, ATPase-dependent mechanisms. *PNAS* **99**:2666–2671. DOI: https://doi.org/10.1073/pnas.052397899, PMID: 11880621

Dürr H, Körner C, Müller M, Hickmann V, Hopfner KP. 2005. X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell* **121**:363–373. DOI: https://doi.org/10.1016/j.cell.2005.03.026, PMID: 15882619

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. "Features and development of Coot.". *Acta Crystallographica Section D* **66**:486–501 . DOI: https://doi.org/10.1107/S0907444910007493

Eustermann S, Schall K, Kostrewa D, Lakomek K, Strauss M, Moldt M, Hopfner KP. 2018. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**:386–390. DOI: https://doi.org/10.1038/s41586-018-0029-y, PMID: 29643509

Farnung L, Vos SM, Wigge C, Cramer P. 2017. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* **550**:539–542. DOI: https://doi.org/10.1038/nature24046, PMID: 29019976

Flaus A, Martin DM, Barton GJ, Owen-Hughes T. 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Research* **34**:2887–2905. DOI: https://doi.org/10.1093/nar/gkl295, PMID: 16738128

Hauk G, McKnight JN, Nodelman IM, Bowman GD. 2010. The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Molecular Cell* **39**:711–723. DOI: https://doi.org/10.1016/j.molcel.2010.08.012, PMID: 20832723

Hauk G, Bowman GD. 2011. Structural insights into regulation and action of SWI2/SNF2 ATPases. *Current Opinion in Structural Biology* **21**:719–727. DOI: https://doi.org/10.1016/j.sbi.2011.09.003, PMID: 21996440

Hopfner KP, Gerhold CB, Lakomek K, Wollmann P. 2012. Swi2/Snf2 remodelers: hybrid views on hybrid molecular machines. *Current Opinion in Structural Biology* **22**:225–233. DOI: https://doi.org/10.1016/j.sbi.2012.02.007, PMID: 22445226

Kabsch W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D Biological Crystallography* **66**:133–144. DOI: https://doi.org/10.1107/S0907444909047374, PMID: 20124693

Kiianitsa K, Solinger JA, Heyer WD. 2003. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Analytical Biochemistry* **321**:266–271. DOI: https://doi.org/10.1016/S0003-2697(03)00461-5, PMID: 14511695

Kim S, Broströmer E, Xing D, Jin J, Chong S, Ge H, Wang S, Gu C, Yang L, Gao YQ, Su XD, Sun Y, Xie XS. 2013. Probing allostery through DNA. *Science* **339**:816–819. DOI: https://doi.org/10.1126/science.1229223, PMID: 23413354

Liu X, Li M, Xia X, Li X, Chen Z. 2017. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**:440–445. DOI: https://doi.org/10.1038/nature22036, PMID: 28424519

Luo Y, North JA, Rose SD, Poirier MG. 2014. Nucleosomes accelerate transcription factor dissociation. *Nucleic Acids Research* **42**:3017–3027. DOI: https://doi.org/10.1093/nar/gkt1319, PMID: 24353316

Moyle-Heyrman G, Viswanathan R, Widom J, Auble DT. 2012. Two-step mechanism for modifier of transcription 1 (Mot1) enzyme-catalyzed displacement of TATA-binding protein (TBP) from DNA. *Journal of Biological Chemistry* **287**:9002–9012. DOI: https://doi.org/10.1074/jbc.M111.333484, PMID: 22298788

Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. 2013. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154**:490–503. DOI: https://doi.org/10.1016/j.cell.2013.07.011, PMID: 23911317

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**:1605–1612. DOI: https://doi.org/10.1002/jcc.20084, PMID: 15264254

Saha A, Wittmeyer J, Cairns BR. 2002. Chromatin remodeling by RSC involves ATP-dependent DNA translocation. *Genes & Development* **16**:2120–2134. DOI: https://doi.org/10.1101/gad.995002, PMID: 12183366

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to ImageJ: 25 years of image analysis. *Nature Methods* **9**:671–675. DOI: https://doi.org/10.1038/nmeth.2089, PMID: 22930834

Sen P, Ghosh S, Pugh BF, Bartholomew B. 2011. A new, highly conserved domain in Swi2/Snf2 is required for SWI/SNF remodeling. *Nucleic Acids Research* **39**:9155–9166. DOI: https://doi.org/10.1093/nar/gkr622, PMID: 21835776

Sprouse RO, Brenowitz M, Auble DT. 2006. Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. *The EMBO Journal* **25**:1492–1504. DOI: https://doi.org/10.1038/sj.emboj.7601050, PMID: 16541100

Sundaramoorthy R, Hughes AL, El-Mkami H, Norman DG, Ferreira H, Owen-Hughes T. 2018. Structure of the chromatin remodelling enzyme Chd1 bound to a ubiquitinylated nucleosome. *eLife* **7**:e35720. DOI: https://doi.org/10.7554/eLife.35720, PMID: 30079888

Tora L, Timmers HT. 2010. The TATA box regulates TATA-binding protein (TBP) dynamics in vivo. *Trends in Biochemical Sciences* **35**:309–314. DOI: https://doi.org/10.1016/j.tibs.2010.01.007, PMID: 20176488

Wang L, Limbo O, Fei J, Chen L, Kim B, Luo J, Chong J, Conaway RC, Conaway JW, Ranish JA, Kadonaga JT, Russell P, Wang D. 2014. Regulation of the Rhp26ERCC6/CSB chromatin remodeler by a novel conserved leucine latch motif. *PNAS* **111**:18566–18571. DOI: https://doi.org/10.1073/pnas.1420227112, PMID: 25512493

Whitehouse I, Stockdale C, Flaus A, Szczelkun MD, Owen-Hughes T. 2003. Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. *Molecular and Cellular Biology* **23**:1935–1945. DOI: https://doi.org/10.1128/MCB.23.6.1935-1945.2003, PMID: 12612068

Wollmann P, Cui S, Viswanathan R, Berninghausen O, Wells MN, Moldt M, Witte G, Butryn A, Wendler P, Beckmann R, Auble DT, Hopfner KP. 2011. Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. *Nature* **475**:403–407. DOI: https://doi.org/10.1038/nature10215, PMID: 21734658

Xia X, Liu X, Li T, Fang X, Chen Z. 2016. Structure of chromatin remodeler Swi2/Snf2 in the resting state. *Nature Structural & Molecular Biology* **23**:722–729. DOI: https://doi.org/10.1038/nsmb.3259, PMID: 27399259

Yan L, Wang L, Tian Y, Xia X, Chen Z. 2016. Structure and regulation of the chromatin remodeller ISWI. *Nature* **540**:466–469. DOI: https://doi.org/10.1038/nature20590, PMID: 27919072

Zentner GE, Henikoff S. 2013. Mot1 redistributes TBP from TATA-containing to TATA-less promoters. *Molecular and Cellular Biology* **33**:4996–5004. DOI: https://doi.org/10.1128/MCB.01218-13, PMID: 24144978

Zofall M, Persinger J, Kassabov SR, Bartholomew B. 2006. "Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome." *Nature Structural &Amp. Molecular Biology* **13**:339 . DOI: https://doi.org/10.1038/nsmb1071

5.2 Structural basis for TBP displacement from the TATA box by the Swi2/Snf2 ATPase Mot1

Stephan Woike, Sebastian Eustermann, James Jung, Simon Josef Wenzl, Götz Hagemann, Joseph Bartho, Katja Lammens, Agata Butryn, Franz Herzog, Karl-Peter Hopfner. Structural basis for TBP displacement from the TATA box by the Swi2/Snf2 ATPase Mot1, *Nature Structural and Molecular Biology*, accepted in principle, Sep 2022.

Summary

In this publication, we present five cryo-EM structures of *Chaetomium thermophilum* Swi2/Snf2 ATPase Mot1, mostly in complex with TBP bound to DNA bearing a TATA motif. The structural work was accompanied by NADH-coupled ATPase assays and EMSAs to analyze the coupling between ATPase activity and TBP:DNA dissociation by Mot1. By using different ATP analogues, we could reconstruct a complete Mot1 ATP-binding and -hydrolysis cycle and visualize the stepwise removal of TBP from DNA. Without bound nucleotide, the Mot1 N-terminal HEAT-repeat ring partly interacts with TBP, whereas only the RecA1 lobe of the C-terminal ATPase stably contacts upstream DNA. In this so-called "substrate recognition state", the RecA2 lobe remains flexible, but closes upon binding of the ATP-ground state analogue ADP-BeF_3^-. Mimicking a "prehydrolysis state", in this conformation the upstream DNA is engaged by the both ATPase lobes and TBP is fully engulfed in an induced fit-like manner by the Mot1 N-terminal domain. Also, TBP is already lifted off the DNA with its DNA binding site occupied by the so-called hook element – a region that we identified as part of the C-terminal bridge via crosslink-mass spectrometry. Further, we could identify the outermost part of the bridge as a partly helical density tracing the inner part of the HEAT-repeat ring. This structure confirms the observation from our earlier publication that the bridge functions as an allosteric regulator of Mot1 remodeling activity. It does so by contacting TBP and simultaneously occluding the concave site of TBP preventing rebinding to DNA. We used the slowly hydrolysable ATP analogue ATPγS, achieving a kinetically retarded ATP-hydrolysis in combination with a dilution of the nucleotide out of the binding pocket by SEC, which yielded a "posthydrolysis state". In this conformation, the Mot1 N-terminal domain has undergone an extensive pivot movement that correlates with the manifestation of a "wedge" element following the brace. The wedge protrudes between Mot1 N- and C-terminus acting like a lever that supports the hinge-like motion. The DNA was translocated by 1-2 base pairs and by bending the DNA through the pivot movement, the TATA motif is rotated away from TBP. Additional binding of ATPγS resulted in the "product state" with complete disruption of Mot1:TBP from DNA. In sum, in this publication

we present structural data that reveal a non-processive, short-range Swi2/Snf2 ATPase activity, where the substrate TBP is remodeled by a concerted activity of gripping, bending and rotating the DNA.

Author contribution

I planned and purified the Mot1 constructs, performed EMSAs and ATPase assays and collected and processed most of the cryo-EM data. I planned the experiments together with Dr. Sebastian Eustermann und Prof. Dr. Karl-Peter Hopfner. Further, I prepared the figures and wrote the manuscript together with Prof. Dr. Karl-Peter Hopfner.

Structural basis for TBP displacement from TATA box DNA by the Swi2/Snf2 ATPase Mot1

Stephan Woike[1,2], Sebastian Eustermann[1,2,3+], James Jung[1,2+], Simon Josef Wenzl[1,2], Götz Hagemann[1,2], Joseph Bartho[1,2], Katja Lammens[1,2], Agata Butryn[1,2,4], Franz Herzog[2,5] & Karl-Peter Hopfner[1,2,*]

[1]Department of Biochemistry, Ludwig-Maximilians-Universität, Munich, Germany

[2]Gene Center, Ludwig-Maximilians-Universität, Munich, Germany

[3]Present address: European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Heidelberg, Germany

[4]Present address: Macromolecular Machines Laboratory, Francis Crick Institute, London, UK.

[5]Present address: Institute Krems Bioanalytics, IMC University of Applied Sciences, 3500 Krems, Austria

[+]These authors contributed equally

*Correspondence:

Contact: Prof. Dr. Karl-Peter Hopfner

Gene Center

Feodor-Lynen-Str. 25

81377 Munich, Germany

Tel.: +49 (0) 89 2180 76953

Fax: +49 (0) 89 2180 76999

Email: hopfner@genzentrum.lmu.de

Keywords: Swi2/Snf2 protein, remodeler, transcription initiation, cryo-electron microscopy

The Swi2/Snf2 family transcription regulator Mot1 uses ATP to dissociate and reallocate the TATA box binding protein TBP from and between promoters. To reveal how Mot1 removes TBP from TATA box DNA we determined cryo-electron microscopy structures that capture different states of the remodeling reaction. The resulting molecular movie reveals how Mot1 dissociates TBP in a process that, intriguingly, does not require DNA groove tracking. Instead, the motor grips DNA in the presence of ATP and swings back after ATP hydrolysis, moving TBP to a thermodynamically less stable position on DNA. Dislodged TBP is trapped by a chaperone element that blocks TBP's DNA binding site. Our results show how Swi2/Snf2 proteins can remodel protein:DNA complexes through DNA bending without processive DNA tracking and reveal mechanistic similarities to RNA gripping DEAD box helicases and RIG-I like immune sensors.

Swi2/Snf2 ATPases remodel protein:DNA complexes in a wide range of genome associated processes such as nucleosome remodeling, DNA repair and transcription regulation and feature chromatin remodelers, DNA repair proteins and several other proteins[1,2]. Swi2/Snf2 proteins are generally suggested to remodel protein:DNA complexes by DNA groove tracking and applying translational and rotational forces on DNA[3–5]. Despite many landmark structures in recent years[6–11] elucidating the mechanisms of remodeling based on DNA translocation, detailed insights into the remodeling mechanisms are still scarce and we also do not mechanistically understand how some family members remodel their targets apparently without processive DNA translocation.

The Swi2/Snf2 protein Modifier of Transcription 1 (Mot1) regulates, represses, and activates transcription in a context dependent manner through reallocating TBP between promoters to assemble RNA polymerase II initiation complexes[12–15]. Mot1 has served as a pioneering model for Swi2/Snf2 proteins and their remodeling and chaperone functions[16,17]. Mot1 contains an N-terminal domain ($Mot1^{NTD}$), comprising an array of 16 HEAT repeats, that serves as a flexible scaffold for TBP via several loop insertions on its convex side[16,18] (Fig. 1 a). The C-terminal domain ($Mot1^{CTD}$) is a Swi2/Snf2-type ATPase and possesses the two Swi2/Snf2 characteristic RecA-like like domains, along with a Mot1-specific, hitherto structurally unresolved 50 amino acid long C-terminal extension. This extension, denoted "bridge", serves as an allosteric regulator of TBP remodeling[19].

The current models of TBP dissociation of Mot1 include both a remodeling as well as a chaperoning function in the sense that it shields interaction sites and thereby regulates complex assembly. Upon remodeling of the TBP:DNA complex through action of $Mot1^{CTD}$, the DNA binding site of TBP is blocked through a "latch" from reassociation with DNA[16]. Like with more

complex remodelers, the structural mechanism of how Mot1 specifically displaces TBP is still unclear as neither the full protein in complex with TBP and DNA, nor remodeling intermediates have been visualized yet. For instance, it is puzzling that measurable processive DNA translocation activity in biochemical assays was not observed[20] and TBP dissociation by Mot1 is enhanced by the cooperation of two Mot1 molecules[21,22]. Remodeling without measurable translocation and transient dimers are also reported for some complex nucleosome remodelers[9,23], begging for mechanistic explanations.

Mot1^{NTD} initially recognizes TBP:DNA by an array of HEAT repeats.

We used cryogenic electron microscopy (cryo-EM) to interrogate the remodeling of DNA bound TBP by Mot1. *Chaetomium thermophilum* (*Ct*)Mot1, *Ct*TBP and TATA box containing DNA were mixed in the absence and presence of various nucleotides (ATP, ADP-BeF_3^-, ATPγS) before and after size exclusion chromatography (SEC) in order to capture different functional stages.

First, we set out to capture the ATP-bound Mot1:TBP:DNA complex by probing an ATP hydrolysis deficient Walker B mutant (Mot1^{E1434Q}) along with ATP. In a native electrophoretic mobility shift assay (EMSA), Mot1^{E1434Q} complex is unable to disrupt TBP:DNA (Fig. 1 b, Extended Data Fig. 1 a) and lacks ATP-hydrolysis activity (data not shown). In the resulting 5.1 Å reconstruction, we for the first time visualize a complete Mot1 protein bound to both DNA and TBP. As in our earlier structures[16,18], TBP is bound to TATA DNA and partially engaged by the horseshoe shaped HEAT repeats of Mot1^{NTD}. Interestingly, the DNA angle at TBP is with ~45° less than the ~90° angle observed in the free TBP:DNA complex (Extended Data Fig. 1 b), suggesting that binding of Mot1 to TBP and upstream DNA already straightens DNA and possibly weakens the TBP:DNA interaction to some extent, consistent with biophysical studies[23]. The C-terminal Swi2/Snf2 domain (Mot1^{CTD}) engages with upstream DNA (Fig. 1 d), with RecA1 binding both strands of the DNA minor groove approximately two helical turns away from the TATA box.

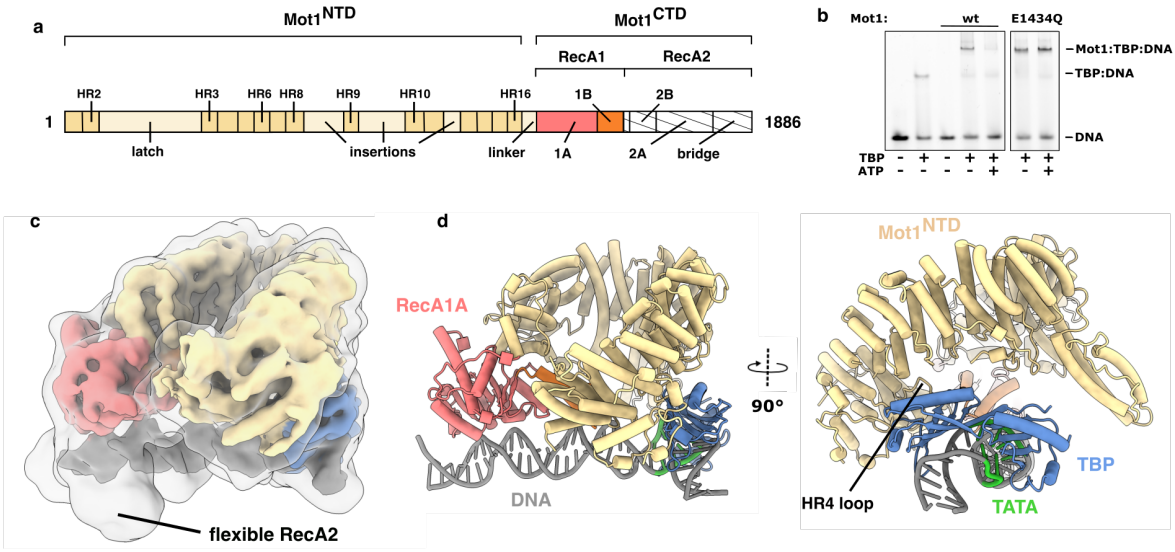


Figure 1: Cryo-EM structure of initial TBP recognition by Mot1 (substrate recognition state)

a) Schematic domain architecture of Mot1 with annotated elements. Not structurally resolved parts are hatched. (HR: HEAT-repeat, NTD: N-terminal domain, CTD: C-terminal domain)
b) Native electrophoretic mobility shift assays with 5'-6-FAM upstream labeled DNA showing TBP:DNA dissociation by Mot1.
c) Cryo-EM density of Mot1^{E1434Q}:TBP:DNA in the substrate recognition state, colored according to the structural model in d). A gaussian filtered density visualizes the flexible RecA2 lobe.
d) Two rotated views of the Mot1^{E1434Q}:TBP:DNA structure in the substrate recognition state.

However, the two RecA-like lobes of the C-terminal Swi2/Snf2 domain (Mot1^{CTD}) do not constitute an ATP-bound conformation. While RecA1 effectively binds upstream DNA and its density is reasonably well defined (Fig. 1 c), density for RecA2 is largely missing and only appears after Gaussian filtering of the map. Based on evidence from other Swi2/Snf2 family members it is known that in resting states of many Swi2/Snf2 proteins (absence of substrate binding) RecA2 exhibits either high flexibility or is locked in an inactive "open" conformation[11,24–26] and needs proper engagement with substrates to adopt a functional state. For instance, empty nucleotide-free Mot1 crystallized in the "open" ATPase conformation with RecA2 locked by engagement with Mot1^{NTD} [19]. In the cryo-EM structure, Mot1^{CTD} is indeed liberated from its auto-inhibited state, but it remains flexible and Mot1^{CTD} does not yet stably enclose DNA with both ATPase lobes. A plausible reason is that the Walker B mutant prevents formation of a proper ATP bound state. We interpret this state as initial recognition of the TBP:DNA complex prior to full DNA engagement of the Swi2/Snf2 domain and denote it "substrate recognition" state.

Nucleotide binding leads to dislodgement and chaperoning of TBP.

Since mutating the Walker B motif in conjunction with ATP did not result in a fully engaged Mot1^{CTD}:DNA complex, we alternatively incubated wild type Mot1:TBP:DNA with the ATP analog ADP-BeF_3^- prior to vitrification. This state could be reconstructed to 3.6 Å resolution and showed a conformation where DNA is now fully gripped by both lobes of Mot1^{CTD}. Although well-defined density for DNA is limited to the base pairs gripped by the RecA-like lobes, gaussian filtering revealed DNA density can be traced downstream towards TBP (Fig. 2 a). Interestingly, in this state we do not find density for DNA at TBP itself, suggesting that TBP has been displaced from DNA, yet DNA did not dissociate from Mot1 yet.

However, the DNA binding site of TBP is not empty, but we observe density for a part of the C-terminal “bridge” element of Mot1 at the concave DNA binding surface of TBP (Fig. 2 a). This element (denoted “hook”, aa1867-1886) acts as chaperone blocking TBP re-association with DNA, once DNA is displaced from TBP (Fig. 2 b, Extended Data Fig. 1 c, see below). In addition, as we explain later in more detail, an interrupted, seemingly flexible density proximal to RecA2A can be identified as the so called “brace” (Fig. 2 a, b). ADP-BeF_3^- readily fits into the nucleotide density in the interlobal cleft of the CTD (Fig 2 c). $BeF3^-$ is only visible at a lower map threshold – a typical phenomenon for anions intrinsic to electron microscopy derived densities and in line with other ADP-BeF_3^- containing Swi2/Snf2 cryo-EM maps[7,9,27] Still, the RecA1 and 2 align very well to other Swi2/Snf2 proteins in ADP-BeF_3^- states[7,27], allowing for an unambiguous assignment of a conformation that mimics ATP-binding (Fig. 2 d). In summary, ADP-BeF_3^- resulted in a full grip of the ATPase on upstream DNA, but already led to a partial dissociation of TBP from DNA. Hence the observed conformation identifies a remodeling intermediate of a Swi2/Snf2 complex where substrate interactions have been partially broken that we denote "prehydrolysis" state.

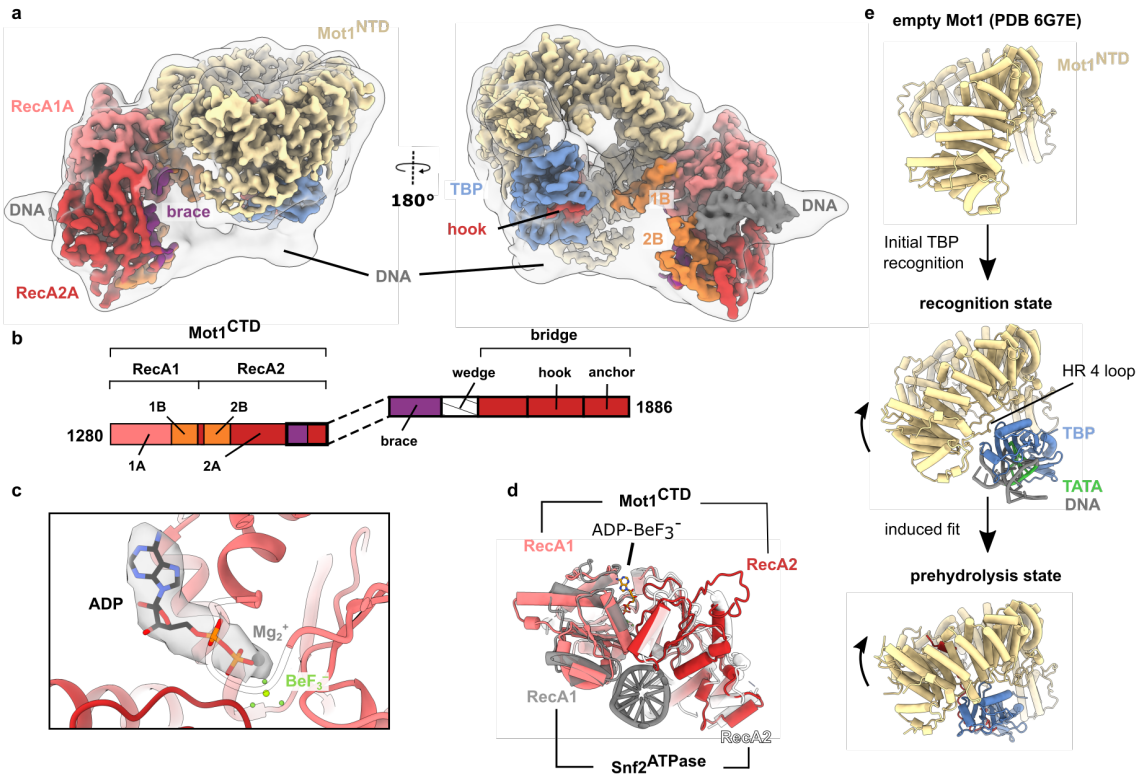


Figure 2: Cryo-EM structure of the ATP-BeF_3^- bound Mot1:TBP:DNA prehydrolysis state

a) Denoised cryo-EM map of Mot1:TBP:DNA prehydrolysis state from two rotated views. The surrounding gaussian filtered map delineates the low resolved DNA path.
b) Schematic domain architecture of Mot1 C-terminus with annotated elements and close up on outermost C-terminus. For the ADP-BeF_3^- containing conformation (prehydrolysis state), the wedge element (hatched) remains structurally unresolved.
c) Close up on the nucleotide binding pocket with ADP-BeF_3^- in its density.
d) Structural alignment of prehydrolysis $Mot1^{CTD}$ and Snf2 in the ADP-BeF_3^- state (PDB 5Z3V)
e) Conformational trajectory of TBP accommodation by the $Mot1^{NTD}$ HEAT repeat array based on three structures. The arrows indicate the lifting movement of the HEAT repeats.

Structural comparison of $Mot1^{NTD}$ in "substrate recognition" and "prehydrolysis" states shed light on allosteric activation of Mot1 by TBP binding. As discussed above, in the absence of ligands, Mot1 resides in an autoinhibited resting state with a flipped, inactive RecA2[19]. We find here that, upon TBP and DNA binding, a spiral-to-planar movement of $Mot1^{NTD}$ releases RecA2 from its autoinhibited conformation to adopt an active state with properly aligned ATPase motifs (Fig. 2 e). Initial TBP recognition by the HEAT repeat (HR) 4 loop already correlates with lifting of the HEAT repeat array compared to the empty Mot1 structure (Extended Data Fig. 1 d). In the nucleotide-bound prehydrolysis state, an induced-fit like conformational change results in a complete recognition of TBP by a now planar HEAT repeat array (Fig. 2 e, Suppl. Movie 1). This is facilitated by disruption of the TBP DNA interface, which allows a rotation of TBP relative to $Mot1^{CTD}$ in order to fully bind $Mot1^{NTD}$. In summary, the dislodgement of TBP from DNA emerges

as a multifaceted process induced by nucleotide binding: an apparent strain on upstream DNA by the closed ATPase accompanied by an induced fit of TBP into Mot1^{NTD} and competitive occupancy of the TBP concave site by the hook.

ATP hydrolysis leads to DNA bending and Mot1^{CTD} pivoting

Next, we aimed at resolving the role of ATP-hydrolysis, which was shown to be essential for complete removal of TBP from DNA[22,28]. Because Mot1 rapidly dissociates TBP:DNA *in vitro*, we set out to kinetically decelerate the reaction using ATPγS prior to SEC, which shows 2.5 fold reduced hydrolysis rate compared to ATP (Fig. 3 a) and also leads to a kinetically restrained TBP:DNA dissociation in gel shift assays (Fig. 3 b). This strategy yielded a 3.9 Å cryo-EM reconstruction of a "posthydrolysis" state (Fig. 3 c).

This state bears some resemblance with the “prehydrolysis” state in that DNA is bound to Mot1^{CTD} but lifted of TBP, which is tightly gripped by Mot1^{NTD}. However closer inspection reveals also remarkable differences to the prehydrolysis state: the entire Mot1^{CTD} is tilted backwards by ~30° and strongly bends DNA by a repositioning of RecA2 and RecA1 compared to the prehydrolysis state. A fully structured helical brace now tethers the two RecA-like lobes. Consistently, the conformation of Mot1^{CTD} RecA1 and A2 resemble Swi2/Snf2 domains in nucleotide-free or posthydrolysis states[8,29] and both RecA-like lobes align with the apo Snf2 ATPase[29] (Fig. 3 d). We do not observe extra density at the Walker A/B motifs (Extended Data Fig. 1 e), which indicates that during SEC remaining ADP (or non-hydrolyzed ATPγS) diluted out from its binding pocket. All these points let us conclude, that we obtained a nucleotide-free posthydrolysis state with the complex still partially bound to DNA. In this conformation, the DNA density between the RecA-like lobes reveals a minor groove widening by insertion of RecA2B (Fig. 3 e), which – as we explain later – emerges as a key element of the remodeling mechanism. Finally, to visualize a full "product state", we added additional ATPγS prior to vitrification. This procedure yielded a state that strongly resembles the prehydrolysis state, which could be rigid-body docked into the product state density, but with fully dissociated DNA (Fig. 3 f). Evidently, repetitive binding and/or hydrolysis cycles of ATPγS led to dissociation of DNA also from Mot1^{CTD} once it is displaced from TBP by action of the hook.

Taken together, these observations have profound consequences for the interpretation of the remodeling mechanism by Mot1, as it appears that instead of a groove-tracking mechanism, Mot1 remodels TBP:DNA by a mechanism that involves DNA gripping (prehydrolysis), bending and Swi2/Snf2 domain tilting (posthydrolysis) and release (additional cycle starting with ATP binding).

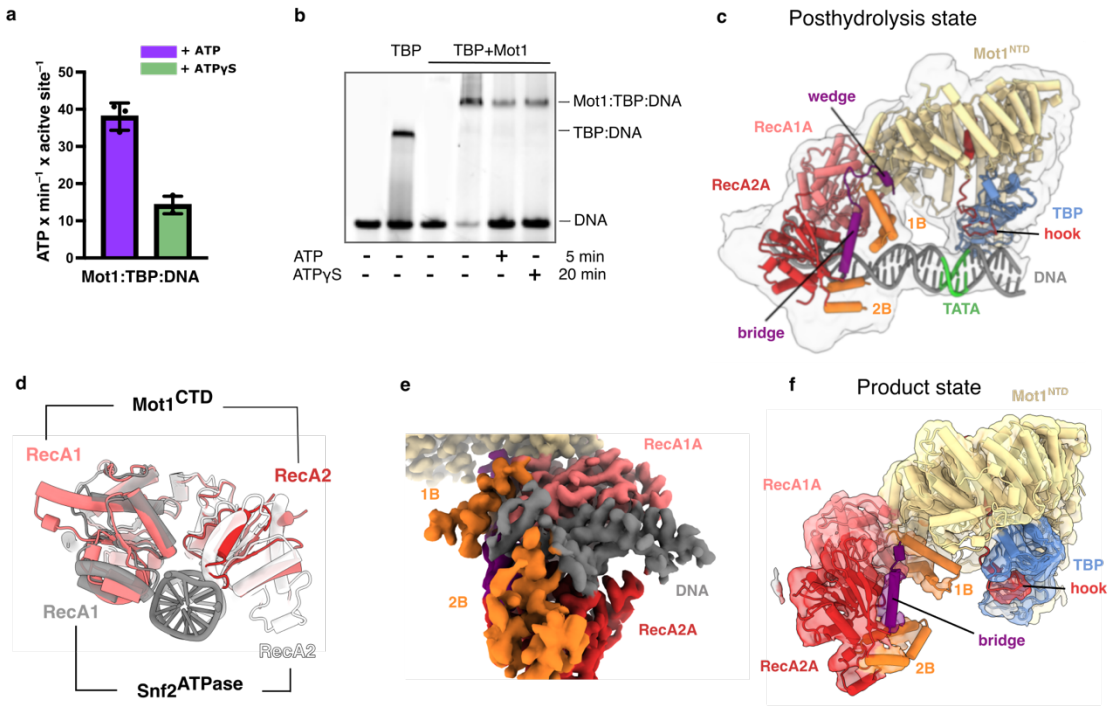


Figure 3: Posthydrolysis and product state of Mot1:TBP:DNA.

a) ATPase assay comparing the hydrolysis rate between ATP and ATPγS for Mot1:TBP:DNA. Horizontal bars represent means from three technical replicates shown as dots, error bars represent the standard deviation.
b) Native electrophoretic mobility shift assays with 5'-6-FAM upstream labeled DNA for showing relative complex stability of Mot1:TBP:DNA upon addition of ATP or ATPγS. Times of incubation with the respective nucleotide are indicated.
c) Structural model of the Mot1:TBP:DNA posthydrolysis state fitted in a gaussian filtered cryo-EM map.
d) Structural alignment of posthydrolysis Mot1^{CTD} and Snf2 in the apo state (PDB 5X0X).
e) Close up view on upstream DNA bound by Mot1^{CTD}, showing minor groove widening by RecA2B.
f) Cryo-EM density of the Mot1:TBP:DNA product state with a rigid-body fitted Mot1 prehydrolysis state.

The brace-wedge element controls remodeling and chaperoning functions.

The C-terminal brace-bridge element emerges as a multifunctional regulator of the conformational cycle and remodeling states of Mot1. The 'brace' is a conserved element in Swi2/Snf2 proteins and connects RecA1B and RecA2B, playing a key role in coupling ATP cycles to remodeling[29,30]. Mot1's brace is similar to that of other Swi2/Snf2 proteins, but continues as a ~70 amino acid long 'bridge' extension. The bridge is functionally critical in *in vivo*[31] and *in vitro*[19]. Deleting the terminal 50 amino acids of Mot1 robustly slowed down TBP displacement kinetics but without a major reduction in ATPase activity (see below), confirming

that the bridge is important to couple ATP dependent conformational changes to TBP dissociation.

The major conformational switch during the TBP displacement, the ~30° pivot movement between Mot1^{NTD} and Mot1^{CTD}, indeed correlates with rearrangements and ordering of the brace-bridge with its three terminal elements “wedge”, “hook” and “anchor” (Fig. 4). In the posthydrolysis state, not only does the brace fully assembly into a helix and stabilizes the conformational of both RecA-like lobes, the subsequent “wedge” binds into a cleft between Mot1^{CTD} and Mot1^{NTD}, locking their backward-tilted conformation in the posthydrolysis state (Fig. 4).

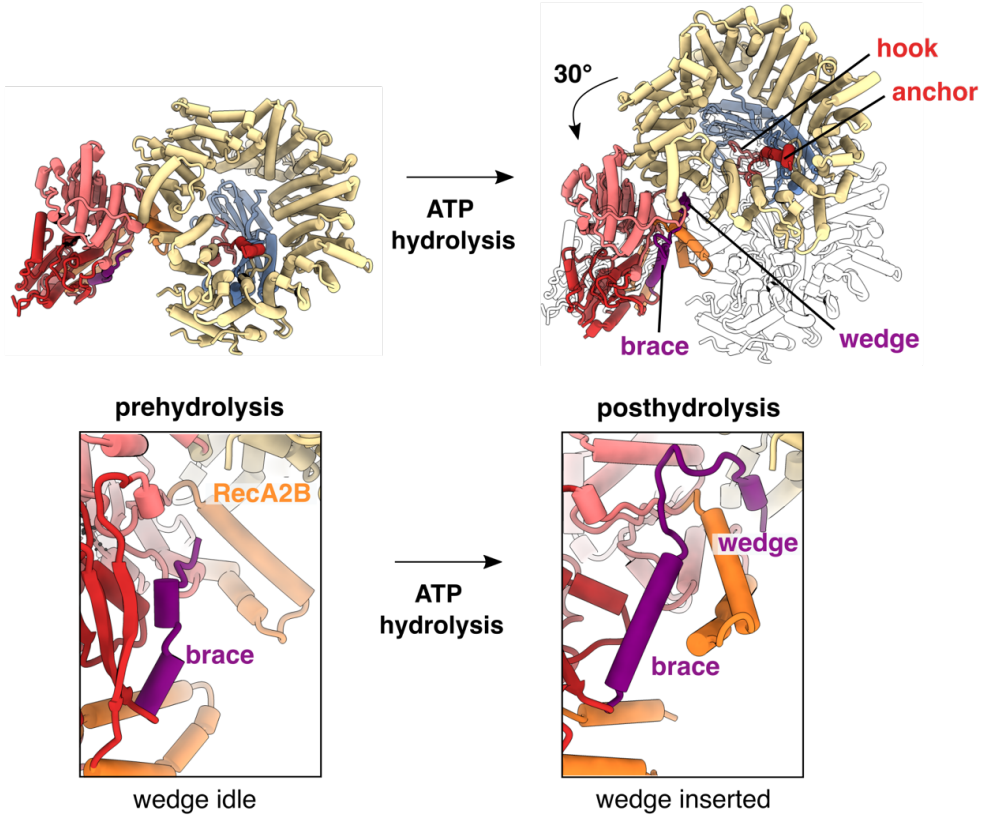


Figure 4: Mot1^{CTD} pivot movement by leverage of the brace-wedge.

Mot1:TBP:DNA prehydrolysis state (top left) with a close up on the partly resolved (idle) brace (bottom left).
Mot1:TBP:DNA posthydrolysis state (top right) with a close up on the fully resolved brace-wedge (bottom right).

Intriguingly, this allosteric lock resembles the pincer 1-pincer 2 domain of antiviral innate immune sensor RIG-I (Extended Data Fig. 1 f). The pincer allosterically activates RIG-I’s

ATPase activity upon recognition of cognate 5'-triphosphate containing dsRNA termini[32,33]. It has been noted that RIG-I like immune sensors, DEAD box RNA helicases and Swi2/Snf2 proteins are evolutionary closely related[1], despite having very different biological functions. Both brace-wedge and pincer1-pincer2 wrap around RecA1B and lock the ATPase lobes, uncovering unexpected substrate-induced regulatory similarities between Swi2/Snf2 proteins and RIG-I family innate immune sensors (Extended Data Fig. 1 f).

The hook, following the wedge after an unstructured linker, stabilizes this state as it binds the DNA binding groove of TBP to prevent re-association of DNA. It thus acts as a chaperone. The anchor located C-terminal to the hook, is a helix that interacts with the inner surface of Mot1^{NTD} (Fig. 5 a, b). With the help of chemical crosslinking and mass spectrometry, we could corroborate the location and *de novo* build the anchor into the cryo-EM map (Extended Data Fig. 2 a – d). For all three complex conformations, most crosslinks satisfy the spatial restraints constituted by the diverse conformations (Extended Data Fig 2 e). A Mot1 mutant with a hook deletion (Mot1$^{\Delta hook}$, Δ1850-1868) showed a strongly reduced remodeling activity and demonstrated its pivotal functional importance since truncation of the outermost C-terminal 11 anchor residues (Mot1$^{\Delta anchor}$) alone did not diminish dissociation activity (Fig. 5 b, Extended Data Fig. 3 a, b). The finding that the hook of *Ct*Mot1 takes on the role of a TBP chaperone element was somewhat surprising, because the chaperone element was previously identified in the N-terminal 'latch' between HR 2 and 3 in microsporidian *Encephalitozoon cuniculi* Mot1 (*Ec*Mot1). In yeast Mot1, the latch between HR 2 and 3 harbors a nuclear localization sequence (NLS, 195 – 211), which is conserved in *Ct*Mot1 (208 - 224) and the human homologue BTAF1 (203 – 219) (Extended Data Fig. 3 c). Due to the high genome compaction of *Encephalitozoon cuniculi*[34], *Ec*Mot1 probably integrated the chaperone activity in the NLS bearing latch. Consistently, a Δlatch/Δhook double mutant showed similar activity as Mot1$^{\Delta hook}$ (Fig. 5 b, Extended Data Fig. 3 a, b).

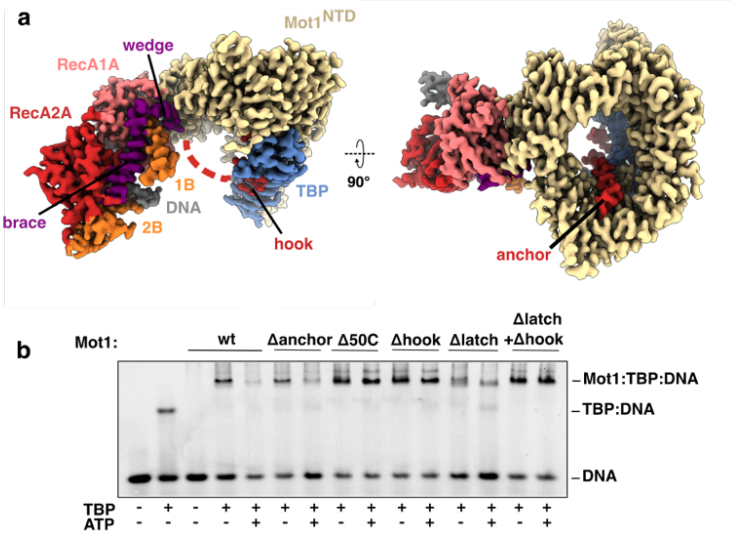


Figure 5: Allosteric regulation of TBP dislodgement by the hook.

a) Denoised cryo-EM map of Mot:TBP:DNA posthydrolysis state at 3.9 Å resolution from two rotated views. The brace extends into the bridge (dashed red line), where it contacts TBP on its concave site via a chaperoning hook. C-terminally, the anchor ascends the inner HEAT-repeat ring.
b) Native electrophoretic mobility shift assays with 5'-6-FAM upstream labeled DNA for Mot1^{wt} and diverse truncations show the effect of diverse bridge deletions on complex stability and TBP dissociation upon ATP addition.

Deleting the hook captures Mot1 in the act of TBP:DNA dislocation.

Identification of the hook as the chaperone element enabled us to trap complexes prior to TBP:DNA dissociation by deleting the last 50 residues of Mot1 (Mot1$^{\Delta 50C}$). Interestingly, Mot1$^{\Delta 50C}$, lacking the chaperone function, much more readily formed complexes with 2:2:2 stoichiometry than Mot1^{wt} (Extended Data Fig. 4 a, b) and showed an inefficient coupling of remodeling and ATPase activity (Extended Data Fig. 1 a, b), consistent with earlier results[19]. A minority of particles in Mot1^{wt} complexes also formed dimers and yielded low resolution reconstructions (Extended Data Fig. 4 c, d). Intriguingly, Mot1$^{\Delta 50C}$ after ATPγS addition and gel filtration yielded a stable C2-symmetric 2:2:2 Mot1$^{\Delta 50C}$:TBP:DNA complex in cryo-EM, which could be refined to a resolution of 2.8 Å. (Extended Data Fig. 4 e - h). Importantly, in contrast to the substrate recognition state (Fig. 6 a) TBP is still bound to DNA, yet the ATPase is in the posthydrolysis state (Fig. 6 b, only one protomer depicted). Thus, this state emerges as a "missing link": it captured a remodeling intermediate prior to hook-DNA competition in which Mot1 is in the process of TBP displacement.

In this intermediate state, all interfaces between Mot1, TBP and DNA are fully established, whereby in prior structures, at least one interface was disrupted. This indicates that the 2:2:2 Mot1$^{\Delta 50C}$:TBP:DNA complex is a transient intermediate, captured by the lack of the chaperone element (Extended Data Fig 4 e - h). The capability to form all interfaces results from a notable ~60° bent in DNA at Mot1^{CTD}, which is even more pronounced than the DNA bending by Mot1^{CTD} binding in the posthydrolysis state with intact hook.

The strong DNA bending is caused by RecA2B helix 2 protruding into the DNA minor groove where a conserved phenylalanine (F1588), preceded by an isoleucine (I1587), widens the minor groove, and even intercalates between adjacent base pairs (Fig. 6 b, c, Extended Data Fig. 5 a). Mot1^{IF1587AA} shows strongly diminished DNA interactions (Fig. 5 d, Extended Data Fig. 1 a), indicating that this interaction is indeed key for formation of stable DNA complexes. Interestingly, a similar functionally critical phenylalanine was found in the Swi2/Snf2 protein Cockayne Syndrome Protein B (CSB) at the RNA Pol II transcription coupled repair complex. There, the phenylalanine inserts at the junction between double-stranded and melted DNA and pushes the DNA duplex across RecA1 upon ADP-BeF_x mediated movement of RecA2 in order to translocate DNA by one base pair[35,36].

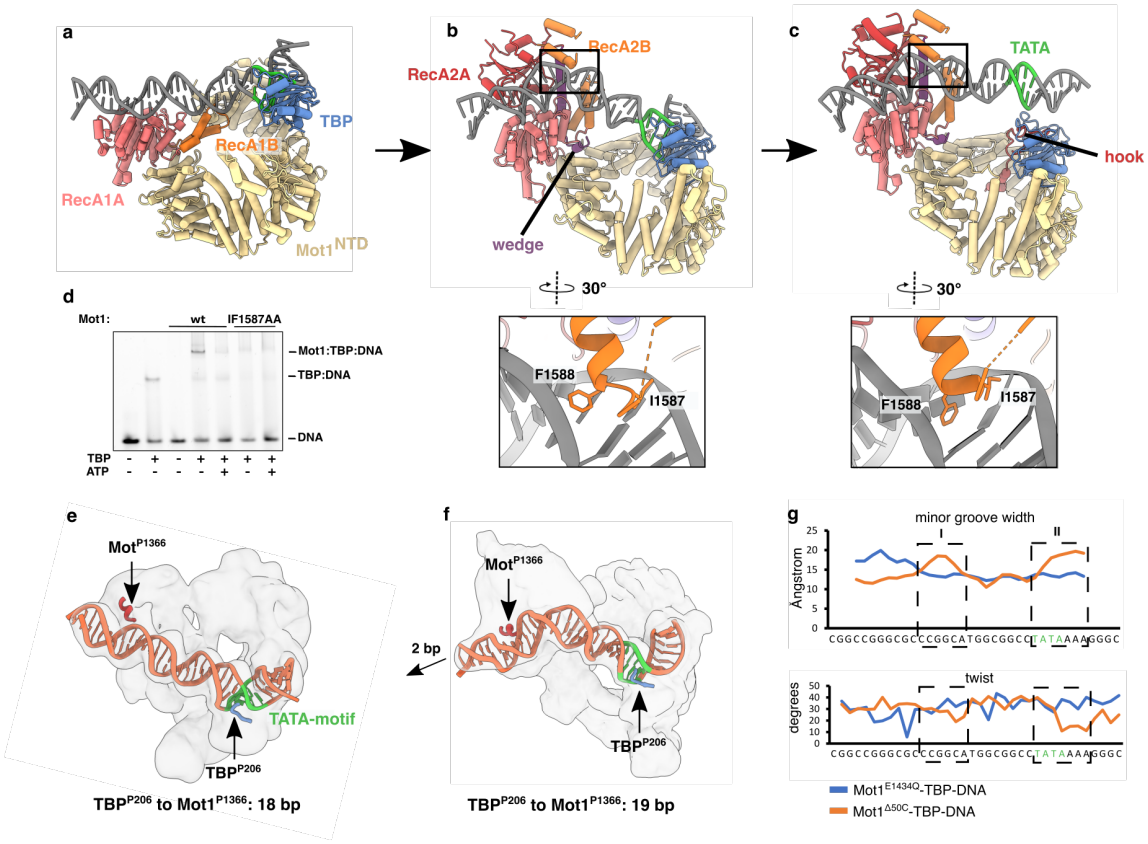


Figure 6: Effects of DNA bending by Mot1^{CTD}

a) Cryo-EM structure of the substrate recognition state (Mot1^{E1434Q}:TBP:DNA)
b) Cryo-EM structure of the posthydrolysis state of Mot1$^{\Delta 50C}$:TBP:DNA (top). The black squares mark the close up (bottom) region below depicting the RecA2B-minor groove interaction.
c) Cryo-EM structure of the posthydrolysis state of Mot1^{wt}:TBP:DNA (top). The black squares mark the close up (bottom) region below depicting the RecA2B-minor groove interaction.
d) Native electrophoretic mobility shift assay with 5'-6-FAM upstream labeled DNA showing the effect of TBP:DNA dissociation by point-mutated Mot1.
e) Complete 36 bp DNA in a gaussian-filtered cryo-EM map of the prehydrolysis state. The DNA register was derived from the TATA motif in a TBP:DNA crystal structure (PDB 1YTB)[37]. Prolines from TBP (P206) and Mot1 motif Ia (P1366) are indicated by arrows used as registry markers.
f) Complete 36 bp DNA in a gaussian-filtered cryo-EM map of the Mot1$^{\Delta 50C}$:TBP:DNA in the posthydrolysis state. An upstream directed 2 bp DNA shift relative to the substrate recognition state in e) allows a complete embedding of the DNA in its entire length. Arrows indicate the same residues as in g)
Comparison of minor groove width and local base twist of DNA between Mot1:TBP:DNA complexes in the pre- (Mot1^{E1434Q}) and posthydrolysis (Mot1$^{\Delta 50C}$) state. Locations of the upstream ATPase induced bending (I) and TATA box region (II) are indicated by dashed squares. The analysis was performed using the web 3DNA server[38].

The observed bending rotates the entire Mot1^{NTD}:TBP module via the DNA "lever" arm, opening space between NTD and CTD for the wedge to insert. To investigate whether this pivot movement might also exert a pulling force on DNA we analyzed the sequence register, using the TATA motif and the DNA termini in the map as guides. The map for Mot1^{E1434Q}:TBP:DNA with its clear outline for the DNA allowed us to model DNA using coordinates from a TBP:TATA-DNA crystal structure[31] (PDB: 1YTB). Orientated at the TATA motif at its defined location, we could delineate the entire 36 bp DNA sequence. This led to a convincing fit of upstream and downstream DNA into the map along their entire lengths and with the TATA box located at its canonical binding site (Fig. 5 e). Comparison with the DNA in the new dimer led to surprising observations. First, properly fitting the whole 36 bp long DNA sequence in the gaussian filtered map required a two base pair shift of the TATA box at TBP towards Mot1^{CTD} Furthermore, we noticed one additional base pair of DNA between TBP (P206) and RecA1 (P1366), which we used as fix points (18 vs 19 base pairs respectively, Fig. 5 e, f). In agreement with this sequence register is the highly resolved Mot1 dimer DNA density, which allows for a discrimination between purine and pyrimidine nucleobases at the TBP binding site (Extended Data Figure 5 b, c). This analysis suggests that rather than DNA groove tracking, CTD stably grips and rotates DNA, likely anchored by the intercalated phenylalanine. Along with the backward tilt, this dislodges the TATA box from TBP by 1-2 bps (Fig. 6 f, Extended Data Fig. 5 d). As a result, the TATA-motif would experience altered DNA mechanical parameters which might lead to a facilitated TBP dissociation by shifting the DNA:TBP complex from its thermodynamic stable minimum at the TATA box (Fig. 6 g).

Discussion

Here, we present the structural basis for TBP dissociation from the DNA TATA box by the Swi2/Snf2 ATPase Mot1, highlighting a path of allosteric activation and substrate dissociation without the need for processive, repetitive minor groove tracking.

Current cryo-EM studies on Swi2/Snf2 remodeler indicate that e.g. nucleosome sliding proceeds via bulge or twist-diffusion models of DNA translocation: the posthydrolysis ATPase induces a local twist defect that propagates around the nucleosome core[39], while multiple steps might sum up to a larger DNA bulge[3]. The twist defect model was recently specified for the Chd1 remodelers, where the ATPase induces a local A-form geometry in DNA[40]. Underlying these models is a groove tracking DNA motor via an inchworm translocation model[2]. Although Mot1 is not acting on a pre-bent nucleosomal DNA, it creates a similar distortion in DNA, revealing a general feature of Swi2/Snf2 proteins to either act on or induce bent DNA. In the case of Mot1, instead of ATP regulated groove tracking and DNA twisting, ATP regulated DNA bending and gripping via an intercalating phenylalanine appears to be at the core of its remodeling mechanism.

Upon initial TBP:DNA recognition, the HEAT repeat solenoidal NTD acts as a spring-like scaffold that can alter between spiral and ring like geometries to activate Mot1^{CTD} and flexibly

grip TBP during remodeling (Fig. 2 e, Suppl. Movie 1). The TBP-induced 90° bending of TATA DNA is decreased after binding of Mot1, which agrees with biophysical studies[28]. ATP binding leads to gripping of upstream DNA which could weaken TBP's DNA interactions further by torsional strain. Transiently dissociated DNA:TBP complexes are already trapped by the hook. Upon ATP hydrolysis, DNA is strongly bent and rotated, leading to a backward-tilting/rotating of $Mot1^{CTD}$ relative to $Mot1^{NTD}$:TBP. In case DNA is still bound at TBP, this movement slides and rotates the TATA box away from TBP to a higher energy state, which likely facilitates dissociation of TBP from DNA and occupation of TBP's DNA binding site by the hook chaperone element (Fig. 7, Suppl. Movie 2). Thus, ATP-hydrolysis confers thermodynamic directionality towards TBP release by inhibiting DNA-rebinding and offering the hook as a competing binding partner. In addition, our data suggest that a final DNA release from Mot^{CTD} is promoted by one or more ATP-binding and/or hydrolysis events. Although this might be specific for Mot1, which has a reduced affinity towards DNA in the presence of ATPγS[28], it has recently been shown that the SWI/SNF nucleosome remodeler undergoes ATP-hydrolysis-dependent remodeling before leaving their nucleosome substrate[41].

An interesting yet unexplained observation is the ability of Mot1 to form dimers in both pre- and posthydrolysis states. While the biochemical role of the Mot1 dimer is probably to stabilize TBP-$Mot1^{NTD}$ interactions and helping pulling TATA DNA away from TBP, the biological role remains to be revealed in future studies. It is conceivable that chromatin regions with a high accumulation and proximity of promoters and TBP, e.g. in transcriptionally active chromatin regions such as topologically associated domains (TADs), would favor dimers. This is an area that would merit future investigations. In any case, the proposed remodeling model is consistent with a large body of prior biochemical work including unbending of DNA[28], the augmentative role of a second Mot1 protein[21,22], a two-step displacement mechanism[28], and lack of processive translocation[20].

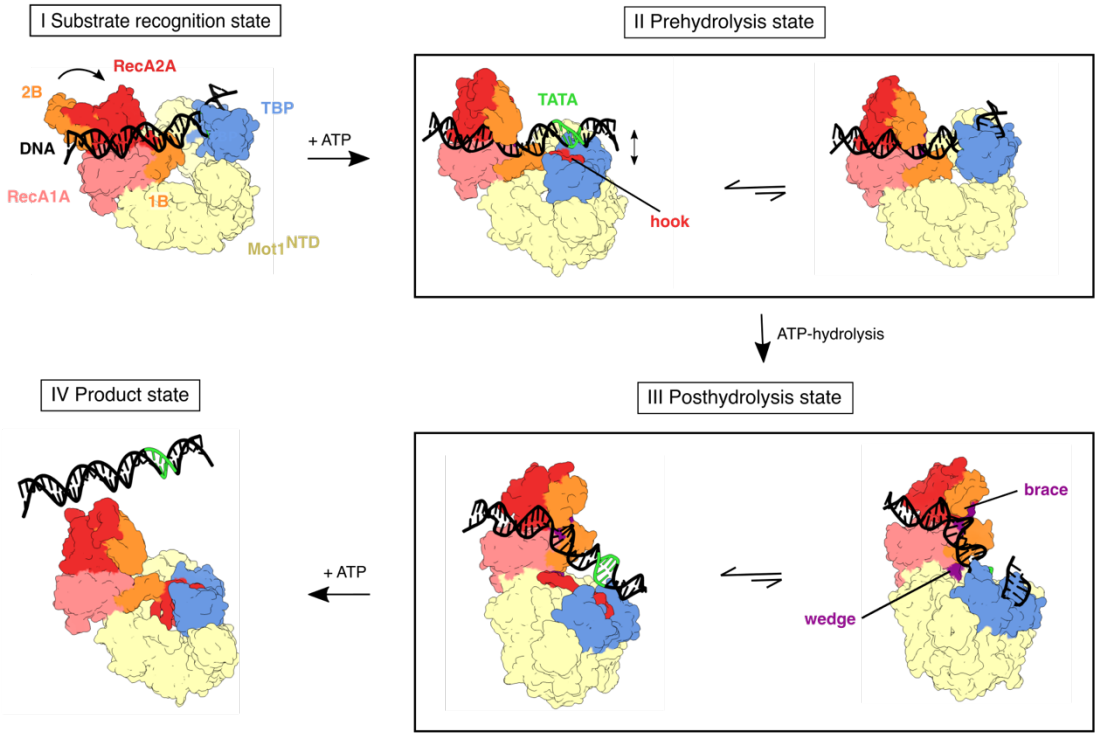


Figure 7: Structure-based model of TBP displacement from DNA by Mot1.

Mot1 binds TBP via the N-terminal HEAT repeats and the C-terminal ATPase contacts upstream DNA (I Substrate recognition state). Binding of ATP leads to tight gripping of DNA (RecA1 and A2) which causes a DNA strain that destabilizes TBP:TATA-motif binding. TBP dissociated from DNA is trapped by the hook (II Prehydrolysis state). ATP hydrolysis further bends DNA accompanied by a pivot movement of Mot1^{CTD}, sliding and rotating the TATA-box away from TBP and shifting the thermodynamic equilibrium further towards a stable hook engagement on TBP (III Posthydrolysis state). Finally, subsequent ATP-binding might release Mot1^{CTD} from upstream DNA leading to complete dissociation of Mot1:TBP (IV Product state).

In summary, our work reveals some unexpected features that considerably expands our understanding of Swi2/Snf2 protein mechanisms. Instead of groove tracking, Mot1 rather appears to grip and pull DNA through an intercalating phenylalanine, which deviates from the inchworm like mechanism proposed for many SF2 translocases[2,4]. Rather, single or repetitive cycles of DNA gripping, bending and release in conjunction with domain tilting could dislodge DNA from its thermodynamic stable position. This remodeling mechanism may help draw general principles that could be relevant for other Swi2/Snf2 proteins that lack processive translocation such as the SWR1 histone exchange complex[9]. In addition, Mot1 mechanistically bridges the gap between Swi2/Snf2 proteins and the closely related DEAD box helicases, which have been shown to be able to locally bend and unwind duplex RNA without translocation[32].

Methods

Protein expression and purification

The full-length *Chaetomium thermophilum* Mot1 sequence (1-1886) was cloned into a pET29 vector (Novagen, Germany) harboring a C-terminal Strep II-tag. The wildtype and mutant constructs were recombinantly expressed in *Escherichia coli* Rosetta (DE3) cells (Novagen) for 16 hours at 18 °C in Turbo Broth™ (Molecular Dimensions). Expression was induced by addition of 0.5 mM IPTG. The cells were resuspended in high salt purification buffer (50 mM Hepes pH 8.0, 800 mM NaCl, 10 mM β-mercaptoethanol, 5% glycerol) sonicated three times for four minutes (50% power, 30% duration) and centrifuged at 18,000 rpm for one hour at 4°C. Firstly, the lysate was run over a StrepTrap HP column (GE Healthcare) and eluted with purification buffer (50 mM Hepes pH 8.0, 400 mM NaCl, β-mercaptoethanol, 5% glycerol) supplemented with 2.5 mM desthiobiotin. Then the main peak fraction of the protein was collected, diluted to 100 mM NaCl and further purified via anion-exchange chromatography (HiTrap Q, GE Healthcare) applying a salt gradient (50 mM Hepes pH 8.0, 100-1000 mM NaCl, 2 mM DTT). The protein was concentrated to ~15 mg/ml in 20 mM Hepes pH 7.5, 60 mM KCl and 10% glycerol, flash-frozen in liquid nitrogen and stored at -80°C.

Chaetomium thermophilum TBP cloned on a pET29 vector with an N-terminal His_6-tag was expressed in *Escherichia coli* Rosetta (DE3) cells (Novagen) for 16 hours at 18°C in Turbo Broth™ (Molecular Dimensions). Expression was induced by addition of 0.5 mM IPTG. The cells were resuspended in high salt purification buffer (50 mM Hepes pH 8.0, 1500 mM NaCl, 10 mM Imidazol, 10 mM β-mercaptoethanol, 5% glycerol) and TBP was purified via a HisTrap HP (GE Healthcare), eluted with 250 mM Imidazol in elution buffer (50 mM Hepes pH 8.0, 200 mM NaCl, 250 mM Imidazol, 10 mM β-mercaptoethanol, 5% glycerol). To cleave off the His-tag, 30:1 (w/w) TEV protease was added to the eluate, which was dialyzed against dialysis buffer (20 mM Hepes pH 8.0, 200 mM NaCl, 2 mM DTT) for at least 12 hours at 4°C. To separate the His-tagged TEV protease and the cleaved His-tags and uncut TBP from untagged TBP, a reverse Ni-NTA purification via HisTrap HP (GE Healthcare) was performed additionally. The TBP-containing flow-through and wash fractions were applied onto a Heparin HP column (GE Healthcare) and eluted with a salt gradient (20 mM Hepes pH 8.0, 200-1000 mM NaCl, 2 mM DTT). At last, the main peak fractions were purified via preparative gel filtration on a Superdex 75 10/300 GL column (GE Healthcare) in storage buffer (20 mM Hepes pH 7.5, 200 mM NaCl, 2 mM DTT). The protein was concentrated to ~10 mg/ml in storage buffer supplemented with 10% glycerol, flash-frozen in liquid nitrogen and stored at -80°C.

Electrophoretic mobility shift assays

Native polyacrylamide gel electrophoresis to analyze Mot1-catalyzed disruption of TBP:DNA were essentially performed as previously described[42] with some modifications. Fluorescently labelled 48 bp dsDNA (10 nM, 5'-CAGTACGGCCGGGCGCCCGGCATGGCGGCCTATAAA

AGGGGGTGGAAT–3', with 6-FAM label on the 5' end of the reverse strand) was incubated with TBP (10 nM) for 10 minutes in sample buffer (0.1 mg/ml bovine serum albumin (BSA), 50 nM poly(dG-dC), 4% glycerol, 0.1% Brij 58, 60 mM KCl, 5 mM Tris pH 8, 5 mM magnesium acetate, 1 mM DTT). Mot1 (50 nM) was added and incubated for 10 minutes, followed by eventual addition of nucleotides. For comparisons of TBP dissociation between Mot1 mutants, 50 µM ATP was added and incubated for 3 minutes. The reaction was quenched with 50 mM ADP.

5% polyacrylamide gels (5% acrylamide from a 30%:0.8% acrylamide:bis-acrylamide stock, 2.5% glycerol, 190 mM glycine, 10 mM magnesium acetate, 2.5 mM Tris pH 8.3) were pre-run for one hour and after sample loading run at 160 V for 50 minutes at 4°C. The gels were imaged at a Typhoon FLA 9000 imager (GE Healthcare), subsequent quantifications were performed with *ImageJ*[43].

ATPase assays

For determination of the ATP hydrolysis rate of Mot1, an assay coupling ATP hydrolysis to NADH oxidation was performed as previously reported[44]. Either Mot1 (150 nM) alone or added to preformed TBP (180 nM):DNA (240 nM) complexes were incubated for 10 minutes at 4°C. Reconstituted complexes were mixed with 0.5 mM phosphoenolpyruvate, 1 mM ATP or ATPγS, 0.1 mM NADH, 25 U/ml lactate dehydrogenase pyruvate kinase mix (Sigma) and ATPase buffer (20 mM HEPES pH 8.0, 60 mM KCl, 1 mM DTT, 5 mM $MgCl_2$, 0.1 mg/ml BSA) in a final volume of 50 µl. Changes in fluorometrical absorbance of NADH at 340 nm were monitored in non-binding black 96-well plates (Greiner Bio-One) with a Tecan Infinite M100 (Tecan) using 343 nm for excitation and emission of 448 nm at 30 °C. ATP turnover was calculated using maximal initial linear rates, corrected for a buffer blank.

Analytical Gel Filtration

The Mot1:TBP:DNA-complex was reconstituted in a stepwise manner: 36 bp TATA-box dsDNA (5'-CGGCCGGGCGCCCGGCATGGCGGCCTATAAAAGGGC-3') (30 µM) was incubated with TBP (25 µM) for 10 minutes at room temperature followed by addition of Mot1 (25 µM). 3 mM ATPγS was added, followed by incubation for another 10 minutes in reconstitution buffer (Hepes pH 7.5, 60 mM KCl, 5 mM CaCl). The complex was loaded on a Superose 6 Increase 3.2/300 column (GE Healthcare) installed at an Äkta Micro (GE Healthcare).

Crosslinking-mass spectrometry

60-70 μg of *in vitro* reconstituted Mot1:TBP:DNA complex incubated with 3 mM ATPγS was crosslinked with 1:1 isotopically labeled (d0/d12) BS3 (Bis[sulfosuccinimidyl] suberate, Creative molecules) in a thermomixer at 35°C for 30 min at 1200 rpm. The reaction was quenched with 100 mM ammonium bicarbonate for 10 min at 20°C and supplemented with 8 M urea to a final concentration of 6 M. After reduction and alkylation crosslinked proteins were digested with Lys-C (1:50 w/w, Wako) for 2 h, diluted with 50 mM ammonium bicarbonate to 1 M urea and digested with trypsin (1:50 w/w, Promega) for 16-18 h. Crosslinked peptides were purified by reversed-phase chromatography using C18 cartridges (Sep-Pak, Waters). Cross-linked peptides were enriched by peptide size exclusion chromatography and analyzed by tandem mass spectrometry (Orbitrap Elite, Thermo Scientific). Fragment ion spectra were searched, and crosslinks identified by the dedicated software program xQuest[45,46]. The quality of all cross-link were manually validated by inspection of the respective mass spectra. For visualization of crosslinks in structures, the program Xlink analyzer was used[47].

Sample and grid preparation for cryo-EM

36 bp TATA-box dsDNA (5'-CGGCCGGGCGCCCGGCATGGCGGCCTATAAAAGGGC-3') (100 μM) was incubated with TBP (90 μM) for 10 minutes at 4 °C followed by addition of wildtype or mutant Mot1 (80 μM) respectively. For the posthydrolysis, bridge-deletion and product conformations, 1 mM ATPγS was added before gel filtration and incubated for 10 minutes in reconstitution buffer (HEPES pH 7.5, 60 mM KCl, 5 mM MgCl) at 4 °C. The complex was run over a Superdex 200 Increase 5/150 GL column or a Superose 6 Increase 3.2/300 GL (GE Healthcare, Germany). The main peak fraction was diluted to a final concentration of 0.5 mg/ml with reconstitution buffer and kept on ice. The product state was achieved by additional incubation with ATPγS (1 mM) 20 minutes before grid preparation in reconstitution buffer. For the prehydrolysis conformation, 1 mM ADP, 1 mM BeF_2 and 3 mM NaF were incubated with Mot1:TBP:DNA for 20 minutes at 25°C before grid preparation. As a detergent, β-octyl glucoside (Roth, Germany) was added to all samples at a final concentration of 0.05%. 4.5 μL of the protein solution were applied to Quantifoil R2/1 holey carbon grids and 20 seconds pre-incubated before blotting, then frozen in liquid ethane using a Leica EM GP (Leica, 10°C and 95% humidity).

Cryo-EM data collection

Movies of particles were embedded in vitreous ice and collected at liquid nitrogen temperature using a Titan Krios G3 transmission electron microscope (ThermoFisher) equipped with a K2 Summit direct electron detector (Gatan) and a GIF Quantum LS Imaging Filter (Gatan). The movies were recorded in counting mode using EPU acquisition software (ThermoFisher) at 130k

x magnification with a pixel size of 1.059 or 1.046 Å/pixel, and nominal defocus range of -1.4 to -3.2 µm. The total electron dosage of each movie was ~45-60 e/Å^2 with a nominal exposure rate of 1.1-1.5 electrons Å^{-2} s^{-1} per frame, fractionated into 40 movie frames with 250 ms/frame exposure time.

Cryo-EM data processing

Beam-induced motions of particles were corrected with MotionCor2[48] and contrast transfer function parameters were estimated using CTFFIND[49] if not stated otherwise.

Mot1:TBP:DNA posthydrolysis state: Particles were picked employing the Laplacian-of-Gaussian picker in Relion[50]. Multiple rounds of successive 2D classification were performed and particles submitted for 3D classification using an initial Mot1 cryo-EM map as a template (downfiltered to 60 Å). The most coherent class was 3D refined, CTF refined and Bayesian polished. Post-processing yielded a map at an average resolution of 3.9 Å.

Mot1:TBP:DNA recognition state: For picking, the template-based picker in cryoSPARC[51] was used, with the posthydrolysis state as a template. After successive rounds of 2D classification Topaz picker[52] was trained two times with interspersed 2D classifications for particle sorting. Then, particle coordinates were transferred to Relion and submitted to 3D classification, using the posthydrolysis state as a template (downfiltered to 60 Å). The most coherent class was 3D refined and submitted to CTF refinement, Bayesian polishing and post processing. The final map was refined to an average resolution of 6.6 Å.

Mot1:TBP:DNA product state: particles were picked using the template picker in cryoSPARC based on the down-filtered posthydrolysis state as a template. Several rounds of 2D classification were followed by two times training of the Topaz picker with interspersed 2D classifications. To enhance the overall quality of the reconstruction, 2D classes were rebalanced before calculation of an ab-initio model. The 3D initial model was submitted to non-uniform refinement, which yielded an average resolution of 4.4 Å.

Mot1:TBP:DNA prehydrolysis state: For particle picking the blob picker in cryoSPARC was used. After several rounds of 2D classifications, monomer and dimer classes were processed separately. Particles belonging to good classes were used to train the Topaz picker two times with interspersed 2D classifications. An ab-initio dimer model was submitted to non-uniform refinement and reconstructed to an average resolution of 4.4 Å. For the monomer classes a 3D variability analysis was performed. Three representative maps were submitted to heterogenous refinement and the map with the most complete ATPase moiety was further non-uniformly refined to an average resolution of 3.6 Å. To enhance the map quality of Mot1NTD and Mot1CTD, the respective regions were masked separately, and focused refinements was performed. This led to a 3.7 Å resolution reconstruction for Mot1^{CTD} and a 3.5 Å resolution reconstruction for Mot1^{NTD}.

Mot1$^{\Delta 50C}$:TBP:DNA: Beam-induced motions of particles were corrected using UNBLUR (*cis*TEM) or MotionCor2 (Relion) on whole frames[48,53]. Contrast transfer function (CTF) parameters were estimated using *cis*TEM[53]. The particles were automatically picked *ab initio* using Warp[54] or soft-edged disk templates internally generated in *cis*TEM. The initial 3D reconstructions were generated *ab initio* using *cis*TEM. The particles for final reconstructions were boxed and extracted from the micrographs in Relion with the particle coordinates exported from Warp using the PyEM scripts developed by Daniel Asarnow (https://github.com/asarnow/pyem).

The initial 2D and 3D classifications were carried out using *cis*TEM. The final 2D and 3D classifications were carried in Relion using the *ab initio* 3D references generated in *cis*TEM. The particles were Bayesian polished in Relion. The initial 3D refinements were carried out in Relion and the final 3D refinements were carried out using *cis*TEM. The cryo-EM maps were *B*-factor sharpened globally using *cis*TEM or *Coot*[55], or locally sharpened using Relion. The number of movies and particles used towards final reconstructions are in Table S1.

Model building and refinement

All protein models were real-space refined using PHENIX[56], and evaluated using *Coot*[55] and the MolProbity server[57]. Intermittently, every model was geometry optimized in a molecular dynamics environment using ISOLDE[58]. In UCSF Chimera[59], the Mot1 crystal structure[19] (PDB code: 6G7E) was split in three parts and rigid-body docked in the density of the prehydrolysis complex. A yeast TBP crystal structure[37] (PDB code: 1YTB) was modified after being docked into the respective part of the map accordingly. Following rebuilding in *Coot*, this model was used for rigid-body docking into the maps of the other conformations. Guided by crosslink mass spectrometry data visualized with Xlink analyzer[47] in UCSF Chimera, a peptide comprising the crosslinked bridge lysine (K1855) was placed in the respective density to yield a permitted Euclidean Cα-Cα crosslinker length of < 30 Å. DNA was built in *Coot* after adding externally generated base pair and parallel plane restraints with LIBG[60] and later refined including base pairing and base stacking restraints in Phenix[56]. For presentation purposes, the cryo-EM maps were denoised using DeepEMhancer[61]. The symmetric and asymmetric reconstruction cryo-EM maps were deposited in the Electron Microscopy Databank (EMDB)[62] and the coordinates of the atomic models were deposited in the Protein Data Bank (PDB)[63]. The figures were generated using UCSF Chimera[59] and ChimeraX[64].

References

1. Fairman-Williams, M. E., Guenther, U. P. & Jankowsky, E. SF1 and SF2 helicases: Family matters. *Current Opinion in Structural Biology* 20, 313–324 (2010).

2. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature Reviews Molecular Cell Biology 2017 18:7* 18, 407–422 (2017).

3. Jungblut, A., Hopfner, K. P. & Eustermann, S. Megadalton chromatin remodelers: common principles for versatile functions. *Current Opinion in Structural Biology* 64, 134–144 (2020).

4. Yan, L. & Chen, Z. A Unifying Mechanism of DNA Translocation Underlying Chromatin Remodeling. *Trends in Biochemical Sciences* 45, 217–227 (2020).

5. Mueller-Planitz, F., Klinker, H. & Becker, P. B. Nucleosome sliding mechanisms: New twists in a looped history. *Nature Structural and Molecular Biology* 20, 1026–1032 (2013).

6. Farnung, L., Vos, S. M., Wigge, C. & Cramer, P. Nucleosome–Chd1 structure and implications for chromatin remodelling. *Nature 2017 550:7677* 550, 539–542 (2017).

7. Li, M. *et al.* Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature 2019 567:7748* 567, 409–413 (2019).

8. Eustermann, S. *et al.* Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature 2018 556:7701* 556, 386–390 (2018).

9. Willhoft, O. *et al.* Structure and dynamics of the yeast SWR1 : nucleosome complex. 199, (2018).

10. He, S. *et al.* Structure of nucleosome-bound human BAF complex. *Science (1979)* 367, 875–881 (2020).

11. Farnung, L., Ochmann, M., Engeholm, M. & Cramer, P. Structural basis of nucleosome transcription mediated by Chd1 and FACT. *Nature Structural and Molecular Biology* 28, 382–387 (2021).

12. Dasgupta, A., Juedes, S. A., Sprouse, R. O. & Auble, D. T. Mot1-mediated control of transcription complex assembly and activity. *EMBO Journal* 24, 1717–1729 (2005).

13. Sprouse, R. O., Brenowitz, M. & Auble, D. T. Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. *EMBO Journal* 25, 1492–1504 (2006).

14. Sprouse, R. O. *et al.* Function and structural organization of Mot1 bound to a natural target promoter. *Journal of Biological Chemistry* 283, 24935–24948 (2008).

15. Zentner, G. E. & Henikoff, S. Mot1 Redistributes TBP from TATA-Containing to TATA-Less Promoters. *Molecular and Cellular Biology* 33, 4996–5004 (2013).

16. Wollmann, P. *et al.* Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. *Nature* (2011) doi:10.1038/nature10215.

17. Hopfner, K. P., Gerhold, C. B., Lakomek, K. & Wollmann, P. Swi2/Snf2 remodelers: Hybrid views on hybrid molecular machines. *Current Opinion in Structural Biology* Preprint at https://doi.org/10.1016/j.sbi.2012.02.007 (2012).

18. Butryn, A. *et al.* Structural basis for recognition and remodeling of the TBP:DNA:NC2 complex by Mot1. *Elife* (2015) doi:10.7554/eLife.07432.

19. Butryn, A., Woike, S., Shetty, S. J., Auble, D. T. & Hopfner, K.-P. Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state. *Elife* 7, 1–12 (2018).

20. Auble, D. T. & Steggerda, S. M. Testing for DNA Tracking by MOT1, a SNF2/SWI2 Protein Family Member. *Molecular and Cellular Biology* 19, 412–423 (2015).

21. Viswanathan, R., True, J. D. & Auble, D. T. Molecular mechanism of Mot1, a TATA-binding protein (TBP)-DNA dissociating enzyme. *Journal of Biological Chemistry* (2016) doi:10.1074/jbc.M116.730366.

22. Heiss, G. *et al.* Conformational changes and catalytic inefficiency associated with Mot1-mediated TBP–DNA dissociation. *Nucleic Acids Research* (2019) doi:10.1093/nar/gky1322.

23. Willhoft, O. *et al.* Crosstalk within a functional INO80 complex dimer regulates nucleosome sliding. *Elife* 6, 1–21 (2017).

24. Dürr, H., Körner, C., Müller, M., Hickmann, V. & Hopfner, K. P. X-Ray structures of the sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell* 121, 363–373 (2005).

25. Lewis, R., Dürr, H., Hopfner, K. P. & Michaelis, J. Conformational changes of a Swi2/Snf2 ATPase during its mechano-chemical cycle. *Nucleic Acids Research* 36, 1881–1890 (2008).

26. Hauk, G., McKnight, J. N., Nodelman, I. M. & Bowman, G. D. The Chromodomains of the Chd1 Chromatin Remodeler Regulate DNA Access to the ATPase Motor. *Molecular Cell* 39, 711–723 (2010).

27. Farnung, L., Vos, S. M., Wigge, C. & Cramer, P. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* (2017) doi:10.1038/nature24046.

28. Moyle-Heyrman, G., Viswanathan, R., Widom, J. & Auble, D. T. Two-step mechanism for modifier of transcription 1 (Mot1) enzyme-catalyzed displacement of TATA-binding protein (TBP) from DNA. *Journal of Biological Chemistry* 287, 9002–9012 (2012).

29. Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* (2017) doi:10.1038/nature22036.

30. Clapier, C. R., Verma, N., Parnell, T. J. & Cairns, B. R. Cancer-Associated Gain-of-Function Mutations Activate a SWI/SNF-Family Regulatory Hub. *Molecular Cell* 1–14 (2020) doi:10.1016/j.molcel.2020.09.024.

31. Auble, D. T., Wang, D., Post, K. a I. W. & Hahn, S. Molecular Analysis of the SNF2 / SWI2 Protein Family Member MOT1 , an ATP-Driven Enzyme That Dissociates TATA- Binding Protein from DNA. 17, 4842–4851 (1997).

32. Rehwinkel, J. & Gack, M. U. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nature Reviews Immunology* 20, 537–551 (2020).

33. Rawling, D. C., Kohlway, A. S., Luo, D., Ding, S. C. & Pyle, A. M. The RIG-I ATPase core has evolved a functional requirement for allosteric stabilization by the Pincer domain. *Nucleic Acids Research* 42, 11601–11611 (2014).

34. Corradi, N., Pombert, J. F., Farinelli, L., Didier, E. S. & Keeling, P. J. The complete sequence of the smallest known nuclear genome from the microsporidian Encephalitozoon intestinalis. *Nature Communications* 1, (2010).

35. Kokic, G., Wagner, F. R., Chernev, A., Urlaub, H. & Cramer, P. Structural basis of human transcription–DNA repair coupling. *Nature 2021 598:7880* 598, 368–372 (2021).

36. Yan, C. *et al.* Mechanism of Rad26-assisted rescue of stalled RNA polymerase II in transcription-coupled repair. *Nature Communications* 12, 1–12 (2021).

37. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. Crystal structure of a yeast TBP/TATA-box complex. *Nature 1993 365:6446* 365, 512–520 (1993).

38. Zheng, G., Lu, X. jun & Olson, W. K. Web 3DNA - A web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Research* 37, 240–246 (2009).

39. Yan, L. & Chen, Z. A Unifying Mechanism of DNA Translocation Underlying Chromatin Remodeling. *Trends in Biochemical Sciences* 1–11 (2019) doi:10.1016/j.tibs.2019.09.002.

40. Nodelman, I. M. *et al.* Nucleosome recognition and DNA distortion by the Chd1 remodeler in a nucleotide-free state. *Nature Structural & Molecular Biology* 29, 121–129 (2022).

41. Tilly, B. C. *et al.* In vivo analysis reveals that ATP-hydrolysis couples remodeling to SWI/SNF release from chromatin. *Elife* 10, 1–27 (2021).

42. Darst, R. P., Wang, D. & Auble, D. T. MOT1-catalyzed TBP-DNA disruption: Uncoupling DNA conformational change and role of upstream DNA. *EMBO Journal* 20, 2028–2040 (2001).

43. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* vol. 9 671–675 Preprint at https://doi.org/10.1038/nmeth.2089 (2012).

44. Kiianitsa, K., Solinger, J. A. & Heyer, W.-D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Analytical Biochemistry* 321, 266–271 (2003).

45. Herzog, F. *et al.* Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science (1979)* 337, 1348–1352 (2012).

46. Rinner, O. *et al.* Identification of cross-linked peptides from large sequence databases. *Nature Methods* (2008) doi:10.1038/nmeth.1192.

47. Kosinski, J. *et al.* Xlink analyzer: Software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *Journal of Structural Biology* 189, 177–183 (2015).

48. Zheng, S. Q. *et al.* MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* 14, 331–332 (2017).

49. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* 192, 216–221 (2015).

50. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7, (2018).

51. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* 14, 290–296 (2017).

52. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods* 16, 1153–1160 (2019).

53. Grant, T., Rohou, A. & Grigorieff, N. CisTEM, user-friendly software for single-particle image processing. *Elife* 7, (2018).

54. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nature Methods* 16, 1146–1152 (2019).

55. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* 66, 486–501 (2010).

56. Afonine, P. v. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallographica Section D: Structural Biology* 74, 531–544 (2018).

57. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* 66, 12–21 (2010).

58. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology* 74, 519–530 (2018).

59. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605–12 (2004).

60. Brown, A. *et al.* Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallographica Section D: Biological Crystallography* 71, 136–153 (2015).

61. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology* 4, 1–8 (2021).

62. Lawson, C. L. *et al.* EMDataBank unified data resource for 3DEM. *Nucleic Acids Research* 44, D396–D403 (2016).

63. Berman, H. M. *et al.* The protein data bank. *Acta Crystallographica Section D: Biological Crystallography* 58, 899–907 (2002).

64. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* 27, 14–25 (2018).

Acknowledgements

We thank Felix Metzner, Franziska Kunert and Kevin Schall for discussions. We thank Achilleas Frangakis for generous help with initial cryo-EM studies. This work was supported by the German Research Council (RTG1721, CRC1064 and Gottfried Wilhelm Leibniz-Prize) and the European Research Council (Advanced Grant 833613) to K.-P.H.

Author contributions

S.W. prepared the proteins and performed biochemical analysis. S.W and S.E prepared cryo-EM samples and performed the structure determination. S.W built atomic models. A.B. investigated Mot1 variants and together with S.E. performed initial cryo-EM analyses. S.W. and J.J. processed and refined electron microscopy data. S.J.W. performed ATPase assays. G.H. and F.H. did the crosslink-mass spectrometry analysis. J.B and K.L. helped with electron microscopy data collection. S.W., S.E. and K.-P.H designed the overall study and analyzed the results. K.-P.H. and S.W. wrote the paper with contributions from all authors. K.-P.H. provided funding.

Competing interest declaration

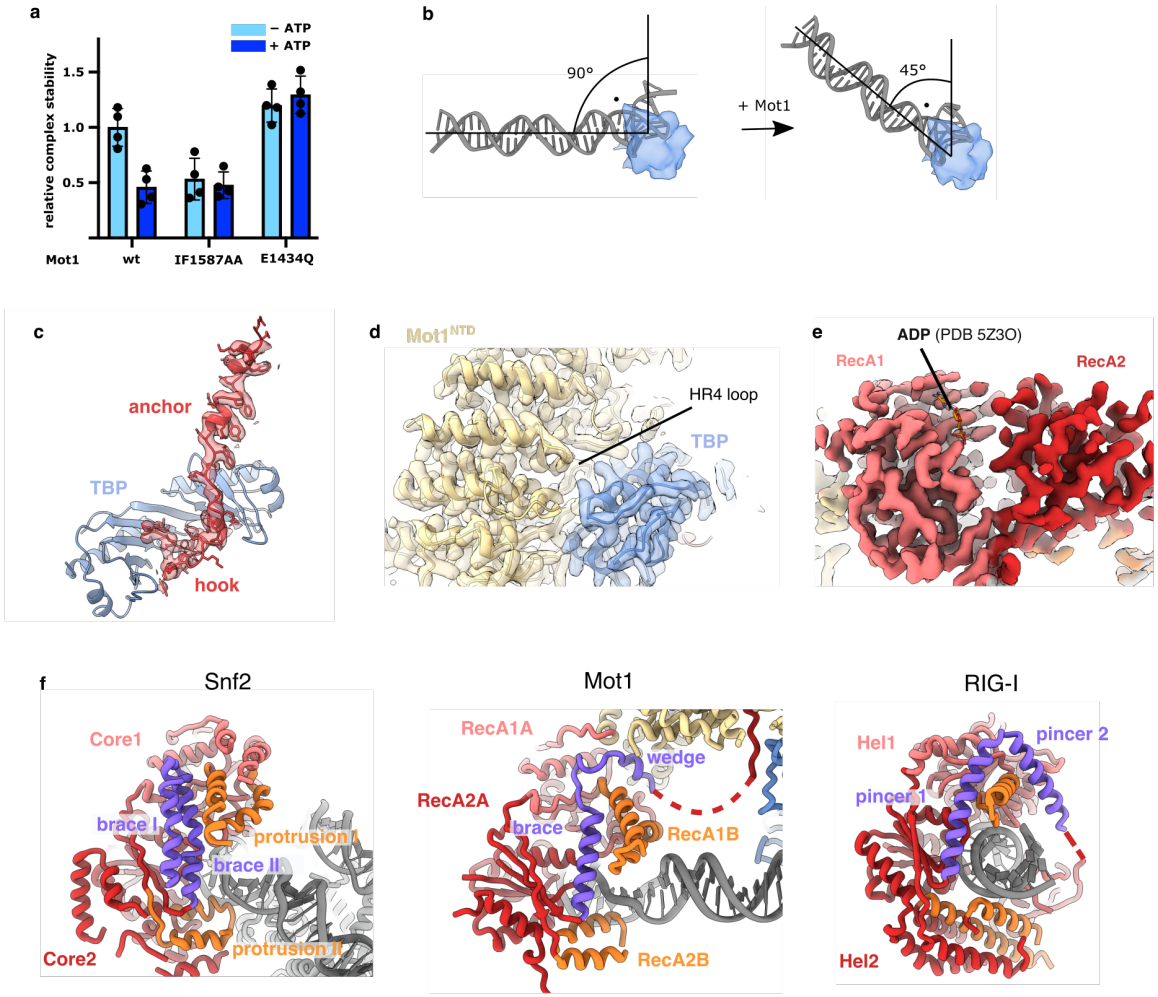
The authors declare no competing interest

Data availability

The electron density reconstruction and final models were deposited at the Electron Microscopy Data Base (EMDB) with accession codes EMD-14534 (recognition state), EMD-14762 (prehydrolysis state), EMD-14562 (posthydrolysis state), EMD-14583 (Mot1$^{\Delta 50C}$ posthydrolysis dimer) and EMD-14554 (product state). The coordinates were deposited at Protein Data Bank (PDB) with accession codes 7Z7N (recognition state), 7ZKE (prehydrolysis state), 7Z8S (posthydrolysis state) and 7ZB4 (Mot1$^{\Delta 50C}$ posthydrolysis dimer).

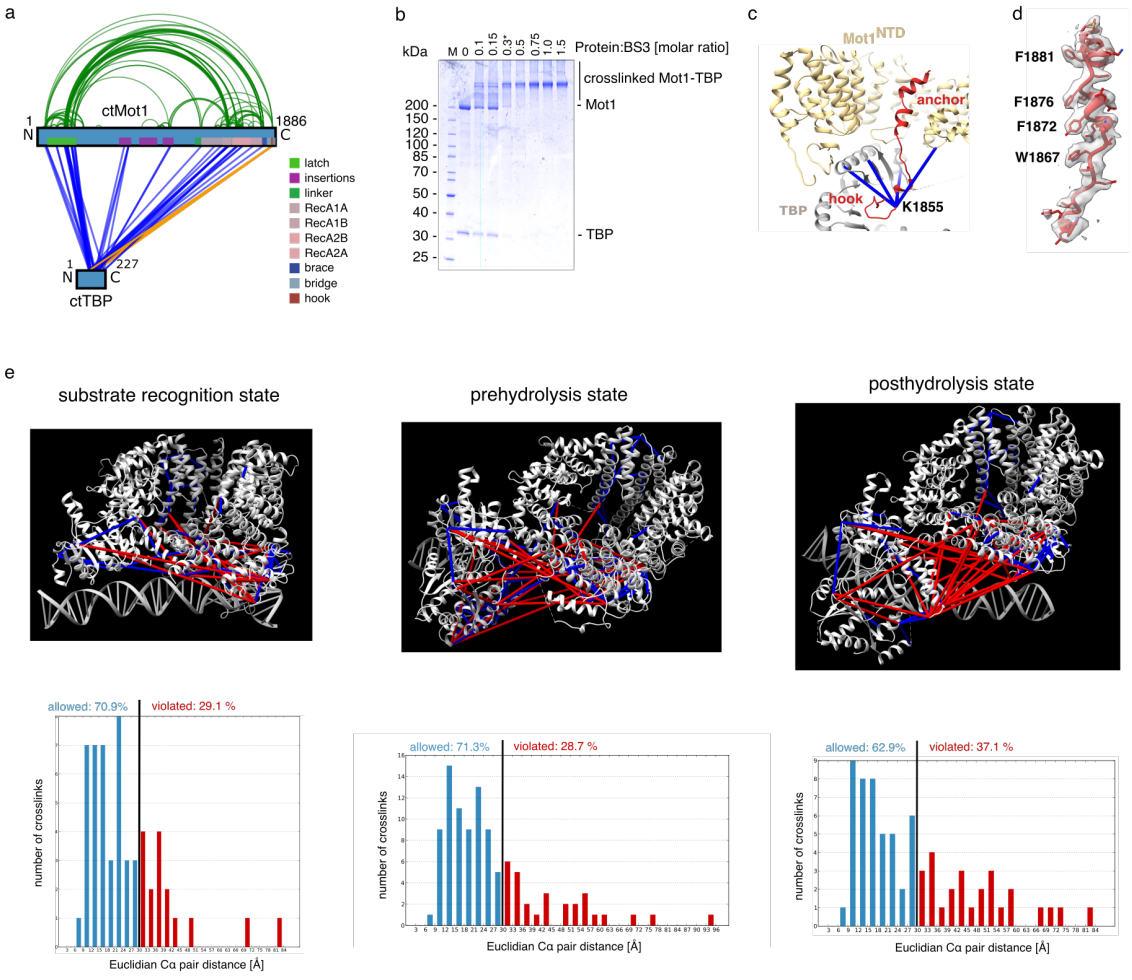
Materials and correspondence

Correspondence and request for materials should be addressed to Karl-Peter Hopfner (hopfner@genzentrum.lmu.de)



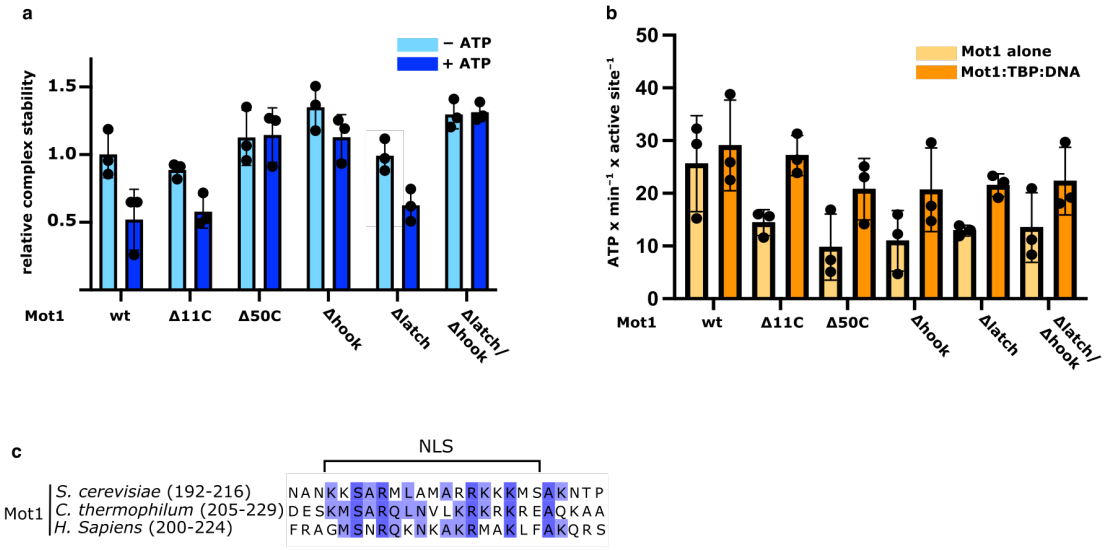
Extended Data Fig. 1: Enzymatic and structural details on Mot1 substrate recognition, pre- and posthydrolysis states and Swi2/Snf2 brace comparison.

a) Quantification of the native electrophoretic mobility shift assay in Fig. 1 b/ Fig 6 d. The bars show the mean from four technical replicates. The error bar represents +/- standard deviation.
b) Structure TBP:DNA (left) and after binding of Mot1 (right), which unbends the DNA by 45°.
c) Example density with built in hook and anchor domains in the Mot1 prehydrolysis complex.
d) Example density with built in part of the HEAT repeat array ($Mot1^{NTD}$), showing the engagement with the convex site of TBP.
e) Cryo-EM density of the $Mot1^{CTD}$ nucleotide binding pocket from the Mot1 posthydrolysis complex. ADP, docked from a Snf2 structure, aligned via RecA1, illustrates the lack of any nucleotide density.
f) Swi2/Snf2 family ATPases in their closed conformation feature a conserved helical domain that folds from lobe 2 back to lobe 1. Yeast Snf2 (PDB 5Z3I)[24] exhibits a double brace. In Mot1 and RIG-I, (PDB 5E3H)[28] a helix is followed by a loop turning around RecA1B (Protrusion I) and a second helical motif which extends into a C-terminal allosteric regulator that relays substrate binding to the ATPase core. The C-terminal continuation is depicted as a dashed red line.



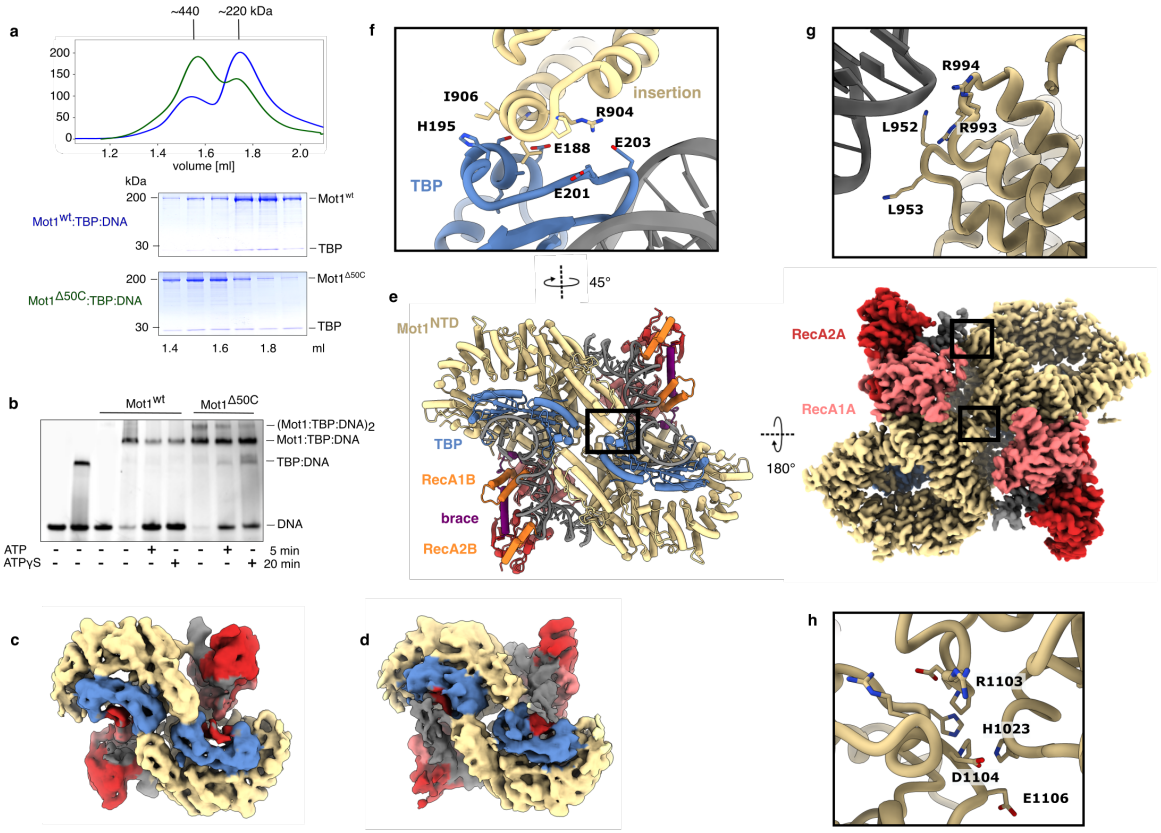
Extended Data Fig. 2: Crosslink-mass spectrometry (XL-MS) of Mot1:TBP:DNA and ATPase assays.

a) Mot1 complex topology as revealed by XLMS. Mot1-TBP inter-links are depicted in blue, Mot1 intra-links in green, with interlinks between the C-terminal bridge of Mot1 and TBP in orange. The legend (right) assigns a color-code for functional domains.
b) BS3 crosslinker titration of Mot1:TBP:DNA on a Coomassie-stained SDS-gradient gel. A protein:BS3 molar ratio of 0.3 was used for the final XLMS experiment.
c) Localization of the anchor domain confirmed by spatially allowed crosslinks (blue) between an anchor lysine (K1855) and TBP as well as a HEAT repeat lysine as analyzed by XLMS.
d) Structural model of the anchor domain built in its cryo-EM density in the posthydrolysis state.
e) Three Mot1 complex conformations with crosslinks visualized in the respective structures. Spatially allowed crosslinks (< 30 Å) are colored blue, crosslinks violating the allowed distance (> 30 Å) are colored red.
f) Histograms depicting the allowed and violated crosslinks from the structures in e).



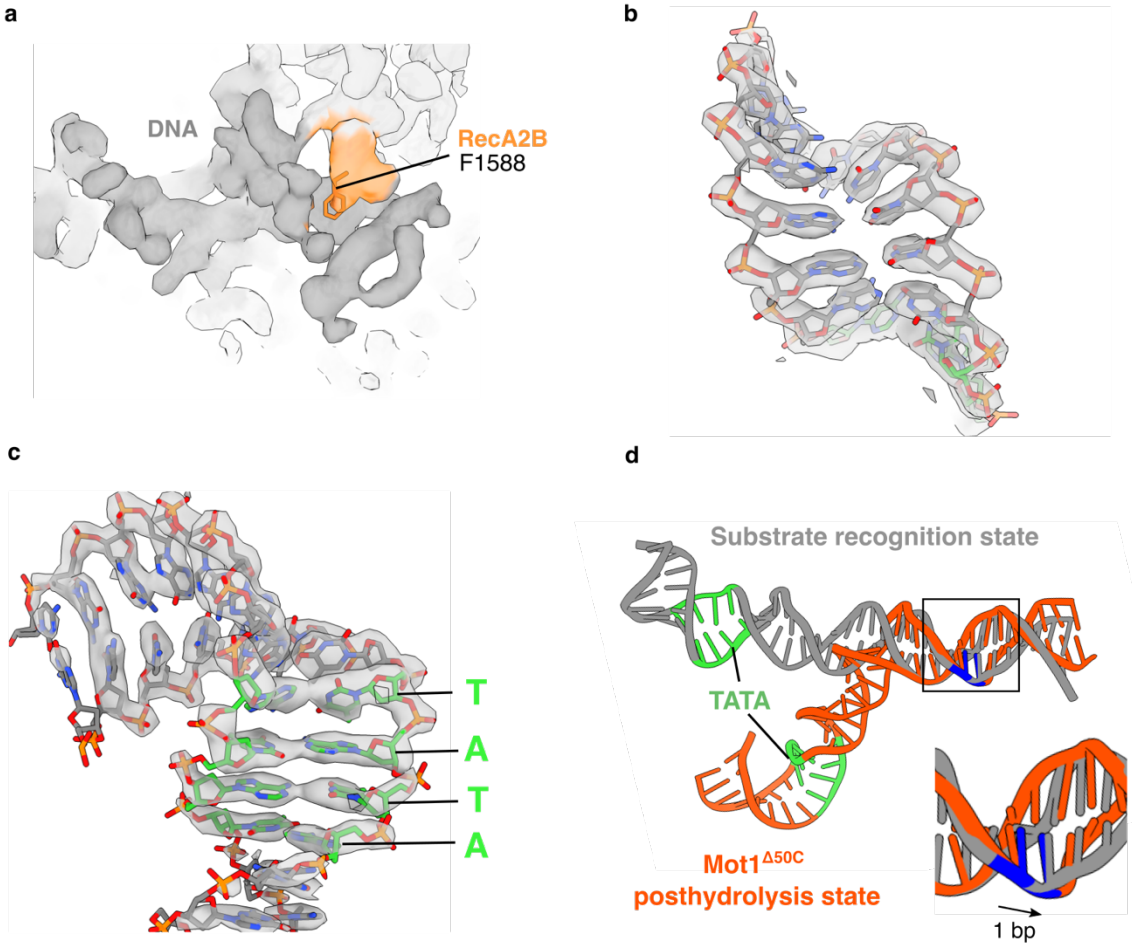
Extended Data Fig. 3: Enzymatic and structural consequences of bridge mutations.

a) Quantification of electrophoretic mobility shift assays as shown in Fig. 5 e. Horizontal bars represent means from three technical replicates shown as dots, error bars represent the standard deviation.
b) ATPase assay of Mot1 alone and in complex with TBP:DNA comprising the truncated constructs from a). Horizontal bars represent means from three technical replicates shown as dots, error bars represent the standard deviation.
c) Sequence alignment of Mot1 from three species comprising the conserved Nuclear Localization Sequence (NLS, degree of identity reflected by shades of blue).



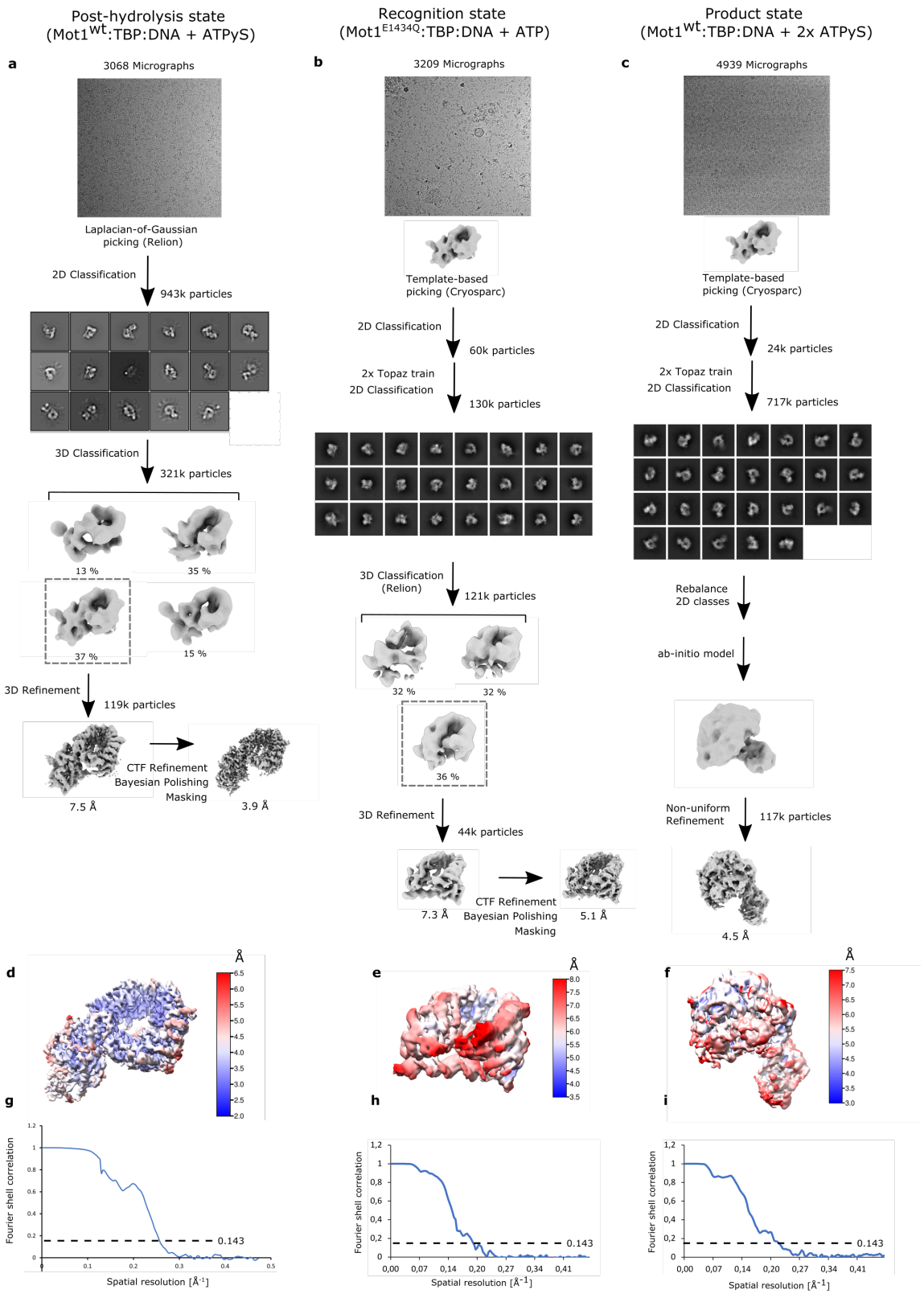
Extended Data Fig. 4: Dimerization of Mot1 complexes

a) Comparative analytical size-exclusion chromatography (SEC) of Mot1^{wt} (blue) and Mot1$^{\Delta 50C}$ (green) trimeric complexes (top) and corresponding samples after SDS-PAGE (bottom).
b) Native electrophoretic mobility shift assay (EMSA) of Mot1^{wt} and Mot1$^{\Delta 50C}$ trimeric complexes with 5'-6-FAM-labelled DNA on the reverse strand showing successive complex formation. A super shift indicates dimerization for the Mot1$^{\Delta 50C}$:TBP:DNA. Addition of ATP or ATPγS leads to complex dissociation and accumulation of free DNA.
c) Low resolved cryo-EM density of Mot1:TBP:DNA dimer in the prehydrolysis state.
d) Low resolved cryo-EM density of Mot1:TBP:DNA dimer in the posthydrolysis state.
e) Structure and denoised 2.8 Å cryo-EM map of Mot1$^{\Delta 50C}$:TBP:DNA complex dimer rotated by 180°. Black squares indicate the close-ups in d), e), and f).
f) Close-up of the interaction between TBP and the opposite insertion domain.
g) Close-up of the Mot1 N-terminus contacting the upstream DNA of the opposite complex via an arginine/lysine basic patch.
h) Close up of the interaction between the N-termini of the opposite monomers



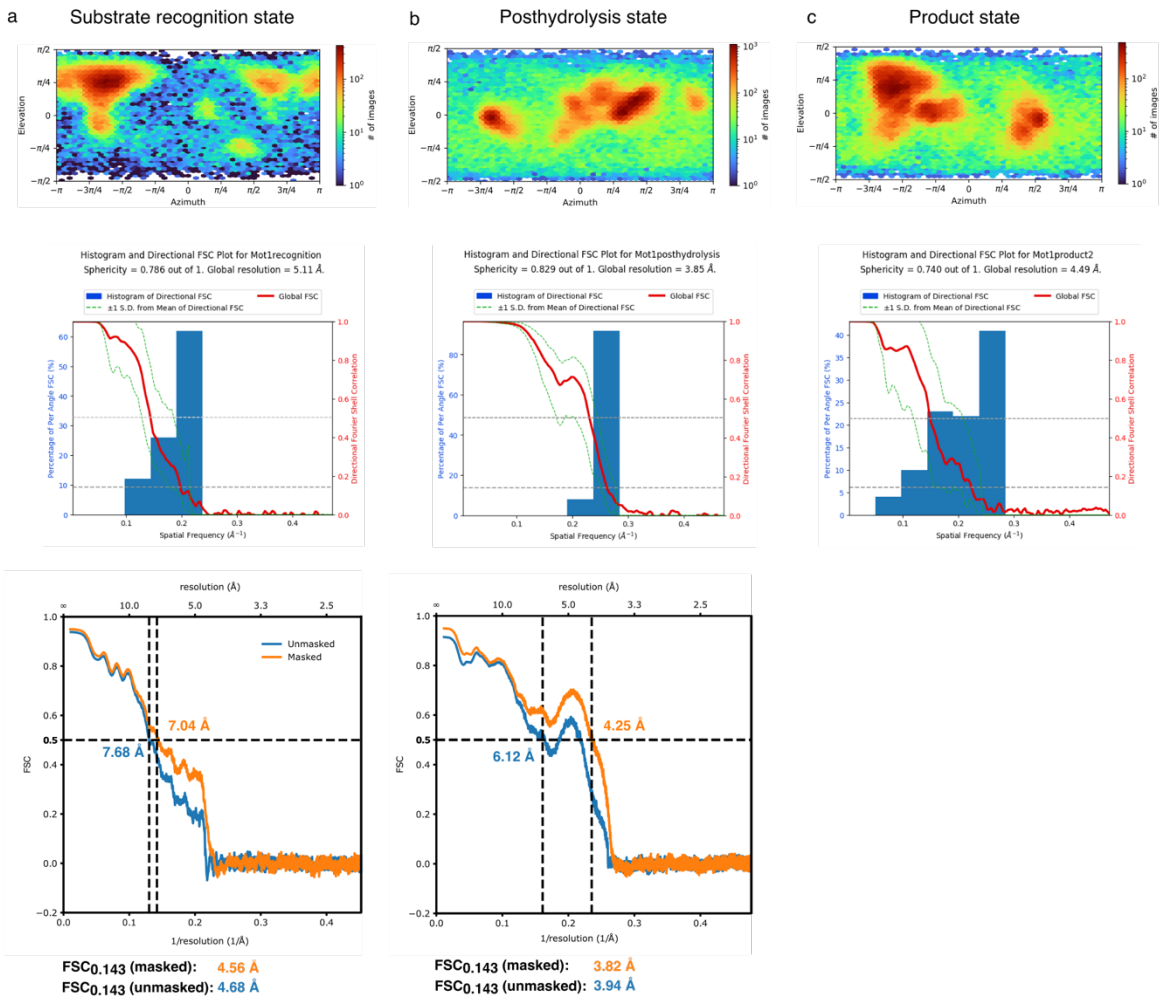
Extended Data Fig. 5: DNA bending and short-range translocation by Mot1^{CTD}

a) Close up of the cryo-EM density of the Mot1^{CTD} residing DNA showing intercalation of a RecA2B phenylalanine (F1588, orange) into the minor groove.
b) Example density and built in model of DNA widened at the concave site of TBP from the Mot1$^{\Delta 50C}$:TBP:DNA reconstruction.
c) Density and built in model of the DNA assigning the TATA-motif bound by TBP from the Mot1$^{\Delta 50C}$:TBP:DNA reconstruction.
d) Comparison of DNA paths between substrate recognition state (grey) and Mot1$^{\Delta 50C}$:TBP:DNA posthydrolysis state (orange) by alignment via RecA1 (not visible). The same upstream cytosine (blue) is shifted upstream by one base pair in the Mot1$^{\Delta 50C}$ complex, indicated in the close up from the black square.


Post-hydrolysis state
(Mot1wt:TBP:DNA + ATPyS)
Recognition state
(Mot1E1434Q:TBP:DNA + ATP)
Product state
(Mot1wt:TBP:DNA + 2x ATPyS)
a
3068 Micrographs
Laplacian-of-Gaussian
picking (Relion)
2D Classification
943k particles
3D Classification
321k particles
13 %
35 %
37 %
15 %
3D Refinement
119k particles
CTF Refinement
Bayesian Polishing
Masking
7.5 Å
3.9 Å
b
3209 Micrographs
Template-based
picking (Cryosparc)
2D Classification
60k particles
2x Topaz train
2D Classification
130k particles
3D Classification
(Relion)
121k particles
32 %
32 %
36 %
3D Refinement
44k particles
7.3 Å
CTF Refinement
Bayesian Polishing
Masking
5.1 Å
c
4939 Micrographs
Template-based
picking (Cryosparc)
2D Classification
24k particles
2x Topaz train
2D Classification
717k particles
Rebalance
2D classes
ab-initio model
Non-uniform
Refinement
117k particles
4.5 Å
d
e
f
Å
g
h
i
Fourier shell correlation
0.143
Spatial resolution [Å-1]

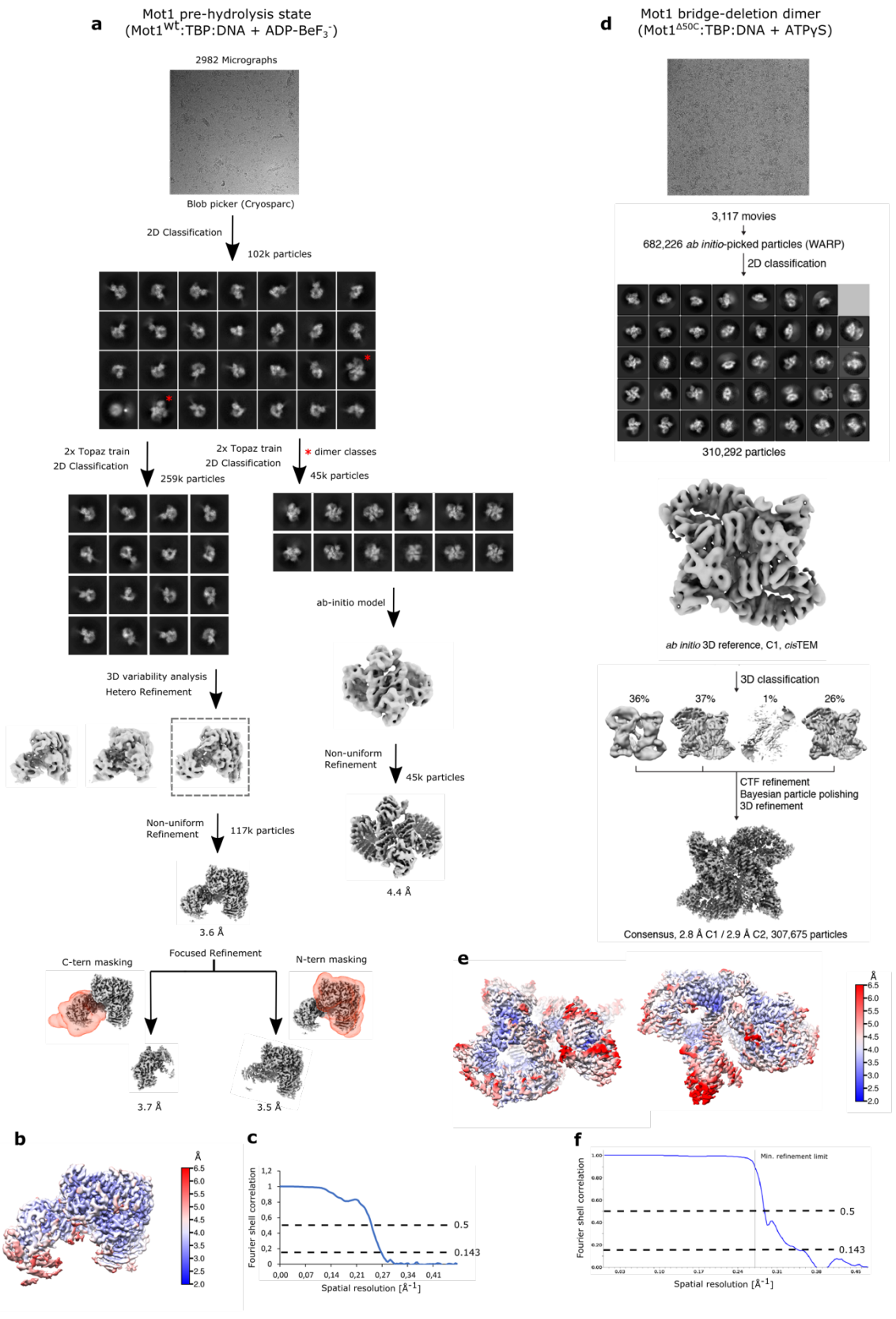
Extended Data Fig. 6: Cryo-EM processing schemes for Mot1 posthydrolysis, recognition and product states.

a) Processing of Mot1^{wt}:TBP:DNA with added ATPγS before SEC, representing the posthydrolysis state. In Relion, particles were picked with the Laplacian-of-Gaussian picker, sorted by repetitive cycles of 2D classifications and 3D classified. The most complete 3D map was 3D refined and submitted to CTF refinement and Bayesian polishing, yielding a reconstruction of 3.9 Å average resolution.
b) Processing of Mot1^{E1434Q}:TBP:DNA with added ATP before SEC, representing the recognition state. The first picking was done with the template picker in cryoSPARC, using the cryo-EM map of the posthydrolysis state. Several rounds of 2D classification yielded particle classes, which were used for training of the Topaz picker. Particles picked by Topaz were again sorted by 2D classifications followed by another round of Topaz training with subsequent 2D classifications. For 3D classification the particle coordinates were transferred to Relion. The most coherent map was 3D refined and submitted to CTF refinement and Bayesian polishing, yielding a reconstruction of 5.1 Å average resolution.
c) Processing of Mot1^{wt}:TBP:DNA with added ATPγS before SEC and before grid preparation, representing the product state. The template picker in cryoSPARC was used with the cryo-EM map from the posthydrolysis state as a template for the first round of picking. 2D classifications alternating with two times training of the Topaz picker respectively led to a set of particles that were used to calculate an ab-initio model, which was non-uniformly refined, yielding a reconstruction of 4.5 Å average resolution.
d)-f) Local resolution of the cryo-EM maps of the three respective Mot1 complex states from a)-c).
g)-i) Fourier shell correlation of the masked cryo-EM maps of the three respective Mot1 complex states from a)-c). The “gold standard” resolution cut-off (0.143) is marked by a dashed line.



Extended Data Fig. 7: Angular distribution (top), 3DFSC curves (middle) and eventual map-to-model FSCs (bottom) for Mot1

a) substrate recognition state
b) posthydrolysis state
c) product state


a
Mot1 pre-hydrolysis state
(Mot1wt:TBP:DNA + ADP-BeF3-)
2982 Micrographs
Blob picker (Cryosparc)
2D Classification
102k particles
2x Topaz train
2D Classification
259k particles
2x Topaz train
2D Classification
* dimer classes
45k particles
ab-initio model
3D variability analysis
Hetero Refinement
Non-uniform
Refinement
45k particles
Non-uniform
Refinement
117k particles
4.4 Å
3.6 Å
Focused Refinement
C-tern masking
N-tern masking
3.7 Å
3.5 Å
d
Mot1 bridge-deletion dimer
(Mot1Δ50C:TBP:DNA + ATPγS)
3,117 movies
682,226 ab initio-picked particles (WARP)
2D classification
310,292 particles
ab initio 3D reference, C1, cisTEM
3D classification
36%
37%
1%
26%
CTF refinement
Bayesian particle polishing
3D refinement
Consensus, 2.8 Å C1 / 2.9 Å C2, 307,675 particles
e
Å
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
b
Å
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
c
Fourier shell correlation
1,2
1
0,8
0,6
0,4
0,2
0
0.5
0.143
0,00
0,07
0,14
0,21
0,27
0,34
0,41
Spatial resolution [Å-1]
f
Fourier shell correlation
Min. refinement limit
0.5
0.143
Spatial resolution [Å-1]

Extended Data Fig. 8: Cryo-EM processing schemes for Mot1 posthydrolysis, recognition and product states.

a) Processing of Mot1^{wt}:TBP:DNA with added ADP-BeF_3^- before SEC, representing the prehydrolysis state. Particles were picked with blob picker in cryoSPARC and sorted via multiple rounds of 2D classification. Particles from good classes were used for two times Topaz training with intermittent rounds of 2D classification. Dimer classes (marked by red asterix) were Topaz trained and sorted separately. A dimer *ab initio* model (C1) was calculated and non-uniformly refined to an average resolution of 4.4 Å. Monomeric particles were submitted to 3D variability analysis followed by hetero-refinement of three distinct density maps. The map with the most distinct ATPase domain was non-uniformly refined to an average resolution of 3.6 Å. To increase quality of Mot1^{NTD} and Mot1^{CTD}, both areas were masked separately and focused-refined to average resolutions of 3.7 Å (C-terminal masked) and 3.5 Å (N-terminal mask) respectively.
b) Local resolution of the cryo-EM maps of Mot1 prehydrolysis complex from a)
c) Fourier shell correlation of the masked cryo-EM map of Mot1 prehydrolysis complex from a). The “gold standard” resolution cut-off (0.143) is marked by a dashed line.
d) Processing of the Mot1$^{\Delta 50C}$:TBP:DNA dimer. For details on processing see Methods section.
e) Local resolution of the cryo-EM maps of Mot1$^{\Delta 50C}$:TBP:DNA dimer from d)
f) Fourier shell correlation of the masked cryo-EM map of Mot1$^{\Delta 50C}$:TBP:DNA from d). The “gold standard” resolution cut-off (0.143) is marked by a dashed line.

Table S1: Cryo-EM data collection, refinement and validation statistics

	#1 Recognition (EMDB-14534) (PDB 7Z7N)	#2 Prehydrolysis (EMDB-14762) (PDB 7ZKE)	#3 Posthydrolysis (EMDB-14562) (PDB 7Z8S)	#4 Δ50C dimer (EMDB-14583) (PDB 7ZB4)	#5 Product (EMDB-14554)
Data collection and processing					
Magnification	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e–/Å^2)	48	45	48	55	44
Defocus range (μm)	-1.2 – -2.8	-1.2 – -2.8	-1.2 – -2.8	-1.2 – -3.2	-1.2 – -2.8
Pixel size (Å)	1.059	1.046	1.059	1.059	1.059
Symmetry imposed	C1	C1	C1	C1/C2	C1
Initial particle images (no.)	129,750	258,877	943,465	682,226	716,738
Final particle images (no.)	44,132	117,422	118,973	307,675	117,066
Map resolution (Å)					
0.143 FSC threshold	5.1	3.6	3.9	3.3	4.5
Map resolution range (Å)	4.2 – 15.2	3.1 – 36.9	3.0 – 48.2	2.8 – 14.0	3.6–58.7
Refinement					
Initial model used (PDB code)	6G7E, 1YTB	6G7E, 1YTB	pre-hydrolysis state	post-hydrolysis state	
Model resolution (Å)					
0.5 FSC threshold	6.9	3.7	4.4	3.4	
Model resolution range (Å)					
Map sharpening *B* factor (Å^2)	-272	-168	-156	-90	
Model composition					
Non-hydrogen atoms	13,144	14,760	14,798	28,550	
Protein residues	1,484	1,818	1,692	3,252	
Nucleotides	72	20	72	144	
Ligand	--	1	--	--	
B factors (Å^2)					
Protein	83.7	36.1	53.5	51.5	
Nucleotide	1.0	77.7	1.0	67.8	
Ligand		79.0			
R.m.s. deviations					
Bond lengths (Å)	0.004	0.005	0.004	0.004	
Bond angles (°)	0.636	0.652	0.745	0.657	
Validation					
MolProbity score	1.61	1.21	0.91	1.12	
Clashscore	6.12	3.41	1.54	2.39	
Poor rotamers (%)	0.24	0.71	1.03	0.64	
Ramachandran plot					
Favored (%)	96.06	97.61	98.03	97.54	
Allowed (%)	3.94	2.39	1.97	2.46	
Disallowed (%)	0.00	0.00	0.00	0.00	

5.3 Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Franziska Kunert, Felix J. Metzner, James Jung, Markus Höpfler, Stephan Woike, Kevin Schall, Dirk Kostrewa, Manuela Moldt, Jia-Xuan Chen, Susanne Bantele, Boris Pfander, Sebastian Eustermann, Karl-Peter Hopfner. Structural mechanism of extranucleosomal DNA readout by the INO80 complex, *Science Advances*, accepted in principle, Oct 2022.

Summary

In the publication, we extended previous structural work on the INO80 complex by employing an INO80 subcomplex comprising the core (C-module) with the motor ATPase and the allosterically regulating Arp8 module (A-module). We assembled INO80 from *Saccharomyces cerevisiae* and *Chaetomium thermophilum* on nucleosomes with an 80 bp entry DNA overhang. High-resolution cryo-EM structures of the A-modules from both species bound to extranucleosomal DNA led to the identification of a DNA-binding "AT-hook" at the Arp8 N-terminus and a conserved binding mode of the A-module's client protein Ies4. Ies4 anchors to Arp4/Actin via a β-hairpin that establishes a hydrophobic pocket by two angled tryptophan residues, denoted "2W motif". Via sequencing alignment and structural comparisons, we defined the 2W β-hairpin as a minimal variant of the abundant "WW domain".

The structural work is accompanied by NADH-coupled ATPase assays, EMSAs, gel-based nucleosome shift assays, anisotropy measurements and yeast growth assays of diverse INO80 DNA-binding mutants. DNA stretches with differing DNA stiffness/bendability proximal to the A-module and between ATPase and Arp5 subunit show a dependence of nucleosome positioning on DNA shape parameters. Structures comprising straight and bend DNA reflect the versatile DNA-binding of the A-module.

Separate expression of the human A-module with its client protein YY1 yielded a low-resolution cryo-EM density, where A-module subunits could be rigid-body docked into. The YY1 binding site was predicted with AlphaFold2 and fitted into the density as well.

Additionally, identifying a lysine-rich site at the Arp5 "grappler" domain contacting entry DNA allowed to propose an integrative model of how INO80 senses and monitors extranucleosomal DNA via the A-module, Arp5 and the ATPase and how conformational changes dependent on DNA shape are relayed between A-module and motor ATPase to achieve context-dependent nucleosome positioning.

Author contribution

I cloned and purified the human A-module and collected and processed the respective cryo-EM data. Further, I contributed to the analysis of the 2W β-hairpin domain and took part in preparing figures and writing of the manuscript.

Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Authors

Franziska Kunert1,†, Felix J. Metzner1,†, James Jung1,2,†, Markus Höpfler3,4,‡, Stephan Woike1,‡, Kevin Schall1,5, Dirk Kostrewa1, Manuela Moldt1, Jia-Xuan Chen6, Susanne Bantele3,7, Boris Pfander3,8,9, Sebastian Eustermann1,9,†, Karl-Peter Hopfner1*

Affiliations

1 Gene Center, Department of Biochemistry, Ludwig-Maximilians-Universität München, Munich, Germany

2 Present address: Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD, Australia.

3 DNA Replication & Genome Integrity, Max-Planck-Institute of Biochemistry, Martinsried, Germany

4 Present address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, United Kingdom

5 Present address: Crelux GmbH, Martinsried, Germany

6 Institute of Molecular Biology (IMB), Mainz, Germany

7 Present address: Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

8 Present address: Genome Maintenance Mechanisms in Health and Disease, Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany

9 Present address: Genome Maintenance Mechanisms in Health and Disease, Institute of Genome Stability in Ageing and Disease, CECAD Research Center, University of Cologne, Cologne, Germany

10 European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

† equal contribution

‡ equal contribution

* Correspondence: Prof. Dr. Karl-Peter Hopfner

Gene Center,

Feodor-Lynen-Str. 25,

81377 Munich, Germany

Electronic address: hopfner@genzentrum.lmu.de

Tel.: +49 (0) 89 2180 76953

Abstract

The nucleosomal landscape of chromatin depends on the concerted action of chromatin remodelers. The INO80 remodeler specifically places nucleosomes at the boundary of gene regulatory elements, which is proposed to be the result of an ATP dependent nucleosome sliding activity that is regulated by extranucleosomal DNA features. Here we use cryo-electron microscopy and functional assays to reveal how INO80 binds and is regulated by extranucleosomal DNA. Structures of the regulatory A-module bound to DNA clarify the mechanism of linker DNA binding. The A-module is connected to the motor unit via an HSA/post-HSA lever element to chemo-mechanically couple the motor and linker DNA sensing. Two notable sites of curved DNA recognition by coordinated action of the four actin/actin-related proteins and the motor suggest how sliding by INO80 can be regulated by extranucleosomal DNA features. Finally, the structures clarify the recruitment of YY1/Ies4 subunits and reveal deep architectural similarities between the regulatory modules of INO80 and SWI/SNF complexes.

Teaser

Cryo-EM and functional analyses reveal how the nuclear actin module senses linker DNA to regulate nucleosome remodeling by INO80

Keywords: chromatin remodeler, nucleosome, cryo-electron microscopy, INO80, nuclear actin

Introduction

Chromosomal DNA is predominantly organized in the form of nucleosome core particles (NCPs) – ~147 base pairs (bp) of DNA wrapped around the histone octamer (two copies of histones 2A, 2B, 3 and 4) – along with interspersed extranucleosomal linker DNA as well as larger nucleosome-free (NFRs) or nucleosome-depleted regions (NDRs) (*1*). NFRs and NDRs are important regulatory regions and are found at promoters, enhancers and origins of replication in *S. cerevisiae* (*2*). Nucleosomal packaging condenses and protects DNA, but also generates epigenetic information in the form of nucleosome occupation, histone modifications and histone variant composition (*2*).

The location, composition and epigenetic modifications of nucleosomes play key roles in the regulation of gene expression, DNA replication and DNA repair, and are shaped by the collective action of chromatin remodelers and epigenetic modifiers. Chromatin remodelers are molecular machines that use the energy of ATP hydrolysis to slide, position, evict or edit nucleosomes (*3, 4*). They are generally grouped into four main families: INO80/SWR1, SWI/SNF, ISWI and CHD. Common to all remodelers is a Swi2/Snf2-type ATPase domain that uses ATP-hydrolysis to translocate DNA. This basal activity is converted into the diverse remodeling reactions by additional, remodeler-specific domains or subunits (*5*).

INO80 is a >1 megadalton chromatin remodeler that is conserved from yeast to human (*6, 7*) and emerges as a central multi-subunit enzyme complex that determines chromatin structure around NDRs/NFRs (*8*). INO80 slides canonical nucleosomes and hexasomes (i.e. nucleosomes lacking one H2A-H2B dimer), forms regularly spaced nucleosomal arrays and exchanges histone variants *in vitro* (*9-11*). Hereby, INO80 shows a uniquely robust ability to position +1 (i.e. transcription start site) and -1 (opposite side) nucleosomes that generate the boundary to the nucleosome free DNA in NDRs/NFRs in genome wide *in vitro* chromatin reconstitution assays (*12*). *In vivo*, INO80 is implicated in NDR/NFR and array formation as well (*13, 14*).

A comprehensive mechanistic framework for the different biochemical activities of INO80 and how they are regulated or work together is still largely elusive. For instance, the detailed structural mechanism by which INO80 determines +1 and -1 nucleosome positions, remains unclear. NFR located barrier factors such as *S. cerevisiae* Reb1, DNA ends, but also NFR features such as promoter DNA mechanics and shape recently emerged as regulators of INO80-mediated nucleosome positioning in whole genome chromatin reconstitutions (*15-17*). In mammals, INO80 might be regulated in part by the DNA sequence since the DNA binding transcription factor YY1 (Yin Yang 1), an early developmental regulator and structuring factor of promoter-enhancer elements, is a component of the *Hs*INO80 complex. Altogether, current evidence suggests that INO80 acts as information processing hub which integrate diverse sources of information to properly shape chromatin around gene promoter regions (*15, 16*).

Structural studies on INO80 and other remodelers revealed basic principles how these molecular machines (or subcomplexes) bind nucleosomes and mobilize nucleosomal DNA

using cycles of ATP binding and hydrolysis (*18-25*). Even in light of this process, we are far from understanding how complex remodeling reactions are carried out in a highly regulated manner, owed in part to their complex, dynamic and modular architecture. INO80 contains more than 15 subunits, organized in three structural modules that we denote "N", "A" and "C". Up to now, structural information is available for the C-module bound to the nucleosome, as well as parts of the A-module in the absence of DNA. The Ino80 polypeptide itself carries the core ATPase motor activity and acts as a scaffold for the three modules. The C-module is the core nucleosome sliding unit: it contains the Swi2/Snf2 ATPase motor domain of Ino80p ($Ino80^{motor}$), the scaffolding AAA+ ATPases Rvb1 and Rvb2, and nucleosome binding subunits Ies2, Ies6 (Ies: Ino eighty subunit) and Arp5 (Arp: actin related protein) (*20, 21*). The NCP is bound by $Ino80^{motor}$-Ies2 at DNA superhelical location SHL-6 and by Arp5-Ies6 at DNA SHL -2. Furthermore, the Arp5 “grappler” insertion domain interacts with the nucleosome “acidic patch”, a motif at the H2A/H2B interface that is a binding site for numerous chromatin proteins (*20*). In this configuration, the $Ino80^{motor}$ pumps extranucleosomal entry DNA into the NCP, a model that can explain its sliding activity (*10, 20, 22*). The function of N- and A-modules is less clear. The N-module is evolutionary rather divergent, binds DNA and has autoregulatory functions to ensure switch-like activation of INO80 by extranucleosomal DNA (*26*). The A-module is highly conserved in evolution and contains an HSA (helicase-SANT-associated) domain ($Ino80^{HSA}$) in the middle of the Ino80p polypeptide chain, along with actin (Act1), Arp4, Arp8, Ies4 and Taf14. The complex of $Ino80^{HSA}$ with Arp4, actin and Arp8 has been crystallized and low resolution structural along with functional analysis suggests that the $Ino80^{HSA}$ domain acts as extranucleosomal DNA sensor, which is required for robust nucleosome sliding (*27, 28*) and positioning in whole genome chromatin reconstitution (*15, 16*).

It is yet unclear how the A-module binds DNA and how it regulates the C-module. A-modules are found in all multi-subunit remodelers of the INO80/SWR1 family, carry nuclear actin, and while their functional importance is well-established, the underlying regulatory and sensing mechanisms are unclear. Here we present cryo-EM structures of the regulatory INO80 A-module (*C. thermophilum, S. cerevisiae, H. sapiens*), the A-module bound to DNA (*C. thermophilum, S. cerevisiae*) and an overall structure of the INO80 A- and C-modules in an extranucleosomal DNA sensing configuration (*C. thermophilum*). Supported by yeast *in vivo* studies, the structures reveal the mode of extranucleosomal DNA binding and identify both $Ino80^{HSA}$ and Arp8 as core DNA binding elements. DNA can bind along the A-module in a notably curved fashion, which, together with biochemical analysis, supports a function as a DNA feature sensor. The overall structure of the A- and C-module-nucleosome complex, along with high-resolution views of the motor domain in nucleotide-free (*apo*) and ADP·BeF_x states suggest how extranucleosomal DNA sensing and DNA mechanical features might regulate INO80 through an allosteric link to the motor domain. Finally, we reveal that yeast/fungal Ies4 and human YY1 are structural homologs. A double tryptophan (2W)-anchored hairpin of Ies4/YY1 emerges as an evolutionarily conserved Arp4-actin anchor motif that unifies core A-module compositions across INO80 and SWI/SNF type remodelers and provides links to polycomb repressive complexes. Altogether, our data provide a structural framework for regulation of INO80 by extranucleosomal DNA.

Results

Architecture of the INO80 regulatory A-module

To determine the complete modular architecture of INO80 A-modules (Fig. 1A) and to gain insight into their interactions with DNA, we used cryo-EM to obtain high-resolution structures of A-modules from *C. thermophilum* and *S. cerevisiae* (Fig. 1B and C). Structures were obtained either directly from recombinantly produced A-modules, or as individually processed and refined A-module classes in cryo-EM datasets on various INO80 or INO80:nucleosome complexes (table S1). The qualities of the maps were good enough to model the polypeptide chain (Fig. 1B, fig. S1A) using prior crystal structures as starting models or de novo (Arp8 N-terminus, Ies4). The releases of AlphaFold2 (*29*) allowed us to interpret less well-defined regions of the maps, as well as interpret a medium resolution map of the *H. sapiens* A-module (Fig. 1D).

The A-modules from all three species revealed similar overall architectures and conformations (Fig. 1B to D). As observed in a previously reported partial crystal structure of *S. cerevisiae* Ino80^{HSA}-Arp4-actin-Arp8$^{\Delta N}$ (Arp8 N-terminus deletion), the cryo-EM structures showed a sequential arrangement of Arp4, actin and Arp8 along approximal 20 helical turns of the Ino80^{HSA} domain. However, the cryo-EM analysis enabled us to define two functionally important elements of the INO80 A-module that were missing in the previous crystallographic analysis, the N-terminal extension of Arp8 (i.e. amino-acids preceding the actin fold), and the Ies4 subunit (Fig. 1B and C).

We could visualize the majority of the *Ct*Arp8 N-terminal extension (residues 14-98) with only residues 1-13 missing. It forms an extended yet defined chain that folds along INO80^{HSA} towards the actin-fold part of Arp8, with additional contacts to Ies4, Arp4, and actin (Fig. 1B). *S. cerevisiae* Arp8's N-terminal extension (residues 1-266) harbors additional 170 amino acids, which are not visible in our structure and are not an evolutionary conserved feature. However, the conserved region of the Arp8 N-terminal extensions adopts a remarkably similar geometry and employ similar contacts along the actin folds and Ino80^{HSA}, despite the lack of secondary structures, suggesting a high degree of evolutionary and functional conservation (Fig. 1B and C).

The C-terminal part of Arp8 N-terminal extension forms a helix that binds into the interface of actin and Arp8 and thus might be affected by the nucleotide state of the actin folds. To test this, we imaged A-modules in the presence of different nucleotides (fig. S1A to C). In the order ADP>ATPγS->ATP we observe a very small conformational change in the *S. cerevisiae* Arp8-actin pair and an ordering of the N-terminal segment of Arp8 along actin and Arp4 in the ATP state (fig. S1D and E). This may indicate a potential differential role of ATP/ADP at *S. cerevisiae* Arp8. Typical for actin-fold proteins, the underlying conformational changes are very subtle, making it difficult to distinguish them from experimental variability in the cryo-EM analyses at this stage. In the case of *C. thermophilum* A-module, imaging without nucleotides (DNA bound classes) or adding ATPγS (without DNA) resulted in the presence of ATP/ATPγS at the nucleotide binding sites of all three actin fold proteins (fig. S2A and B). In any case, in both *S.*

cerevisiae and *C. thermophilum* A-modules, we observed constitutive ATP binding at Arp4 and actin, while nucleotide binding to *S. cerevisiae* Arp8 is at least variable (fig. S1A to C, fig. S2A and B).

The resolution of the maps allowed us to unambiguously define and model the central part of the Ies4 subunit and define its interaction within the A-module (Fig. 1B). *Ct*Ies4$^{173\text{-}192}$ forms a β-hairpin that binds across actin (subdomain I) and Arp4, stabilizing and fixing their mutual arrangement. The same β-hairpin structure and interaction architecture is seen in the case of the *S. cerevisiae* Ies4$^{35\text{-}74}$, despite low sequence similarity (Fig. 1C). Comparing both structures sheds light onto two tryptophan residues (2W), which emerge as key anchor points to actin and are highly conserved among Ies4 homologs (Fig. 1E to G). While the β-hairpin element (denoted 2W-hairpin) and some flanking parts are defined in the structures, further N- and C-terminal parts of Ies4 are not resolved.

The two tryptophans bind a Gly366-Pro367 linker between the last two helices of actin (subdomain I) (Fig. 1G, fig. S2C, fig. S3A). Here, Pro367 is situated in an aromatic “corner” formed by the nearly right-angled tryptophan side chains. Interestingly, a similar type of interaction to human β-actin P367 through two tryptophans organized in a β-stranded structure is seen in the extracellular actin sensor C-type lectin DNGR-1, suggesting a more widely evolved actin interaction principle (fig. S3B) (*30*). Furthermore, the β-stranded fold and the presence of two tryptophans is broadly similar to the WW domains that bind proline-rich peptides (fig. S3C) (*31, 32*).

Ies4 also interacts with the N-terminal tail of Arp8 as well as with Ino80^{HSA} (Fig. 1B, fig. S2D and E). These contacts are mediated by the tip of the β-hairpin element and are probably important to assemble a “defined” INO80 A-module, since the Arp4-actin pair is also present in BAF/PBAF, SWR1 and NuA4 complexes as part of different molecular assemblies. The 2W-hairpin motifs of Ies4 are furthermore interesting as they resemble the structure of Rtt102 bound to Arp7-Arp9. Arp7-Arp9 are the orthologs of Arp4-actin in *S. cerevisiae* SWI/SNF family remodelers SWI/SNF and RSC (Fig. 1E and F). Rtt102 displays a similar 2W attachment to Arp9 as Ies4 (Fig. 1E, fig. S3D), revealing an architectural conservation of A-modules across INO80 and SWI/SNF remodelers that goes well beyond the Arp4-actin pair and the HSA domain.

Mammalian YY1 is the structural homolog of yeast and fungal Ies4

Mammalian INO80 does not have a clearly recognizable Ies4 homolog, based on sequence conservation. However, YY1, a GLI/Krüppel-like transcription factor associated with chromosome loop formation, stem cell biology and early development, has been shown to interact with a module of human INO80 containing Ino80^{HSA}, Arp4 and Arp8 (*33*). To see whether YY1 could be the evolutionary ortholog of Ies4, we produced recombinant *H. sapiens* A-module *Hs*INO80^{HSA}, ACTL6A (Arp4 homolog), β-actin, ACTR8 (Arp8 homolog) and YY1.

These proteins assemble in a stoichiometric and stable complex that we used for cryo-EM analysis (fig. S4A). From 25,652 particles, we obtained a map with a resolution of 7.5 Å (Fig. 1D, fig. S4B), but a high degree of particle orientation bias impeded a higher resolution reconstruction. Still, it allowed unambiguous interpretation with models derived from the crystal structure of hArp8 and AlphaFold2 models of hArp4 and β-actin. In general, the arrangement of actin related proteins and β-actin along *Hs*INO80^{HSA} is very similar to that found in fungal and yeast complexes (Fig. 1B to D). Interestingly, after docking of the actin fold proteins, residual density at the hydrophobic rim of ACTL6A-β-actin matches very well the density corresponding to the hairpin region of Ies4 on the surface of yeast and fungal Arp4-actin (Fig. 1D). Sequence analysis (Fig. 1E) and AlphaFold2 prediction of YY1 indicated that residues 201-226 have the appropriate β-hairpin structure with two conserved, flanking tryptophans. This part has also been crystallized in a complex with the polycomb group protein MBTD1 and indeed shows a 2W-hairpin motif (fig. S3E) (*34*). AlphaFold2 modeling of a complex of ACTL6A and the 2W-hairpin of YY1 (Fig. 1H) resulted in a complex that matches the corresponding surface density of the *Hs*A-module. Chemical crosslinking and mass spectrometry (CX-MS) also identifies a crosslink, consistent with this location of YY1 (fig. S4C and D). Of note, binding of the 2W-hairpin motif (denoted also REPO domain) (*35*) to ACTL6A-β-actin is distinct from its interaction with MBTD1. Superposition of both complexes via the YY1 element indicates partially overlapping binding sites to the hairpin region (fig. S3F), which may explain partitioning of the Drosophila YY1 ortholog Pho into INO80 and *Drosophila melanogaster (Dm)* polycomb group protein Sfmbt (*35*).

Besides the INO80 complex, we identified the 2W-hairpin motif in the AlphaFold2 predictions of complex subunits of INO80 family (*S. cerevisiae*: Ies4 in INO80, Swc4 in SWR1 and NuA4; *H. sapiens*: YY1 in INO80, DMAP1 in SRCAP and TIP60) and SWI/SNF family remodelers (*S. cerevisiae*: Rtt102 in SWI/SNF and RSC; *H. sapiens*: BCL7 in BAF and PBAF) (*36*), hinting at a pervasive binding motif between remodeler families (fig. S3G to I).

Altogether, we conclude that mammalian YY1 is the ortholog of fungal and yeast Ies4 and that actin (or Arp9 in the case of *S. cerevisiae* SWI/SNF and RSC) along with Arp4 orthologs (or *Sc*Arp7) recruit a REPO/2W-hairpin element protein client (YY1, Ies4, Rtt102 and others) to assemble a conserved hetero-trimeric element in SWI/SNF and INO80/SWR1 chromatin modifying complexes (fig. S3J).

HSA$^{\alpha1}$ and HSA$^{\alpha2}$ are critical for INO80 function in yeast

Prior biochemical work established that the INO80 A-module is important for extranucleosomal DNA recognition and nucleosome sliding *in vitro* (*27, 28*). To this end, we previously identified a series of positively charged residues on HSA$^{\alpha1}$ and HSA$^{\alpha2}$ that, upon mutation to glutamines, severely affected the nucleosome sliding *in vitro* (denoted HSA^{Q1} and HSA^{Q2}) (*27*). We introduced these mutants, along with *arp8Δ, arp8ΔN* (*28*), and a Walker B mutation in Ino80 that affects ATP hydrolysis (*ino80^{E842A}*) into *S. cerevisiae* (W303 background) (table S2 and S3).

Since these mutants were designed prior to experimental DNA complex (see below), we generated an additional set of more structure-informed (*C. thermophilum*) K/R->A mutants in *S. cerevisiae* HSA$^{\alpha 2}$ (denoted HSA^{A2}), which led to similar effects as the HSA^{Q2}. We tested for viability under unchallenged conditions as well as in the presence of different stresses that had previously been linked to the INO80 function (*37*-*39*). While a WT *INO80* construct was able to complement the *INO80* deletion, *ino80-HSA^{Q1}* gave poor growth already at unchallenged conditions and was unable to support growth upon heat stress, in the absence of inositol, under anaerobic conditions or upon induction of a DSB (DNA double-strand break) (Fig. 2A, see fig. S5A to C for expression levels of mutant proteins). *ino80-HSA^{Q2}* and *HSA^{A2}* cells showed similar but slightly less severe phenotypes. The *ino80-HSA^{Q1+Q2}* double mutant was unable to support viability in W303 background, similar to strains lacking *INO80* or the *ino80^{E842A}* mutant (Fig. 2B), suggesting an additive contribution of DNA binding by HSA$^{\alpha 1}$ and HSA$^{\alpha 2}$. Furthermore, deletion of *ARP8* showed a growth phenotype under all stresses, but was only mildly affecting growth under non-perturbed conditions (Fig. 2C). Interestingly, expression of *arp8ΔN* partially rescued the *arp8Δ* heat stress phenotype, but not the homologous recombination dependent DSB repair function as tested in growth and ectopic recombination assays (Fig. 2C and D). Altogether, these data validate the importance of putative DNA interacting residues of the HSA domain in rendering INO80 functional and indicate that the INO80 DNA binding surfaces might affect the diverse functional roles to different degrees.

Structural basis of DNA binding by the INO80 A-module

Having established the structure of INO80 A-module and the critical functional role of the positively charged Ino80^{HSA} surface residues *in vivo*, we set out to reveal the way the A-module interacts with extranucleosomal DNA. We utilized a subset of 2D classes in our *Ct*INO80 dataset (ADP·AlF$_x$ and *apo*) that showed well-defined A-module:DNA complexes (Fig. 3A to C). Using extensive 2D and 3D classification, 3D variability analysis (movie S1) and refinement, we classified and refined two states that differ somewhat in the way they bind DNA (Fig. 3D). One state was refined to 3.3 Å resolution and showed ~25 bp linear DNA. In a second state, refined to 3.4 Å resolution, additional protein DNA contacts result in binding of ~35 bp DNA that exhibits curved conformation.

In both states, the majority of DNA interactions are formed by the Ino80^{HSA}, consistent with the robust effects of Ino80^{HSA} mutations in the *in vivo* analysis. Additional interactions are contributed by the N-terminal extension of Arp8 and by Ies4 (Fig. 3A and C). While most Ino80^{HSA}-mediated DNA interactions appear to be peripheral electrostatic interactions between Lys and Arg side chains and DNA, a central contact side is at HSA^{L1}, a loop which disrupts the Ino80^{HSA} element into two helices HSA$^{\alpha 1}$ and HSA$^{\alpha 2}$ (Fig. 3B). Here, the N-terminal turn of HSA$^{\alpha 2}$ binds a DNA backbone phosphate through main chain amide nitrogens. This interaction shows a remarkable similarity to the DNA interactions of the innate immune sensor cGAS and could provide a phosphate “registry-lock” (*40*). The central contact side is reinforced by a “hook” element of the Arp8 N-terminus that binds to the DNA backbone as well as to two minor groove

base pairs (Fig. 3C). The hook element is stabilized by Ies4 (residues 148-156), which is also in direct binding distance to the DNA backbone and may contribute further interactions. Similar folds of the hook-region in the *apo* states of the *C. thermophilum* and *S. cerevisiae* A-modules suggest evolutionary conservation of this DNA binding element.

In the case of curved DNA, we also observe DNA contacts around SHL-11, mediated predominantly via the $HSA^{\alpha1}$ region and a helix near the very N-terminus of $Arp8^{N}$ (Fig. 3D). Binding of curved DNA is noteworthy as it might be influenced by DNA mechanical properties. Geometrically, it is a result of the curved shape of the $Ino80^{HSA}$ at the Arp4-actin pair, which is incompatible with binding of linear DNA along the entire length of the A-module (Fig. 3E and F).

A 7.5 Å resolution structure of *S. cerevisiae* A-module bound to DNA could also be reconstructed from 69226 particles (fig. S1F). Interestingly, here we see predominantly contacts at $HSA^{\alpha1}$ and the Arp8 N-terminal extension. Again, DNA appears to be curved at this side, but the rather low resolution prevents a more detailed analysis. Predominant binding of DNA at $HSA^{\alpha1}$ is consistent with the somewhat stronger growth defects of HSA^{Q1} mutations in *S. cerevisiae in vivo* (Fig. 2A).

In summary, we provide a structural mechanism for extranucleosomal DNA binding of the INO80 A-module, revealing multiple DNA contact sites along the entire A-module and the possibility to interact with both curved and linear DNA through a modular set of interaction sites.

Biochemical analysis

The observation that *Ct*INO80 A-module can bind both linear and curved DNA prompted us to perform more detailed biochemical studies to analyze the role of different DNA binding sites (Fig. 4A to F) on *Ct*$INO80^{\Delta N}$ remodeling (Fig. 4G, fig. S6A and B), *Ct*$INO80^{\Delta N}$ ATP hydrolysis (Fig. 4H, fig. S6C) and the A-modules DNA binding activities (Fig. 4I, fig. S6D) *in vitro*. To this end, we evaluated various structure-derived mutations in the $Ino80^{HSA}$ and the Arp8 subunit (fig. S6B, E and F).

Mutations in $HSA^{\alpha1}$ (R740A, K741A, K745A, R748A, K763A, K770A, R774A), $HSA^{\alpha2}$ (K781A, K784A, K791A, R792A, R795A, K802A, R803A, R806A), or truncating of the Arp8 N-terminal extension ($Arp8^{\Delta N}$) did not significantly influence the ATPase rate of *Ct*$INO80^{\Delta N}$, but reduced ($HSA^{\alpha1}$, $Arp8^{\Delta N}$) or nearly abolished ($HSA^{\alpha2}$) nucleosome sliding. They also reduced the A-module DNA binding efficiency (Fig. 4G to I). This suggests that DNA contacts of the HSA domain add proper grip or induce a particular geometry to couple ATP hydrolysis cycles with nucleosome sliding. The severe effect of the $HSA^{\alpha2}$ mutant in sliding, but moderate effect in DNA binding, argues for a geometric function at least for this region, but does not rule out a function as grip as well.

Arp8.1 (N34A, Q35A, K36A, N37A, Y38A, K44A), carrying mutations in the hook as well as the N-terminal helix, leads to a reduction in sliding in the same range as $Arp8^{\Delta N}$, but this effect

appears to be caused by defects other than a simple reduction of DNA affinity (Fig. 4G and I). Again, this argues for a defective geometry of the active complex or a particular conformational state. The most remarkable effect showed the Ino80.1 mutant (K721A, K725A, R736A, R740A), which carries mutations in the very distal extranucleosomal DNA binding region of the Ino80^{HSA}. Ino80.1 strongly reduces binding to the DNA, but in fact increases sliding and ATPase rate of *Ct*INO80$^{\Delta N}$ (Fig. 4G to I). These effects could be explained if the A-module can also negatively regulate INO80 and that such a role is affected by the Ino80.1 mutation.

In summary, the mutations all affect various functions and the biochemical properties of INO80, validating our structural results. However, they indicate that the A-module plays a more complex, pleiotropic regulatory role with activating and inhibitory roles on remodeling.

Overall structure of INO80 A- and C-modules bound to a nucleosome

The complex regulation of nucleosome sliding by the INO80 A-module suggests an intricate regulatory coupling between A- and C-modules (Fig. 5A and B). To reveal how the A-module could chemo-mechanically communicate with the C-module, we recorded and analyzed various datasets of *Ct*INO80$^{\Delta N}$ bound to 0N80 nucleosomes in the absence and presence of the ATP analogue ADP·BeF_x. We used masking, particle subtraction and focused refinement procedures to obtain well-resolved maps at various regions of the complex. Aided by AlphaFold2 modeling of structural elements, we could substantially improve our prior analysis (*20, 27*) and add previously missing parts such as the architecture of the grappler and the post-HSA domain bound to the Ino80^{motor}.

We first focused on the motor domain to see what effects ATP binding has on the way the Ino80^{motor} interacts with DNA. In the *apo* state (no nucleotide), the Ino80^{motor} is well resolved and substantially bends DNA as previously described (Fig. 6A). In the *apo* state, we now see clear density for the post-HSA domain, which was missing in previous analyses. It interacts as a continuous, long helix with the N-lobe of the motor domain. It occupies the same region on the motor as the regulatory elements auto-N of ISWI and the post-HSA domain of Snf2 (*41, 42*) showing a high degree of conservation of motor regulatory elements among different remodelers (fig. S7A). However, we note that the interactions of post-HSA domains of Ino80 and Swi2 is somewhat shifted, although other parts of the motor superimpose and match very well. It was previously suggested that movements of the post-HSA could be coupled to motor activation (*42, 43*).

ADP·BeF_x binding leads to a straightening of the DNA at the motor compared to the bent conformation with widened minor groove in the apo state (Fig. 6A and B, fig. S7B and C and movie S2). Furthermore, in the ADP·BeF_x bound state, the post-HSA domain is not visible anymore, suggesting it is a rather dynamic feature that could be either coupled to the nucleotide state of the motor and/or the relative location of the A-module with respect to the motor (see below) (Fig. 5A).

Comparing the nucleotide-free with the ADP·BeF_x bound state, we observe a conformational transition in the Ino80^{motor} that is very similar to what has been described for Snf2 and ISWI bound to the nucleosomes (*41, 42*). Upon transitioning from ADP·BeF_x to the *apo* conformation, a step that could resemble ATP hydrolysis and ADP + P_i release, extranucleosomal DNA is rotated and pulled towards the nucleosome, consistent with one part of a translocation step.

In any case, the dynamics of the post-HSA motor contacts point towards a more profound allosteric communication between A- and C-modules on both ends of the HSA/post-HSA domain and we set out to obtain an overall structure to see how A- and C-modules could communicate. While A- and C-modules appear to be generally mobile with respect to each other in most picked particles, we could identify a subset of particles in the dataset without nucleotide addition that showed a more defined orientation between the A- and C-modules. This set of particles resulted in a 7.7 Å map that allowed us to place high-resolution structures of A- and C-modules and model the entire HSA/post-HSA helix that links A- and C-modules (Fig. 5A and B). In this structure, the A-module is situated at SHL-9 to -11, orientated such that Arp8 faces the Ino80^{motor} domain, while Arp4 points away. The HSA/post-HSA region forms a continuous helix all the way from Arp8 to the N-lobe of the motor domain. In this state, the Ino80^{HSA} domain could even employ further DNA contacts between Ino80^{motor} and Arp8, supported by several Lys/Arg side chains in the vicinity of DNA. Due to the structural flexibility and moderate resolution of this state, the DNA grooves are not well defined. However, modeling canonical B-DNA into the map indicates that the A-module is not exactly bound to DNA in the same way as we observe in the high-resolution individual reconstructions. It appears to be shifted along DNA by approximal ½ helical turn, suggesting that the A-module might not be fully engaged and aligned with the DNA grooves in this state as observed on the majority of DNA bound classes of the A-module alone (Fig. 3A to C). The observed configuration could be a nucleosome “sliding” state, where loose DNA binding of the A-module does not slow down nucleosome sliding, yet promotes post-HSA-motor contacts. Such an interpretation would be consistent with the observation that some mutants in the HSA/A-module actually lead to a speeding up of nucleosome sliding, while others slow down or abolish sliding. In the ADP·BeF_x dataset, we were not able to identify an equivalent subset of particles with well-defined arrangements of A- and C-modules. The absence of the post-HSA domain density may suggest a more dynamic mutual orientation of A- and C-modules. From sterically considerations, the different angle of entry DNA at the motor in the ADP·BeF_x state may not allow mutual binding of the A-module to DNA along with a linear HSA/post-HSA domain between A-module and motor. It is therefore possible that post-HSA-motor interactions are quite dynamic in the ATPase cycle or additional conformations of HSA/post-HSA and their attachment to the motor exist. Indeed, recent functional and structural studies on the RSC complex (*25*) identified an intriguing structural arrangement of the post-HSA domain at Protrusion I of the motor. Considering the conserved arrangement of these regulatory domains in Swi2/Snf2 ATPases, it is likely that an equivalent regulatory hub exists in INO80. Notably, the conserved QTELY motif, a homologue of the conserved SWI/SNF QTXX[F/Y] motif forms the post-HSA contact site towards Protrusion I, hinting at a critical interface for modular allostery by the A-module (Fig. 6C).

Taken together, the mode of interaction between A- and C-modules through HSA/post-HSA, and its modulation by nucleotide binding at the Ino80^{motor} provides an obvious direct chemo-mechanical link between the Ino80^{motor} and binding of the A-module to extranucleosomal DNA.

The Arp5 grappler interacts with entry DNA and regulates the motor domain.

The improved maps and AlphaFold2 structure predictions allowed us to model the complete Arp5 protein, in particular its unique "grappler" insertion element (Fig. 7A). This led to clarification of the way the grappler "foot" binds the acidic patch of the nucleosome and allowed us to identify two additional critical DNA contacts (fig. S8A). As described previously (*20*), we observe two remarkably distinct grappler configurations (fig. S8B and C). In the "parallel" state, its two main helical arms are arranged in a near parallel fashion and bind DNA around the nucleosome dyad. In the "cross" configuration, one helical arm binds along the DNA gyre, placing its tip at the entry DNA opposite the motor domain. Using 3D variability analysis (movie S3), formation of contacts between the tip of the cross arm appear to coincide with a better-defined HSA/post-HSA and a properly curved entry DNA, suggesting a functional link. We noticed two patches of Arg/Lys residues in loop regions that are properly placed to interact with the entry DNA and may account for this effect. Although the density map is not good enough to directly visualize these loops, the supporting helical elements are nevertheless defined well enough to confidently provide a location for the positively charged loops using AlphaFold2 models (Fig. 7B and C).

We generated two sets of point mutations in these Arp5 loop regions, Arp5.1 (R527A, K528A, R529A, R530A and R531A) and Arp5.2 (K362A, K363A, K366A, R367A) (Fig. 7B and C), and analyze their effects on nucleosome remodeling (Fig. 7D, fig. S8D) and ATP hydrolysis (Fig. 7E, fig. S8E). Both sets of Arp5 mutations nearly abolished nucleosome sliding activity and led to a dramatically reduced ATPase rate of *Ct*INO80$^{\Delta N}$. This might indicate a functional interplay between the Ino80^{motor} and the "grappler" on opposing sides of the entry DNA, enabling a geometry necessary for proper activation of the Ino80^{motor}, or by stabilizing the "unwrapped" (from H3/H4) geometry of entry DNA.

Regulation of remodeling by DNA features

Both the path of DNA around the Ino80^{motor}/Arp5 and the A-module show curved DNA regions, which are geometrically linked with the relative placement of A- and C-modules and a linear HSA/post-HSA helix, or the binding of extranucleosomal DNA along the entire Ino80^{HSA} domain. Prior experimental and statistical analysis indicated that the *S. cerevisiae* INO80 remodeling activity is influenced by DNA shape/mechanical features in extranucleosomal DNA. To test the generality of these observations for the *C. thermophilum* complex and also clarify the

contribution of different modules of INO80 to DNA feature readout, we replaced the sequence of our model substrate with an A/T rich, rigid sequence cassette derived from the URA3 promoter at four different locations (Fig. 8A and B, fig. S9A), probing contributions of distal (SHL -10/11) and proximal (SHL -8/9) extranucleosomal DNA binding sites of the A-module, the motor domain (SHL -6/7), as well as inside the nucleosome behind the Ino80^{motor} (SHL -4/5).

Interestingly, rigid DNA at SHL -8/9 and SHL -10/11 and inside the nucleosome (SHL -4/5) had a nearly equal, moderately reducing effect on nucleosome sliding by the *Ct*INO80$^{\Delta N}$ complex whereas changing the DNA at the Ino80^{motor} binding site more dramatically reduced sliding efficiency (Fig. 8C, fig. S9B). The sliding efficiencies did not correlate well with alterations in the ATPase rates of *Ct*INO80$^{\Delta N}$, since only the SHL -4/5 insertion had a reduced ATPase rate while all nucleosome variants showed similar binding efficiency (Fig. 8D, fig. S9C and D). ATP hydrolysis by the Ino80^{motor} and sliding efficiency were also not correlated in the analysis of A-module mutations, hinting towards futile ATP cycles when stiff DNA is located at the Ino80^{motor} or in extranucleosomal DNA. When DNA is inserted into the nucleosome it is plausible that the underlying nucleosome is structurally weakened, leading to proficient sliding despite reduced ATPase rates. It should be noted that in this analysis, DNA elements are obviously pushed along different DNA binding sites during the remodeling reaction and thus the effects might be to some extent integrated. Nevertheless, the strongest effect is observed at the motor domain, which is also the site where DNA shows the most profound bend.

In summary, these data show that inserting DNA cassettes with a stiff DNA sequence leads to a general reduction of nucleosome sliding, corroborating not only the influence of extranucleosomal DNA sequence, but in revealing that in particular the motor domain is sensitive to DNA features as well.

Discussion

In the past years groundbreaking structures of different remodelers bound to the nucleosome shed light on the basic principles of nucleosome recognition (*18-25*), while structural and functional analyses of selected single subunit remodelers (*4, 19*) suggested paths of allosteric activation by core nucleosome binding. While we begin to understand from these studies how remodelers grip and move DNA at nucleosomes, revealing the large-scale nucleosome reconfiguration steps and their regulation at atomic detail is the next frontier. For instance, the INO80 complex shows pleiotropic biochemical activities such as nucleosome spacing and editing, as well as the positioning of nucleosomes at NFR flanking regions (*12*). These diverse reactions depend on a basic nucleosome or hexasome sliding/mobilization activity (*11*), where the ATPase motor pumps extranucleosomal entry DNA into the nucleosome. In order to place a nucleosome at the +1 position, rather than sliding it further into the NFR, however, requires a regulation of the ATPase activity itself, or the coupling between Ino80^{motor} and nucleosome sliding. Regulatory signals could arise when the remodeler encounters a neighboring nucleosome, a barrier factor at the NFR/NDR and, at least in the case of *S. cerevisiae*, DNA

with particular mechanical or shape features such as those found in NFRs/NDRs (*15-17*). Functional and previous structural work suggested that a key regulative principle is the sensing of extranucleosomal DNA by the INO80 A-module (*15, 16, 27, 28*)

Here we provide a structural basis for this regulation and reveal how INO80 interacts simultaneously with nucleosomal and extranucleosomal DNA (Fig. 9). This work extends the analysis of multi-subunit remodelers from NCP binding to recognition of linker DNA and reveals how binding of extranucleosmal DNA by the A-module is chemo-mechanically coupled to the remodeling motor. We provide details of the linker DNA binding and identify multiple sites where DNA shape features might tune the biochemical activity. Hereby, in particular Arp8 and Arp5 subunits emerge as critical regulators, but also the Ino80^{motor} domain itself might play a central role in DNA feature readout. These properties of the motor domain could provide an explanation for the unique way INO80 has evolved to interact with the nucleosome, compared to other remodelers (*5*).

A first important outcome of our analysis is the extension of the A-module architectures from the previously recognized Arp4-actin-Ino80^{HSA} core element to include highly diverse REPO/2W-hairpin containing client proteins. The similarity of the INO80 Ies4/YY1 subunits with the RSC Rtt102 subunit (*44*) with respect to the 2W-hairpin, and the apparent exceptional conservation of the 2W-motif identifies a conserved nuclear actin anchor that is evolutionary conserved among INO80/SWR1 and SWI/SNF family remodeler (Fig. 1E). Although still lacking structural evidence, the human BAF complex subunit BCL7A is predicted by AlphaFold2 to harbor a similar 2W-motif (fig. S3I) further unifying A-module compositions across INO80 and SWI/SNF type remodeler. Interestingly, the 2W-hairpin bears similarity to the abundant and structurally well characterized WW domains (*45*), which binds proline-rich regions of their target proteins (*31, 32*). The classic WW domain is predominantly found in protein complexes involved in cell signaling, most prominently in the Hippo pathway (*46*). The two-stranded β-sheet in Ies4 and YY1 comprises two conserved tryptophans, but they assemble on opposing sites on the respective β-strand and form a hydrophobic pocket that accommodates one proline of an actin helix-turn motif. This minimal WW domain, which we denote 2W-hairpin, was described in Rtt102, where it tethers it to RSC A-module constituents Arp7/Arp9 (*25, 44*).

The 2W-hairpin also provides an interesting, unanticipated connection between DNA sequence feature recognition and remodeler regulation. Mammalian YY1 and Drosophila Pho both have additional DNA binding domains, which are absent in e.g. Ies4 and Rtt102 orthologs/paralogs. Ies4 is linked to roles of INO80 in the DNA damage response (*47, 48*) and in targeting to centromeric chromatin (*49*), but does not have recognizable DNA binding domains such as YY1 and Pho. We speculate that while the Arp4-actin-Ino80^{HSA} module serves as a core regulator of remodelers, the REPO/2W-hairpin clients provide a rapidly evolvable, variable adaptor to add remodeler-specific and species-specific regulatory and/or targeting features to the core A-module.

In YY1 the 2W-hairpin was characterized as the REPO (REcruitment of POlycomb) domain, because it facilitates recruitment of polycomb group complexes (PcGs) (*34*). This mirrors the

bivalent nature of YY1's context-dependent transcriptional activation and repression: YY1 either recruits activating (INO80) or repressive (PcG) complexes to their respective genomic loci dependent on the cellular context. As an integral subunit of the *Hs*INO80 complex, YY1 has been implicated in the recruitment of INO80 to promoter sites. A co-activation between YY1 and INO80 was proposed, because as a transcription factor acting on accessible DNA, YY1 also relies on the INO80 nucleosome sliding activity to gain access to its cognate promoter sites (*33*). Intriguingly, epigenetic regulation of YY1 binding to DNA could also influence INO80 recruitment. YY1 binding is inhibited by methylation of certain CpG sites (*50, 51*), which might conceivably control INO80 engagement, or activity, at promoter sites. Recently, a YY1-dependent recruitment not only of INO80, but also the BAF complex was shown in embryonic stem cells (*52*). Since BAF and INO80 share the BAF53^{Arp4}/β-actin pair, our structural results offer a possible mechanistic explanation, although whether YY1 binds BAF's A-module like it binds the INO80 A- module needs further investigation.

The way YY1/Pho interacts with the INO80 A-module or polycomb associated factors (*34*) suggests that it cannot bind two complexes at the same time, which explains the partitioning and also the different roles in e.g. cell survival (*53*). To this end, however, our structures might be useful to design point mutations that selectively perturb the YY1 interactions with either INO80 or polycomb complexes, thus helping to functionally dissect its different roles *in vivo*.

Prior functional evidence identified the Ino80^{HSA} domain as well as the Arp8 N-terminus as critical for extranucleosomal DNA sensing (*27, 28*). Here we provide a structural basis for this activity, showing how the A-module specifically recognizes DNA. Interestingly, both the Arp8 N- terminal extension as well as the Ino80^{HSA} domain directly bind DNA, altogether spanning three helical turns. We observed binding of curved DNA, which is of interest in the context of distinguishing nucleosome depleted promoter DNA elements from DNA in nucleosome bound gene bodies. Since the A-module binds to the concave side of the curved DNA, like the histones in the nucleosome, it might help distinguish nucleosome-receptive DNA in gene bodies from more rigid DNA in nucleosome depleted regions. We also find that the hook element of Arp8 widens the minor groove upon DNA binding and could contribute to DNA feature readout.

In addition to the A-module, the Arp5 subunit emerges as a critical regulator of the remodeling reaction, thus identifying all actin related proteins of the INO80 as core regulators. We consistently observe two major configurations in the helical insertion domain of Arp5 denoted "grappler" (*20*), which might point to rather complex functional roles in sliding or exchange reactions. While understanding the function of the "parallel" state and the precise role of the conformational switch needs to be addressed in future work, the "cross" state of the Arp5 grappler binds to the entry DNA opposing the Ino80^{motor} domain. This interaction appears to stabilize the path of entry DNA, allowing a continuous HSA helix to chemo-mechanically couple the extranucleosomal DNA bound A-module to the N-lobe of the motor domain. In order to enable this configuration, DNA needs to be bent at or ahead of the Ino80^{motor}, which provides a possible additional DNA feature sensing mechanism. In fact, placing rigid DNA at this region severely affects sliding efficiency, thus the Ino80^{motor} directly, or indirectly through the Ino80^{HSA}-

A-module geometry could be responsive to DNA mechanical properties. For instance, the extended Ino80^{HSA} domain and the A-module could act like a lever arm in this regard. Such a scenario might also help rationalize the peculiar NCP recognition mode of INO80 complexes as opposed to other remodelers. In INO80, the motor is placed at SHL-6 on the entry DNA, while in others, the motor is placed within the nucleosome at SHL 2. When the motor is positioned at SHL-6 it is able to not only pump DNA into the nucleosome but can in fact monitor DNA features at the same time. In contrast, a motor at SHL 2 might be more blind to shape features since the histones pre-bend DNA anyway.

A-module and Ino80^{motor} are chemo-mechanically coupled to the HSA/post-HSA helix. The allosteric regulation of Swi2/Snf2 motor domains by helical regulatory elements at the N-lobe is well founded (*54, 55*). In structural studies, these elements are often not visible, and might be rather transient, or show large conformational variabilities (*25*). In our structures, we observed that switching the Ino80^{motor} from *apo* to nucleotide bound states affects the interaction of the post-HSA with the N-lobe, a feature that could be intimately linked to remodeling. Since mutations in the Ino80^{HSA} severely reduce remodeling without substantially affecting the ATPase rate, it is plausible that the post-HSA might not switch the motor on or off, but rather provides a critical functional connection in a remodeling step. It could couple the motor activity to productive, directional DNA translocation, and reduce futile ATP hydrolysis steps without DNA translocation. If this is the case, it is unlikely that the A-module is simply a floating lever arm on extranucleosomal DNA, but could undergo positional changes to help translocate DNA.

In summary, we provide a detailed mechanism for extranucleosomal DNA binding by the Ino80^{HSA} and A-module and reveal how it is chemo-mechanically coupled to the motor of the C-module. The overall architecture reveals multiple instances of extranucleosomal curved DNA, indicating an integrative monitoring of DNA features (propeller twist, (*15*)) as one way to tune INO80 sliding (Fig. 9). Future studies need to address how the INO80 complex interacts with other substrates such as hexasomes and nucleosomal arrays. This will allow us to gain further insights into the conformational spectrum of the complex, the way INO80 possess “ruler” functions in the generation of nucleosomal arrays (*16*), and possibly understand the suggested histone exchange activities as well (*10*). To this end, it will be important to visualize the evolutionary highly variable N-terminal modules, which may add additional targeting but also negative regulatory activities (*26*). Nevertheless, our analysis provides an important step forward in the mechanistic understanding of these complex and fascinating chromatin shaping molecular machines.

Materials and Methods

Expression and purification of the INO80 complex from *C. thermophilum*

Subunits of the *Ct*INO80 complex and mutants were cloned and expressed by using the MultiBac technology. Gene coding for Ino80$^{718\text{-}1848}$ (*Ct*Ino80$^{\Delta N}$) -2 x FLAG was cloned in

pACEBac1, genes for Rvb1 and Rvb2 in pIDC and Arp5, Ies6 and Ies2 in pIDK vector. All together they were combined in one bacmid. Ino80$^{545-850}$ (INO80 A-module) was also cloned in pACEBac1. Ino80$^{1-850}$ (INO80ΔC) and Ies1 were also cloned in pACEBac1. Genes coding for HMG and Iec3 were cloned in pIDC. Genes coding for ZnF and FHA were cloned in the pIDS vector. Genes coding for Ies4, Taf14 were cloned in pACEBac1, Arp8, actin and Arp4 were cloned in a pIDK vector and combined on a separate bacmid. *PirHC* (Geneva Biotech) and *Escherichia coli* XL1-Blue (Stratagene) cells were used for all recombination steps by the addition of the Cre recombinase (NEB). From each bacmid (generated in *Escherichia coli* DH10 multibac cells), baculoviruses were generated in *Spodoptera frugiperda* (Sf21) insect cells (Thermo Fisher, #11497013). 1/100 of each baculovirus was transferred to 1 L *Trichoplusia ni* High Five cells (Invitrogen, #B85502) and thereby co-infecting the cells. Cells were cultured for 60 h at 27°C and harvested by centrifugation at 4°C.

Cells were disrupted in lysis buffer (30 mM HEPES, pH 7.8, 400 mM NaCl, 10% glycerol, 0.25 mM DTT, 0.28 µg/mL Pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL Benzamidine, 2 mM $MgCl_2$) for complex purification and gently sonified for two min (Duty cycle 50 % and Output control 5). Raw lysate was cleared by centrifugation at 20500 g and 4°C for 30 min. Supernatant was incubated with 2 mL Anti-FLAG M2 affinity gel (Sigma-Aldrich) for 1 h and washed with 50 mL lysis buffer and 75 mL wash buffer (30 mM HEPES, pH 7.8, 150 mM NaCl, 5 % glycerol, 0.25 mM DTT). The protein was eluted from the matrix by incubation with 4.5 mL wash buffer (supplemented with 0.2 mg/mL FLAG peptide) in three incubation steps of 20 min each.

The elution fractions were loaded onto a Mono Q 5/50 GL column (GE Healthcare) and eluted by an increasing salt gradient (200 mM NaCl to 1 M NaCl), resulting in highly pure INO80.

The Ino80$^{\Delta N}$ and A-module mutants were generated by site directed mutagenesis and expressed and purified as described above (table S4).

Expression and purification of the INO80 complex from *S. cerevisiae*

The coding sequences of the INO80 subunits were cloned into pFBDM vectors. One vector contained the C-terminally 2xFLAG tagged Ino80 coding sequence Ino80$^{1-598}$ (INO80ΔC) and a second vector the remaining subunits of the A- and N-module (actin, Arp4, Arp8, Taf14, Ies4, Ies1, Ies3, Ies5 and Nhp10) (table S4). Bacmids were generated using *Escherichia coli* DH10 multibac cells. Baculoviruses were generated in *Spodoptera frugiperda* (Sf21) insect cells (Thermo Fisher, #11497013). *Trichoplusia ni* High Five cells (Invitrogen, #B85502) were co-infected with two viruses (1:100 v/v) and cultured for 60 h at 27°C. The cells were harvested by centrifugation at 4°C.

For purification of the INO80 complexes, cells were resuspended in lysis buffer (50 mM Tris pH 7.9, 500 mM NaCl, 10% glycerol, 0.25 mM DTT, 0.28 µg/mL Pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL Benzamidine) and disrupted by sonication (4 x 1 min, duty cycle 50 % and

output control 5). The lysate was cleared by centrifugation at 20500 g and 4°C for 40 min. Supernatant was incubated with 3 mL Anti-FLAG M2 affinity gel (Sigma-Aldrich) for 1 h and washed with 50 mL wash 1 buffer (25 mM HEPES pH 8, 500 mM KCl, 10 % glycerol, 0.05% IGEPAL CA630, 4 mM $MgCl_2$ and 0.25 mM DTT), 50 mL wash 2 buffer (25 mM HEPES pH 8, 200 mM KCl, 10 % glycerol, 0.05% IGEPAL CA630, 4 mM $MgCl_2$ and 0.25 mM DTT) and 10 mL buffer A (25 mM HEPES pH 8, 150 mM KCl, 2 mM $MgCl_2$ and 0.25 mM DTT). The protein was eluted from the matrix by incubation with 4.5 mL buffer A (supplemented with 0.2 mg/mL FLAG peptide) in four incubation steps of 15 min each.

The elution fractions were loaded onto a Mono Q 5/50 GL column (GE Healthcare) and eluted by a linear salt gradient (150 mM KCl to 1 M KCl), resulting in highly pure INO80. The A-module was generated by site directed mutagenesis and expressed and purified as described above (table S4).

Expression, purification and grid preparation of *Hs*A-module + YY1

Human A-module (ACTR8, ACTB, ACTL6A and YY1) ORFs were ordered and optimized for insect cell expression at GeneArt (Thermo Fisher Scientific) and assembled on a single pBIG1ab vector using the BiGBac cloning system. The 2x Flag-tagged $Ino80^{HSA}$ ($Ino80^{267-487}$-2xFlag) was cloned separately on a pBIG1a vector. After virus generation in Sf21 cells (*Spodoptera frugiperda,* Thermo Fisher, #11497013) the complexes were recombinantly expressed in High Five insect cells (*Trichoplusia ni,* Invitrogen, #B85502) by adding the two viruses at 1:150 (volume virus:medium) to 3 L insect cell culture. The cells were incubated for 60 h at 27 °C and harvested by centrifugation at 4 °C. For lysis, the pellet was resuspended in lysis buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 0.25 mM DTT, 1x Protease inhibitor (complete, Roche)) and gently sonicated 3x 1.5 min. The lysate was incubated with Anti-FLAG M2 affinity gel (Sigma-Aldrich) for 1.5 h and submitted to a gravity flow column. First, the agarose beads were washed with 10 CV lysis buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 0.25 mM DTT) followed by 20 CV wash buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.25 mM DTT). Next, the protein complex was eluted three times by incubation with 1 CV wash buffer supplemented with 0.4 mg/mL 1xFlag peptide for 15 min each. The elution fractions were applied onto a CaptoHighRes Q 5/50 column (Cytiva) and the protein complex was separated via a salt gradient (100 mM NaCl to 1000 mM NaCl) using buffer A (20 mM HEPES pH 8.0, 100 mM NaCl, 20 mM $ZnCl_2$, 4 mM $MgCl_2$, 0.25 mM DTT) and buffer B (20 mM HEPES pH 8.0, 100 mM NaCl, 20 mM $ZnCl_2$, 4 mM $MgCl_2$, 0.25 mM DTT). Protein target peak fractions were concentrated to 2 mg/mL in centrifugal filters (Centricon 70 kDa cut-off, Millipore) and flash-frozen in liquid nitrogen.

For cryo-EM analysis, the purified A-module + YY1 was vitrified on glow-discharged R2/1 copper mesh 200 grids (Quantifoil). β-octyl glucoside (Roth, Germany) was added at a final concentration of 0.05%. 4.5 µl of sample was pre-incubated on the grid for 20 sec before blotting.

Purification of nucleosomes

Canonical human histones (HistoneSource) were resuspended in an unfolding buffer (7 M guanidinium chloride, 20 mM Tris pH 7.5, 1 mM DTT) respectively under rotation for 30 min at room temperature. Histones were mixed in 1.1-fold excess of H2A and H2B and dialyzed against 4x 1 L refolding buffer (20 mM Tris pH 7.5, 2 M NaCl, 1 mM DTT, 0.5 mM EDTA pH 8) for 16 h at 4 °C**.** Histone octamers were purified by size exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare). After concentrating to 4 mg/mL in centrifugal filters (Centricon 10 kDa cut-off, Millipore) histone octamers were stored in 50% glycerol at -20°C.

Widom 601 DNA with 80 bp extranucleosomal DNA in the 0N80 orientation for reconstituting nucleosomes was used as a DNA template (table S5). DNA was amplified by PCR, followed by purification using anion exchange chromatography and the DNA was concentrated in vacuum after the DNA was dialyzed to H_20 overnight. DNA was mixed at a 1.1-fold excess with the histone octamer at 2 M NaCl. The NaCl concentration was decreased to 50 mM over 16 h at 4°C. After this, nucleosomes were purified by anion exchange chromatography using a SourceQ 1 mL column, fractions containing nucleosomes were pooled and dialyzed to 50 mM NaCl, concentrated to 1 mg/mL (Centricon 10 kDa cut-off, Millipore) and stored at 4°C.

Nucleosome sliding assays

0N80 nucleosomes with 5`-Fluorescein-labeled extranucleosomal DNA were used for monitoring the sliding activity of *Ct*INO80$^{\Delta N}$ and mutants. 150 nM nucleosome was incubated with 50 nM *Ct*INO80$^{\Delta N}$ in sliding buffer (25 mM HEPES, pH 8, 60 mM KCl, 7% glycerol, 0.10 mg/mL BSA, 0.25 mM DTT, 2 mM $MgCl_2$) at 25°C. By the addition of 1 mM ATP the sliding reaction was started and stopped at several time points (15, 30, 45, 60, 120, 300, 600, 1200 sec) by addition of 0.2 mg/mL Lambda DNA (NEB). Nucleosome species were separated by native PAGE on a 3-12% acrylamide BIS-Tris gel (Invitrogen) and visualized using the Typhoon imaging system (GE healthcare). Experiments were performed in triplicates. For gel band quantification ImageJ was used and the fraction of remodeled band was plotted against the reaction time in percent. Data describes a saturation curve and was fitted in Prism (GraphPad) using an exponential equation.

NADH-coupled ATPase assay

NADH-coupled ATPase assays were used to determine the ATPase rate of *Ct*INO80$^{\Delta N}$ and mutants. 30 nM *Ct*INO80$^{\Delta N}$ was incubated in assay buffer (25 mM HEPES, pH8, 50 mM KCl, 1 mM DTT, 2 mM $MgCl_2$, 0.1 mg/mL BSA) with 0.5 mM phosphoenolpyruvate, 1 mM ATP, 0.1 mM NADH and 25 U/mL lactate dehydrogenase and pyruvate kinase (Sigma-Aldrich) at 25°C in a final volume of 50 µL. Decreasing NADH concentration was monitored fluorescently

over 1 h in nonbinding, black, 384-well plates (Greiner BioOne) using 340 nm for excitation and an emission of 460 nm with a Tecan Infinite M100 (Tecan). Where indicated, ATPase rate was determined in the presence of 200 nM nucleosome. Experiments were performed in triplicates. ATP turnover was calculated using maximal initial linear rates, corrected for a buffer blank.

Affinity measurement by fluorescence anisotropy

Increasing protein concentrations of the *Ct*INO80 A-module and mutants (final concentrations for *C. thermophilum* A-module in [nM]: 0, 3.125, 6.25, 12.5 25, 50, 100, 200) were prepared in assay buffer (25 mM HEPES pH 8, 100 mM KCl, 2 mM $MgCl_2$, 2% glycerol, 0.01% Triton X-100, 1 mM DTT) and mixed with 50 bp 6-FAM labeled DNA (table S5) in assay buffer (final concentration 5 nM) in a 1:1 (v/v) ratio (final volume: 20 µL, Greiner Flat Bottom Black 384 well plate). The reaction was incubated for 30 min at room temperature and the fluorescence anisotropy was subsequently measured at an excitation wavelength of 470 nm and an emission wavelength of 520 nm using a TECAN Infinite M1000 plate reader. Experiments were performed in triplicates. The background signal (no protein sample) was subtracted from each value of a dilution series and the data sets were analyzed with Prism (GraphPad Software). Data were analyzed and fitted to a non-linear, non-cooperative 1:1 binding model ($y = Af - (Af - Ab) \times (x/(Kd + x))$; *y* anisotropy; Af anisotropy of free ligand; Ab Anisotropy of bound ligand; Kd dissociation constant; *x* receptor concentration) to calculate the apparent dissociation constants.

Electrophoretic mobility shift assay

Electrophoretic-mobility shift assays were used to monitor the interaction between INO80 and 0N80 nucleosomes. Nucleosomes were labeled at the 5′-end of their extranucleosomal DNA with fluorescein. Nucleosome (40 nM) was incubated with 80 nM INO80 in electrophoretic mobility shift assay buffer (25 mM HEPES, pH 8, 60 mM KCl, 7% glycerol, 0.25 mM DTT, 2 mM $CaCl_2$) for 30 min on ice. Samples were analyzed at 4 °C by native PAGE on a 3–12% acrylamide BIS-Tris gel (Invitrogen) and visualized using the Typhoon imaging system (GE healthcare).

Purification and vitrification of *Ct*INO80$^{\Delta N}$:0N80 complex and INO80 A-module

*Ct*INO80$^{\Delta N}$ and 0N80 nucleosomes were mixed in a ratio of 2:1 and dialyzed to binding buffer (20 mM HEPES, pH 8, 60 mM KCl, 0.25 mM $CaCl_2$, 20 µM $ZnCl_2$, 0.25 mM DTT) for 1 h in slide-a-lyzer dialysis tubes (Thermo Fisher Scientific). The complex was vitrified at a concentration of 1 mg/mL on Quantifoil R2/1 grids in the presence of 0.05% octyl-β-glucoside using a Leica EM GP (Leica). The same was done for INO80ΔC and INO80A with or without DNA and

nucleotides. The *Ct*INO80A with ATPγS bound was purified further and mildly crosslinked by GraFix using an SW40-Ti rotor (Beckman Coulter). The glycerol (10-30%) and glutaraldehyde (0-0.025%) co-gradient was generated using Gradient Station *ip* 153 (BioComp Instruments) The samples were fractionated and monitored for 280/260nm absorbance using Triax UV Flow Cell (BioComp Instruments). The fractions were visually inspected and selected by uranyl acetate (2%) negative staining.

S. cerevisiae N-A-module and DNA (58bp) (table S5) were mixed in an equimolar ratio (1.5 µM each) in cryo-EM buffer (20 mM HEPES pH 8, 60 mM KCl, 2 mM $MgCl_2$, 1 mM DTT). The respective nucleotide was added (final concentration: 1 mM) and the sample was incubated on ice for 10 minutes. Octyl-β-glucoside was added (0.045%) and 4.5 µL was applied onto a glow discharged Quantifoil R2/1 Cu200 grid. The sample was vitrified in liquid ethane using an EM GP plunge freezer (Leica, 10 °C and 90% humidity).

Data collection

Movies of *Ct*INO80$^{\Delta N}$-nucleosome or A-module particles embedded in vitreous solution were collected at liquid nitrogen temperature using a Titan Krios G3 transmission electron microscope (ThermoFisher) equipped with a K2 Summit direct electron detector (Gatan) and a BioQuantum LS Imaging Filter (Gatan). The movies were recorded in counting mode using EPU acquisition software (ThermoFisher) at 130,000x magnification with a pixel size of 1.059 Å/pixel, and nominal defocus range of 1.1 – 2.9 µm. The total electron dosage of each movie was ~40-46 e/Å^2, fractionated into 40 movie frames with 250 ms/frame exposure time.

Cryo-EM data processing of *S. cerevisiae* A-modules

The movie frames were motion corrected using MotionCor2 (*56*). All subsequent cryo-EM data processing steps were carried out using cryoSPARC v3.3.1 (*57*) or relion-3.0 (*58*) and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143).

For the *S. cerevisiae* A-module bound to ATP (fig. S10A and C), the CTF parameters of the dataset (4543 micrographs) were determined using patch CTF estimation (multi). The exact processing scheme and data collection and refinement statistics are summarized in table S1. Initial particle picking was done on 2048 micrographs using Blob picker. The particles were subject to 2D classification and *ab initio* reconstruction and classes with clearly defined features were selected. The selected particles were used as input for a Topaz train job on 4543 micrographs. After three rounds of Topaz, 1 028 485 particles were extracted with a box size of 256px and a pixel size of 1.059 Å. The particles were subject to multiple rounds of 2D classification and heterogenous refinement. The class that showed the most defined features was selected (327 293 particles) and used for further refinement. The final resolution of the ATP

bound A-module reconstruction after non-uniform refinement was 3.3 Å. To identify a subset of DNA bound particles, particles were re-extracted in relion (fig. S10B) and subject to three rounds of 3D classification. The A-module bound to ATP and DNA was reconstructed from 69 226 particles and the final resolution after 3D refinement was 7.5 Å.

For the *S. cerevisiae* A-module bound to ADP (fig. S11A and B), the CTF parameters of the dataset (5550 micrographs) were determined using CTFFIND4.1. All subsequent cryo-EM data processing steps were carried out using relion-3.0 (*58*). Data collection and refinement statistics are summarized in table S1. 2 264 013 particles were picked using Autopicking and particles were extracted with a box size of 256px and a pixel size of 1.059 Å. 3D classification with five classes was performed using the A-module bound to ATPγS (filtered to 40 Å) as reference. After another round of 3D classification, 970 407 particles were selected and used for further refinement. The final resolution of the ADP bound A-module reconstruction after post processing was 3.2 Å.

Cryo-EM Data processing of *S. cerevisiae* A-module in the ATPγS-state

Beam-induced motions of particles were corrected using MotionCor2 (Relion-3.0) in 5 X 5 patches per frame (*56, 58*). Contrast transfer function (CTF) parameters were estimated from sums of three movie frames using CTFFIND4.1 (*59*). The particles were automatically picked *ab initio* and qualitatively filtered using WARP (fig. S12A and B) (*60*). The particles were boxed and extracted from the micrographs in Relion with the particle coordinates exported from WARP using the PyEM scripts developed by Daniel Asarnow (https://github.com/asarnow/pyem). The initial 3D reconstructions were carried out *ab initio* using *cis*TEM (fig. S12C). Iterative rounds of 3D classifications were carried out using Relion (*58*). The initial 3D refinements were carried out in Relion3 using the *ab initio* 3D reference generated in *cis*TEM (*61*). The final resolution of the ATPγS bound A-module reconstruction after post processing was 3.2 Å. The exact processing scheme is depicted in fig. S12A to D. Data collection and refinement statistics are summarized in table S1.

Cryo-EM Data processing of *C. thermophilum* A-modules and INO80$^{\Delta N}$

Beam-induced motions of particles were corrected using MotionCor2 (Relion-3.0) in 5 X 5 patches per frame (*56, 58*). Contrast transfer function (CTF) parameters were estimated from sums of three movie frames using CTFFIND4.1. The particles were automatically picked *ab initio* and qualitatively filtered using WARP (*60*). The particles were boxed and extracted from the micrographs in Relion with the particle coordinates exported from WARP using the PyEM scripts developed by Daniel Asarnow (https://github.com/asarnow/pyem). The initial 3D reconstructions were carried out *ab initio* using *cis*TEM. Iterative rounds of 3D classifications were carried out using Relion3 to remove unbound nucleosomes and separate subtly different C- and A-module conformations. The initial 3D refinements were carried out in Relion3 using the

ab initio 3D reference generated in *cis*TEM (*61*). The exact processing schemes are depicted in fig. S13 to S15. Data collection and refinement statistics are summarized in table S1.

Cryo-EM data processing of *C. thermophilum* INO80 C-module and nucleosome (ADP·BeF_x)

The movie frames were motion corrected using MotionCor2 (*56*). All subsequent processing steps were performed in cryoSPARC v3.2.0 (*57*) and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143). The CTF parameters of the datatset (6064 micrographs) were determined using patch CTF estimation (multi) in cryoSPARC (v3.2.0). The exact processing scheme is depicted in fig. S16A. Data collection and refinement statistics are summarized in table S1.

Initial particle picking was done using Blob picker. Particles were subject to 2D classification and *ab-inito* reconstruction. Classes with clearly defined features were selected and used as input for a Topaz train job on all micrographs, followed by particle extraction and 2D classification. After three rounds of Topaz, 304k particles were extracted with a box size of 360px and a pixel size of 1.059 Å. After selecting 2D classes with clearly defined features, one round of *ab initio* reconstruction with three classes was performed. Classes with most defined features were selected and subject to heterogeneous refinement with two classes. The *ab initio* reconstructions were used as input volumes for the heterogeneous refinement job. Both classes were selected for further refinement. The final resolution of the reconstruction after non-uniform refinement were 3.5 Å for parallel grappler and 3.8 Å for the cross grappler (fig. S16C).

For a detailed analysis of the $Ino80^{motor}$ beam-induced motions of particles were corrected using MotionCor2 (Relion-3.0) in 5 X 5 patches per frame (*56, 58*). Contrast transfer function (CTF) parameters were estimated from sums of three movie frames using CTFFIND4.1. The exact processing scheme is depicted in fig. S16B. 13704k particles were automatically picked *ab initio.* 1242248 manually picked particles were extracted with a box size of 360px and a pixel size of 1.059 Å. Iterative rounds of 3D classifications and 3D refinement were carried out. After the last round of 3D classification, 137900 particles were selected and used for further refinement. The final resolution of the $Ino80^{motor}$ after post processing was 3.6 Å (fig. S16D). Data collection and refinement statistics are summarized in table S1.

Cryo-EM data processing of *H. sapiens* INO80 A-module

The movie frames were motion corrected using MotionCor2 (*56*). All subsequent processing steps were performed in cryoSPARC v3.2.0 (*57*) and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143). The exact processing scheme is depicted in fig. S17A. Data collection and refinement statistics are summarized in table S1.

Initial particle picking was done using Blob picker. Particles were subject to 2D classification. Classes with clearly defined features were selected and used as input for a Topaz train job on all micrographs, followed by particle extraction and 2D classification. After three rounds of Topaz, 15k particles were extracted with a box size of 256px and a pixel size of 1.059 Å. After selecting 2D classes with clearly defined features, one round of *ab initio* reconstruction with one class was performed. The final resolution of the reconstruction after non-uniform refinement was 7.5 Å (fig. S17B).

Model Building & Refinement

A-modules for *S. cerevisiae* and for *C. thermophilum* were built with the crystal structure of the *S.* cerevisiae Arp8 module (*27*) as initial template. For each data set, the model was manually placed into the unsharpened cryo-EM map followed by rigid-body refinement with CHIMERAX (*62*). The model was then initially modified and corrected with COOT (*63*) against the sharpened cryo-EM map. Reciprocal space refinement using jelly-body restraints was done with SERVALCAT (*64*) against maximum-likelihood weighted structure factors calculated from cryo-EM half-maps. Further model building was done with COOT against the maximum-likelihood estimate of the expected true map calculated with SERVALCAT. Final model corrections were done with ISOLDE (*65*) against the same SERVALCAT map, followed by a final round of reciprocal space refinement using jelly-body restraints with SERVALCAT.

The structures of *C. thermophilum* INO80 C-module and *S. cerevisiae* A-module (PDB: 6FML and 5NBN) were docked into the cryo-EM densities using MOLREP (CCP-EM) (*66*), and manually mutated and built in previously unobserved regions using COOT (*63*). All protein models were real-space refined using PHENIX (*67*), and evaluated using COOT and the MolProbity server. The reconstruction cryo-EM maps were deposited in the Electron Microscopy Databank (EMDB) and the coordinates of the atomic models were deposited in the Protein Data Bank. The figures were generated ChimeraX (*62*).

Yeast manipulation and methods

All strains used (listed in table S3) were isogenic to W303 and were constructed via a diploid derivative of YCL076 (*39*). Briefly, knock-outs of *INO80* and *ARP8* were generated in a diploid strain using a PCR-based strategy and confirmed by PCR with locus-specific primers (*68, 69*). Mutant or WT alleles of either gene were cloned into the YIplac211 vector with endogenous promoter sequences and mutations as indicated, and were then integrated at the *URA3* locus. Single copy integration was tested by PCR. Diploid strains were subsequently sporulated and tetrads were dissected for tetrad analysis and to obtain haploid knock-out and point mutant strains for phenotypic analysis.

For growth assays, cells were grown overnight and adjusted to 0.5 OD600 units and five-fold serial dilutions were spotted on YPD plates (1% yeast extract, 2% peptone, 2% glucose), YP+Gal plates (2% galactose) or SD-inositol plates (6.9 g/l yeast nitrogen base without inositol, Formedium CYN37CFG, supplemented with 40 mg/l adenine, uracil, tryptophan and histidine, 80 mg/l leucine and 2% glucose). Cells were then grown at 30°C for 2–5 days unless indicated otherwise. For anaerobic growth conditions, plates were incubated in an anaerobic chamber.

Protein expression levels were determined by total protein extraction from a logarithmic culture using alkaline lysis followed by trichloroacetic acid precipitation as described (*68*). Proteins were separated by SDS-PAGE and analyzed by western blotting using anti-FLAG (Sigma, A8592) and Pgk1 (Invitrogen, #459250).

Recombination assay

To measure the efficiency of homologous recombination, a qPCR-based gene conversion assay was used as described previously (*39*). Briefly, yeast strains were deficient of endogenous HO endonuclease cleavage sites and engineered with galactose-inducible HO endonuclease, a single HO cut site at ChrIV 491 kb and a recombination donor sequence at ChrIV 795 kb with a mutated HO cut site and an addition unique 23 bp sequence to allow qPCR analysis. Yeast cells of the indicated genotypes were grown to logarithmic phase in YP+2% raffinose medium and HO endonuclease expression was induced by addition of 2% galactose. Aliquots equivalent to one OD600 unit were harvested at the indicated time points and genomic DNA was isolated using the Epicentre MasterPure Yeast DNA Purification Kit (MPY80200). qPCR was performed on a Light Cycler 480 instrument (Roche) using LightCycler 480 SYBR Green I Master (Roche 04707516001) with primers designed to detect the completed recombination product (5'-CATACTGTCTCACTCGCTTGGA-3' and 5'-TTGTTTGCCATTTCGTCAGCTAG-3'). Data were normalized to an unrelated control locus (*MDV1* locus, primers 5'-GCGTGCCTGGTCACAGGTTCATACGAC-3' and 5'-TCATACGGCCCAAATATTTACGTCCC-3') and plotted using the GraphPad Prism software as the relative amount of recombination product over time (where 100% recombination = 1). Of note, yeast growth on YP+Gal plates in spot dilution provided a qualitative readout for homologous recombination efficiency as well.

Protein crosslinking

Snap-frozen stock solutions of *H. sapiens* INO80 complex (20 mM HEPES/NaOH pH 8.0, 200 mM NaCl, 0.5 mM $CaCl_2$, 20 µM $ZnCl_2$, 0.5 mM DTT) and 0N80 nucleosome (20 mM HEPES/NaOH pH 7.5, 50 mM NaCl, 0.5 mM DTT, 10% glycerol) were thawed on ice and mixed in equimolar amounts in reconstitution buffer (20 mM HEPES/NaOH pH 7.9, 60 mM KCl, 0.5 mM $CaCl_2$, 20 µM $ZnCl_2$, freshly added 0.5 mM DTT). The INO80-nucleosome complex mixture was incubated on ice for 30 min to allow for reconstitution. Afterwards, 4 µl of freshly prepared BS3 crosslinker stock solution (Thermo Fisher Scientific; 2 µg/µl in reconstitution

buffer) was added to the reconstituted complex. The complex was crosslinked at 4°C for 2 h. After that, the reaction was quenched by adding 4 µl of 2 M ammonium bicarbonate, followed by incubation at 4°C for 30 min. Thereafter, half of the crosslinked product was processed by in-gel digestion and the other half by ethanol precipitation and in-solution digestion.

In-gel digestion

One half of the crosslinked product was mixed with LDS sample buffer and separated in a 4-12% NuPAGE Bis-Tris gel, stained with coomassie blue (Colloidal Blue Staining Kit; Thermo Fisher Scientific). The highlighted area of the gel (fig. S4C) was excised and cut into small gel cubes, followed by destaining in 50% ethanol/50 mM ammonium bicarbonate. The proteins were then reduced in 10 mM DTT at 56°C and alkylated by 50 mM iodoacetamide in the dark at room temperature. Afterwards, proteins were digested by trypsin (1 µg per sample) in 50 mM ammonium bicarbonate at 37°C overnight. Following peptide extraction sequentially using extraction buffer (0.1% formic acid in 30% acetonitrile) and 100% acetonitrile, the sample volume was reduced in a centrifugal evaporator to remove residual acetonitrile. The peptides were then acidified with 0.1% formic acid and purified by solid phase extraction in C18 StageTip (*70*).

Ethanol precipitation and in-solution digestion

The other half of the crosslinked product was mixed with 1 µL of 15 µg/µl GlycoBlue coprecipitant (Thermo Fisher Scientific), filled with reconstitution buffer to 100 µL and then transferred to a new 2-ml Eppendorf tube. The tube was then filled with pure ethanol to a final sample volume of 2 ml and incubated at 4°C overnight. Following centrifugation at 4°C for 1 h, the supernatant was aspirated, and the protein pellet was allowed to air-dry.

The protein pellet was resolubilised in 8 M urea/50 mM ammonium bicarbonate. The proteins were reduced in 5 mM DTT for 30 min and alkylated in 15 mM iodoacetamide for 25 min. Afterwards, an additional 5 mM of DTT was used to quench the iodoacetamide. The proteins were first digested by 0.5 µg of Lys-C for 3 h. After diluting the urea concentration to 2 M with 50 mM ammonium bicarbonate, 1 µg of trypsin was added to digest the proteins overnight. All procedures were carried out at room temperature of 22 °C. Following acidification to 0.5% trifluoroacetic acid, the resultant peptide solution was purified by solid phase extraction in C18 StageTip.

Liquid chromatography tandem mass spectrometry

Crosslinked peptides were analyzed using an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific) coupled to EASY-nLC 1200 UHPLC system (Thermo Fisher

Scientific). Peptides were separated in an in-house packed 55-cm analytical column (inner diameter: 75 µm; ReproSil-Pur 120 C18-AQ 1.9-µm silica particles, Dr. Maisch GmbH) by online reversed phase chromatography through a 90 min gradient of 2.4-33.6% acetonitrile with 0.1% formic acid at a nanoflow rate of 250 nl/min. The eluted peptides were sprayed directly by electrospray ionization into the mass spectrometer. Each sample was injected twice and measured using two different combinations of collision energies in stepped mode (*71*). Mass spectrometry measurement was conducted in data-dependent acquisition mode using a top15 method with one full scan (resolution: 60,000, scan range: 300-1650 m/z, target value: 3×10^6, maximum injection time: 40 ms) followed by 15 fragmentation scans via higher energy collision dissociation (HCD; normalised collision energy in stepped mode: 25, 30, 35 % or 27, 30, 33 %; resolution: 15 000, target value: 1×10^5, maximum injection time: 40 ms, isolation window: 1.4 m/z). Only precursor ions of +3 to +8 charge state were selected for fragmentation scans. Additionally, precursor ions already isolated for fragmentation were dynamically excluded for 25 sec.

Mass spectrometry data analysis

Raw data files were pre-processed by MaxQuant software package (version 1.6.5.0) (*72*) as described (*73*). The peak lists (*.HCD.FTMS.sil0.apl files) were searched using xiSEARCH (version 1.7.4) (*74*) against a target-decoy database consisting of the protein sequences of the *Hs*INO80 complex and nucleosome members. The following settings were used: enzyme specificity, trypsin; allowed maximum number of missed cleavages, 3; BS3 specificity linking K, S, T, Y and protein N-terminus; fixed modification, carbamidomethyl (C); variable modifications, oxidation (M) and mono-links for linear peptides on K, S, T, Y with dead-ends amidated or hydrolysed; MS1 tolerance, 6 ppm; MS2 tolerance, 20 ppm; boosting option activated for residue pairs; residue-level FDR was set at 5%.

References

1. G. C. Yuan *et al.*, Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**, 626-630 (2005).

2. W. K. M. Lai, B. F. Pugh, Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* **18**, 548-562 (2017).

3. C. R. Clapier, J. Iwasa, B. R. Cairns, C. L. Peterson, Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407-422 (2017).

4. I. M. Nodelman *et al.*, Nucleosome recognition and DNA distortion by the Chd1 remodeler in a nucleotide-free state. *Nat Struct Mol Biol* **29**, 121-129 (2022).

5. A. Jungblut, K. P. Hopfner, S. Eustermann, Megadalton chromatin remodelers: common principles for versatile functions. *Curr Opin Struct Biol* **64**, 134-144 (2020).

6. J. Jin *et al.*, A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *J Biol Chem* **280**, 41207-41212 (2005).

7. X. Shen, G. Mizuguchi, A. Hamiche, C. Wu, A chromatin remodelling complex involved in transcription and DNA processing. *Nature* **406**, 541-544 (2000).

8. K. Yen, V. Vinayachandran, B. F. Pugh, SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246-1256 (2013).

9. M. Udugama, A. Sabri, B. Bartholomew, The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol Cell Biol* **31**, 662-673 (2011).

10. S. Brahma *et al.*, INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nature Communications* **8**, 15616 (2017).

11. L. J. Hsieh *et al.*, A hexasome is the preferred substrate for the INO80 chromatin remodeling complex, allowing versatility of function. *Mol Cell*, (2022).

12. N. Krietenstein *et al.*, Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709-721 e712 (2016).

13. A. A. Reyes, R. D. Marcum, Y. He, Structure and Function of Chromatin Remodelers. *J Mol Biol* **433**, 166929 (2021).

14. Y. Yin *et al.*, A basal-level activity of ATR links replication fork surveillance and stress response. *Mol Cell* **81**, 4243-4257 e4246 (2021).

15. E. Oberbeckmann *et al.*, Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nat Commun* **12**, 3231 (2021).

16. E. Oberbeckmann *et al.*, Ruler elements in chromatin remodelers set nucleosome array spacing and phasing. *Nat Commun* **12**, 3232 (2021).

17. A. Basu *et al.*, Measuring DNA mechanics on the genome scale. *Nature* **589**, 462-467 (2021).

18. X. Liu, M. Li, X. Xia, X. Li, Z. Chen, Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440 (2017).

19. L. Farnung, S. M. Vos, C. Wigge, P. Cramer, Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* **550**, 539-542 (2017).

20. S. Eustermann *et al.*, Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature*, (2018).

21. R. J. Aramayo *et al.*, Cryo-EM structures of the human INO80 chromatin-remodeling complex. *Nat Struct Mol Biol* **25**, 37-44 (2018).

22. R. Ayala *et al.*, Structure and regulation of the human INO80–nucleosome complex. *Nature*, (2018).

23. A. B. Patel *et al.*, Architecture of the chromatin remodeler RSC and insights into its nucleosome engagement. *Elife* **8**, (2019).

24. S. He *et al.*, Structure of nucleosome-bound human BAF complex. *Science* **367**, 875-881 (2020).

25. R. W. Baker *et al.*, Structural insights into assembly and function of the RSC chromatin remodeling complex. *Nat Struct Mol Biol* **28**, 71-80 (2021).

26. C. Y. Zhou *et al.*, The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding. *Mol Cell* **69**, 677-688 e679 (2018).

27. K. R. Knoll *et al.*, The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nat Struct Mol Biol* **25**, 823-832 (2018).

28. S. Brahma, M. Ngubo, S. Paul, M. Udugama, B. Bartholomew, The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling. *Nat Commun* **9**, 3309 (2018).

29. J. Jumper *et al.*, Applying and improving AlphaFold at CASP14. *Proteins* **89**, 1711-1721 (2021).

30. P. Hanc *et al.*, Structure of the Complex of F-Actin and DNGR-1, a C-Type Lectin Receptor Involved in Dendritic Cell Cross-Presentation of Dead Cell-Associated Antigens. *Immunity* **42**, 839-849 (2015).

31. X. Huang *et al.*, Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat Struct Biol* **7**, 634-638 (2000).

32. M. A. Verdecia, M. E. Bowman, K. P. Lu, T. Hunter, J. P. Noel, Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* **7**, 639-643 (2000).

33. Y. Cai *et al.*, YY1 functions with INO80 to activate transcription. *Nat Struct Mol Biol* **14**, 872-874 (2007).

34. C. Alfieri *et al.*, Structural basis for targeting the chromatin repressor Sfmbt to Polycomb response elements. *Genes Dev* **27**, 2367-2379 (2013).

35. F. H. Wilkinson, K. Park, M. L. Atchison, Polycomb recruitment to DNA in vivo by the YY1 REPO domain. *Proc Natl Acad Sci U S A* **103**, 19296-19301 (2006).

36. C. Kadoch *et al.*, Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* **45**, 592-601 (2013).

37. J. Poli, S. M. Gasser, M. Papamichos-Chronakis, The INO80 remodeller in transcription, replication and repair. *Philos Trans R Soc Lond B Biol Sci* **372**, (2017).

38. W. Yao *et al.*, The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation. *Mol Cell Biol* **36**, 979-991 (2016).

39. C. A. Lademann, J. Renkawitz, B. Pfander, S. Jentsch, The INO80 Complex Removes H2A.Z to Promote Presynaptic Filament Formation during Homologous Recombination. *Cell Rep* **19**, 1294-1303 (2017).

40. F. Civril *et al.*, Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* **498**, 332-337 (2013).

41. L. Yan, H. Wu, X. Li, N. Gao, Z. Chen, Publisher Correction: Structures of the ISWI-nucleosome complex reveal a conserved mechanism of chromatin remodeling. *Nat Struct Mol Biol* **26**, 389 (2019).

42. M. Li *et al.*, Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature* **567**, 409-413 (2019).

43. C. R. Clapier, N. Verma, T. J. Parnell, B. R. Cairns, Cancer-Associated Gain-of-Function Mutations Activate a SWI/SNF-Family Regulatory Hub. *Mol Cell* **80**, 712-725 e715 (2020).

44. H. L. Schubert *et al.*, Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler. *Proc Natl Acad Sci U S A* **110**, 3345-3350 (2013).

45. M. Sudol, Structure and function of the WW domain. *Prog Biophys Mol Biol* **65**, 113-132 (1996).

46. Y. A. Chen *et al.*, WW Domain-Containing Proteins YAP and TAZ in the Hippo Pathway as Key Regulators in Stemness Maintenance, Tissue Homeostasis, and Tumorigenesis. *Front Oncol* **9**, 60 (2019).

47. A. J. Morrison *et al.*, Mec1/Tel1 phosphorylation of the INO80 chromatin remodeling complex influences DNA damage checkpoint responses. *Cell* **130**, 499-511 (2007).

48. P. Kapoor *et al.*, Phosphorylation-Dependent Enhancement of Rad53 Kinase Activity through the INO80 Chromatin Remodeling Complex. *Mol Cell* **58**, 863-869 (2015).

49. P. P. Singh *et al.*, Hap2-Ino80-facilitated transcription promotes de novo establishment of CENP-A chromatin. *Genes Dev* **34**, 226-238 (2020).

50. A. G. Chapman, A. M. Cotton, A. D. Kelsey, C. J. Brown, Differentially methylated CpG island within human XIST mediates alternative P2 transcription and YY1 binding. *BMC Genet* **15**, 89 (2014).

51. J. Kim, A. Kollhoff, A. Bergmann, L. Stubbs, Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum Mol Genet* **12**, 233-245 (2003).

52. J. Wang *et al.*, YY1 Positively Regulates Transcription by Targeting Promoters and Super-Enhancers through the BAF Complex in Embryonic Stem Cells. *Stem Cell Reports* **10**, 1324-1339 (2018).

53. P. Elizarev, K. Finkl, J. Muller, Distinct requirements for Pho, Sfmbt, and Ino80 for cell survival in Drosophila. *Genetics* **219**, (2021).

54. B. Turegun, R. W. Baker, A. E. Leschziner, R. Dominguez, Actin-related proteins regulate the RSC chromatin remodeler by weakening intramolecular interactions of the Sth1 ATPase. *Commun Biol* **1**, 1 (2018).

55. C. R. Clapier *et al.*, Regulation of DNA Translocation Efficiency within the Chromatin Remodeler RSC/Sth1 Potentiates Nucleosome Sliding and Ejection. *Mol Cell* **62**, 453-461 (2016).

56. S. Q. Zheng *et al.*, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).

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

58. J. Zivanov *et al.*, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, (2018).

59. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015).

60. D. Tegunov, P. Cramer, Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146-1152 (2019).

61. T. Grant, A. Rohou, N. Grigorieff, cisTEM, user-friendly software for single-particle image processing. *Elife* **7**, (2018).

62. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82 (2021).

63. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

64. K. Yamashita, C. M. Palmer, T. Burnley, G. N. Murshudov, Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallogr D Struct Biol* **77**, 1282-1291 (2021).

65. T. I. Croll, ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018).

66. T. Burnley, C. M. Palmer, M. Winn, Recent developments in the CCP-EM software suite. *Acta Crystallogr D Struct Biol* **73**, 469-477 (2017).

67. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).

68. M. Knop *et al.*, Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast* **15**, 963-972 (1999).

69. C. Janke *et al.*, A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947-962 (2004).

70. J. Rappsilber, Y. Ishihama, M. Mann, Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75**, 663-670 (2003).

71. C. E. Stieger, P. Doppler, K. Mechtler, Optimized Fragmentation Improves the Identification of Peptides Cross-Linked by MS-Cleavable Reagents. *J Proteome Res* **18**, 1363-1370 (2019).

72. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).

73. S. Lenz, S. H. Giese, L. Fischer, J. Rappsilber, In-Search Assignment of Monoisotopic Peaks Improves the Identification of Cross-Linked Peptides. *J Proteome Res* **17**, 3923-3931 (2018).

74. M. L. Mendes *et al.*, An integrated workflow for crosslinking mass spectrometry. *Mol Syst Biol* **15**, e8994 (2019).

75. C. Notredame, D. G. Higgins, J. Heringa, T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-217 (2000).

76. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).

77. S. Li, W. K. Olson, X. J. Lu, Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res* **47**, W26-W34 (2019).

Acknowledgments

We thank K. Lammens and J. Bartho for help with data collection and V. Doebler for discussion and biochemical support in nucleosome purification. We also thank T. Fröhlich from the LMU functional genome analysis. This work is supported by the Deutsche Forschungsgemeinschaft DFG (CRC1064 to K.-P.H. and B.P., CRC1361 to K.-P.H and J.C, PFA794-5/1 to B.P., RTG1721 and the Gottfried-Wilhelm-Leibniz Prize to K.-P.H), the European Research Council (ERC Advanced Grant 833613 INO3D to K.P.H) and the Max-Planck-Society (to B.P.). S.E. acknowledges an EMBO long-term fellowship and K.S. acknowledges funding by Quantitative Biosciences Munich (QBM).

Author contributions:

Conceptualization: KPH, SE

Construct design SE, KS, FK, FJM, JJ

Protein purification: FK, FJM, MM, SW, KS, JJ

Cryo-EM sample preparation: FK, FJM, SW, JJ, SE, KS

Processing: JJ, KPH, FK, FJM, SW, SE, KS

Model building: DK, JJ, KPH, SE, KS

In vitro biochemistry: FK, FJM, JC

In vivo experiments: MH, SB, BP

Visualization: FK, FJM, SW, MH

Supervision: KPH, BP

Writing-original draft: KPH, BP, FK, FJM, SW, MH, JJ, DK

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

The coordinates of the *S. cerevisiae* A-modules (ATP- and ATPγS-state) have been deposited in the Protein Data Bank (PDB) under the accession codes 8A5A and 8A5O, respectively. The *S. cerevisiae* A-module (ATP, ATPγS-state and ADP-state) and the DNA bound A-module cryo-

EM reconstructions are available at the Electron Microscopy Data Bank (EMDB) under the EMDB accession codes EMD-15163, EMD-15179, EMD-15177 and EMD-15186. *The C. thermophilum* A-module (ATPγS-state), the A-module bound to curved DNA (*apo*) and straight DNA (ADP·AlF_x), the INO80 C-module bound to nucleosome (cross grappler, *apo*) and the nucleosome bound $Ino80^{motor}$ (ADP·BeF_x) coordinates have been deposited in the PDB under the accession codes 8A5D, 8A5P, 8A5Q, 8AV6 and 8ATF, respectively. *The C. thermophilum* A-module (ATPγS-state) and the A-module bound to curved DNA (*apo*) and straight DNA (ADP·AlF_x), as well as the INO80 C-module bound to nucleosome (cross grappler, ADP·BeF_x) and the nucleosome bound $Ino80^{motor}$ (ADP·BeF_x) cryo-EM reconstructions are available at the EMDB under the EMDB accession codes EMD-15165, EMD-15180, EMD-15184, EMD-15688 and EMD-15647. The *C. thermophilum* A-module (bound to curved DNA in the ADP·AlF_x-state) cryo-EM reconstruction as well as the INO80 C-module bound to nucleosome (parallel grappler, ADP·BeF_x) cryo-EM reconstruction have been deposited at the EMDB under the EMDB accession codes EMD-15187 and EMD-15188, respectively. *C. thermophilum* $INO80^{\Delta N}$ bound to a nucleosome cryo-EM reconstruction has been deposited at the EMDB under the accession code EMD-15211. The *H. sapiens* A-module cryo-EM reconstruction has been deposited at the EMDB under the accession code EMD-15164.

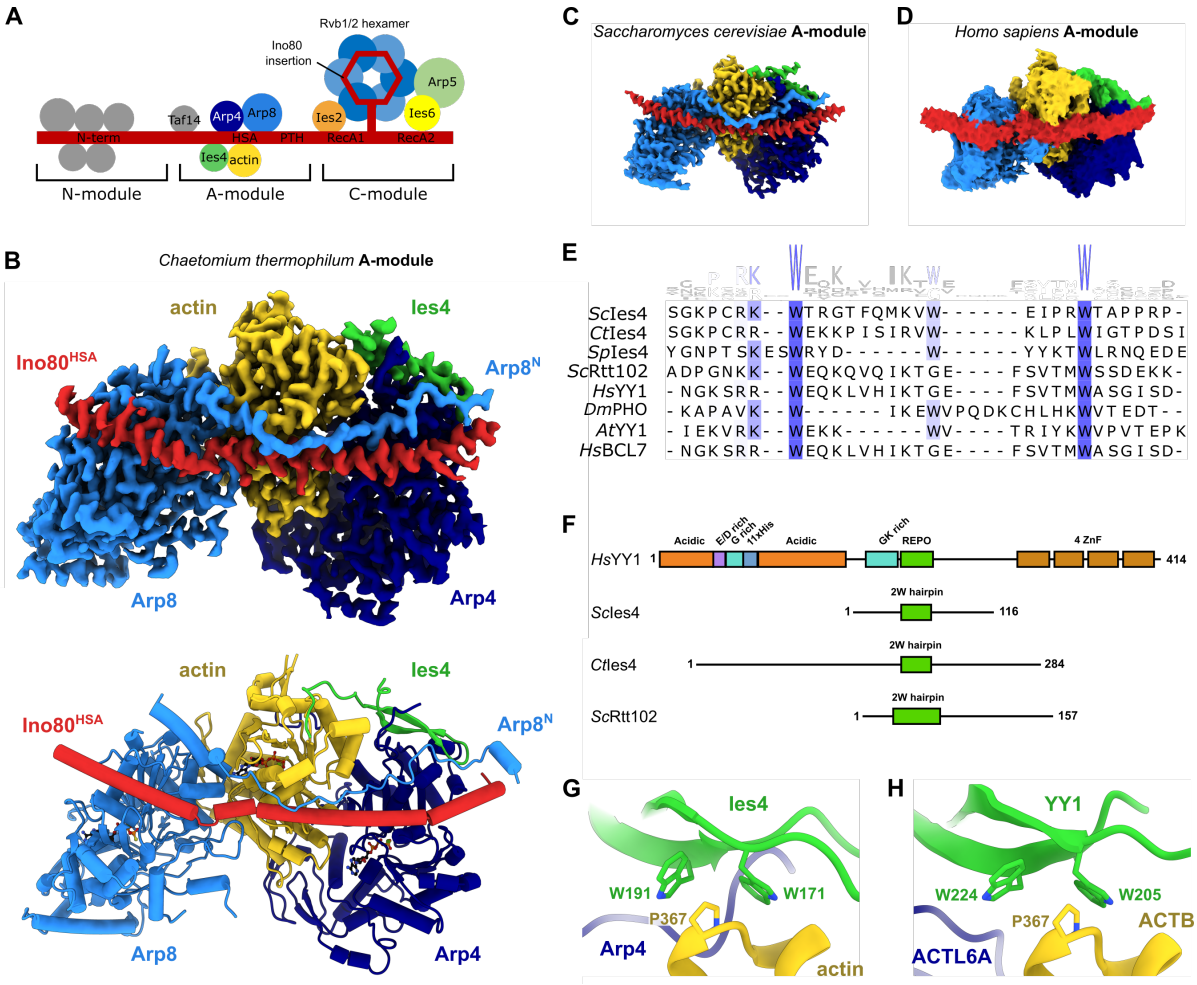


Fig. 1. Structure of the INO80 A-module. (**A**) Schematic of INO80 complex submodule and subunit organization. (**B**) Cryo-EM reconstruction (top) and structural model (bottom) of *C. thermophilum* (Ct) A-module. The protein subunits are color coded and annotated. (**C**) Cryo-EM reconstructions of *S. cerevisiae* (*Sc*) and (**D**) Cryo-EM reconstitution of *H. sapiens* (*Hs*) A-modules color coded as in (B). (**E**) Multiple sequence alignment (*75*) of the REPO/2W-motif of *S. cerevisiae* Ies4 and related actin/Arp-interacting proteins. Abbreviations: *Sc*, *Saccharomyces cerevisiae*; *Ct*, *Chaetomium thermophilum*; *Sp*, *Schizosaccharomyces pombe*; *Hs*, *Homo sapiens*; *Dm*, *Drosophila melanogaster*; *At*, *Arabidopsis thaliana.* (**F**) Domain architectures of *H. sapiens* YY1, *S. cerevisiae* Ies4, *C. thermophilum* Ies4 and *S. cerevisiae* Rtt102. The positions of the REPO/2W-motifs are indicated in green. (**G**) Detailed view of the Ies4-actin interface in *C. thermophilum*. The conserved tryptophan and proline residues are shown. (**H**) Detailed view of the proposed YY1-ACTB interface in *H. sapiens*. An AlphaFold multimer model of YY1-ACTB was used as guidance for rigid-body docking into the A-module density. The conserved tryptophan and proline residues are shown.

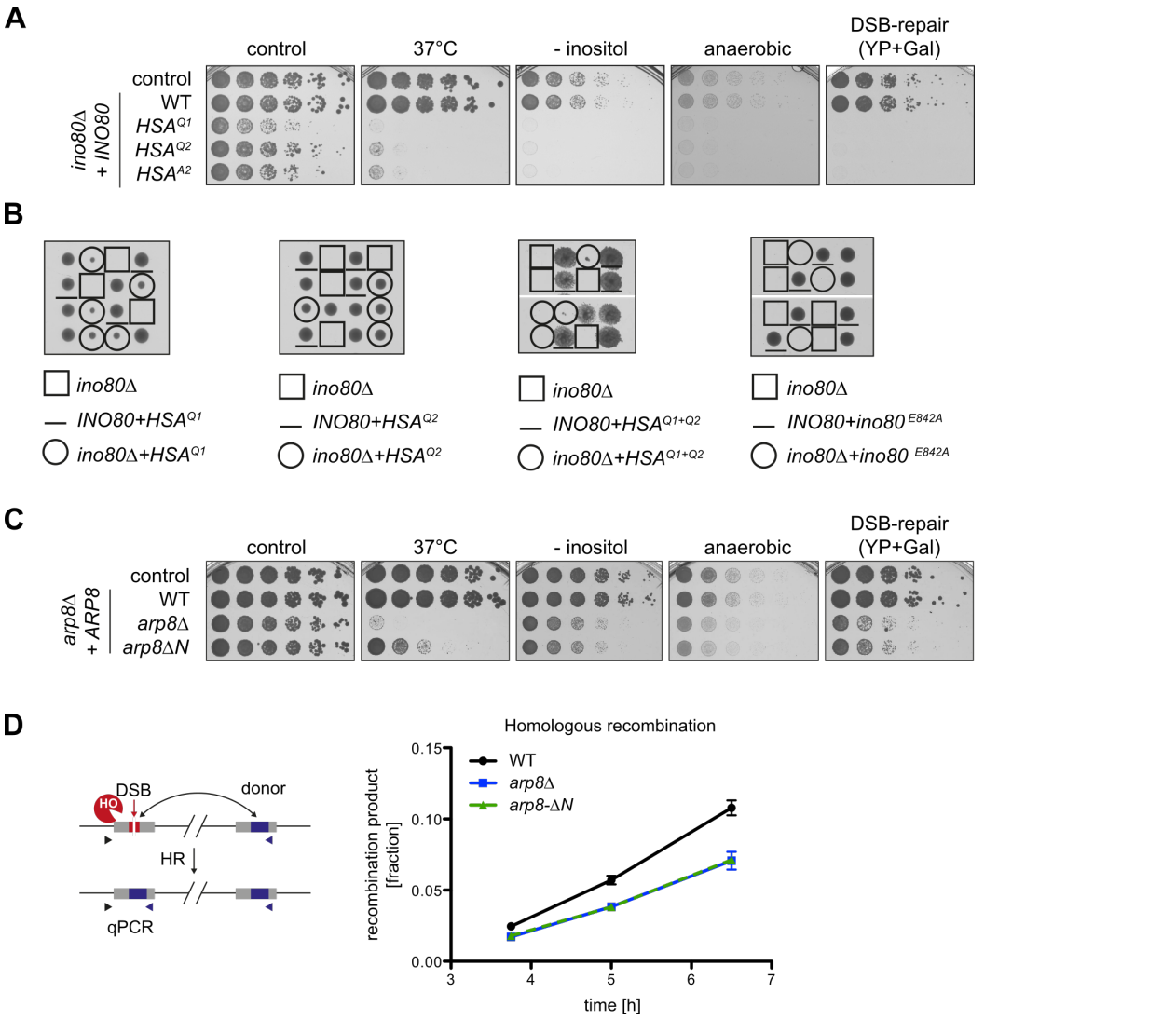


Fig. 2. HSA surface residues are critical for INO80 function in budding yeast. (**A**) Five-fold serial dilutions of budding yeast expressing indicated Ino80 variants were grown for 2–7 days. (**B**) Tetrad analysis of yeast cells bearing the indicated *INO80* alleles, where each row represents four colonies of a tetrad from a single diploid progenitor cell. *ino80-HSA*^{Q1} and *ino80-HSA*^{Q2} alleles (circled colonies in left two panels) partially rescue the *ino80*Δ lethality. The *ino80-HSA*$^{Q1+Q2}$ double mutant allele showed severely impaired or no growth (circled, third panel). The Walker B mutation (*ino80*^{E842A}) is lethal (fourth panel). (**C**) The Arp8 N-terminal region is critical for tolerance to elevated temperatures (37°C), growth on media lacking inositol, under anaerobic conditions, and for DSB repair via homologous recombination. *arp8*Δ cells were complemented with a full-length *ARP8* allele (WT) or an allele lacking the N-terminal 197 amino acids (*arp8*Δ*N*) and subjected to spot dilution growth assays as in (A). (**D**) The N-terminal region is required for Arp8 function in DSB repair by homologous recombination (HR). Left panel: Schematic of the quantitative real-time PCR (qPCR)-based analysis of HR (*39*). Cells express a galactose-inducible HO endonuclease that cuts a single defined HO-cleavage site (red, ChrIV

491 kb). The DSB can be repaired by HR using a non-cleavable donor site as repair template (blue, ChrIV 795 kb), and HR can be quantified by amplifying a recombination-dependent PCR product (triangles indicate primer positions). Right panel: Emergence of the recombination product after HO endonuclease induction (t = 0) was normalized to completed recombination (value = 1) for the strains indicated. n = 3 with error bars denoting SD.

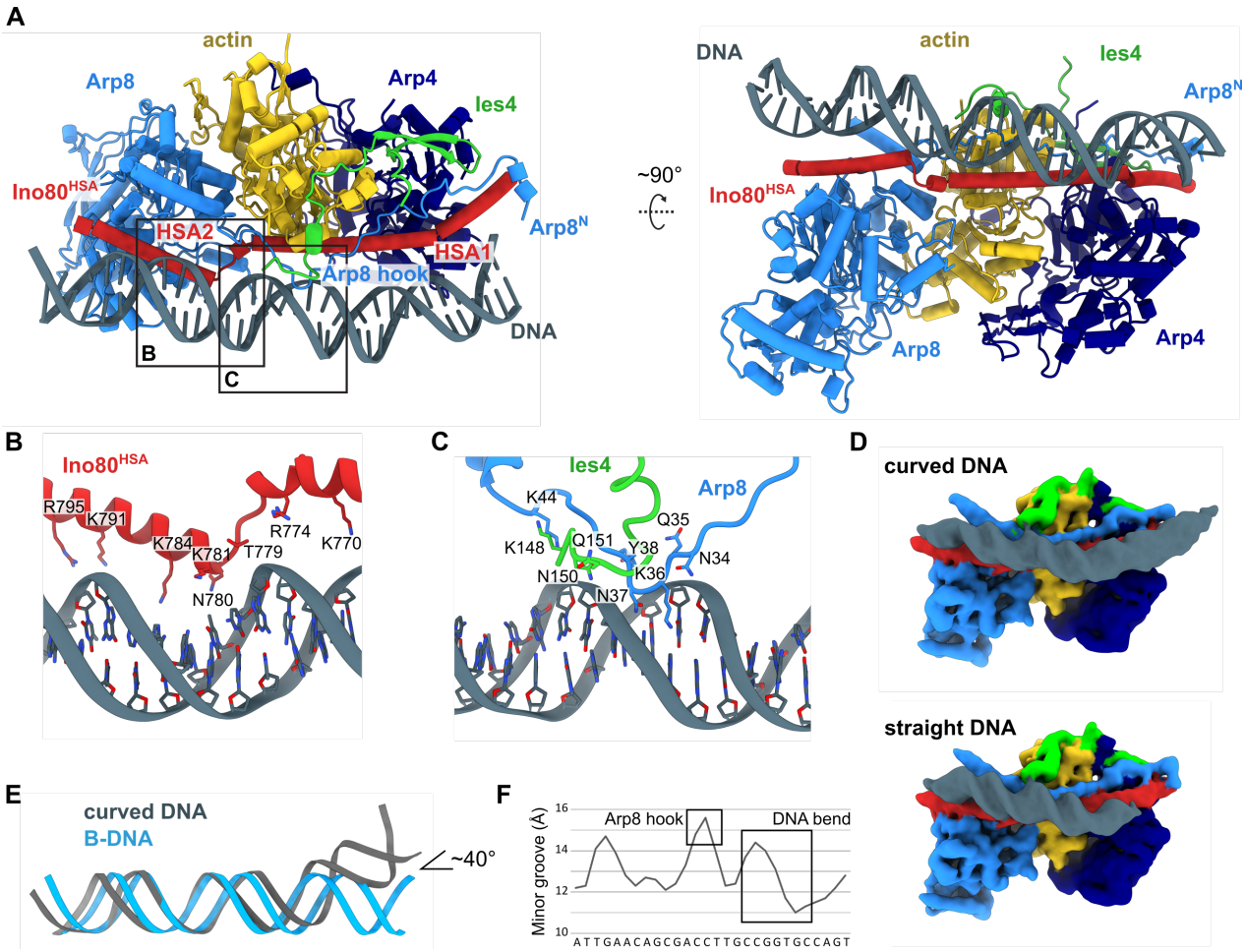


Fig. 3. Structure of the *C. thermophilum* A-module bound to DNA. (**A**) Structural model of *C. thermophilum* A-module bound to DNA. (**B**) Detailed view of the Ino80^{HSA}-DNA interaction. (**C**) Detailed view of the Arp8 hook and Ies4-DNA interaction. (**D**) Cryo-EM reconstructions of A-module bound to (top) curved and (bottom) straight DNA. Gaussian filtering was applied with a width of 1.25 (*76*) (**E**) Structural comparison of A-module bound DNA (curved DNA) and B- DNA. (**F**) Analysis of minor groove width of curved DNA (*77*). Positions of Arp8 hook interaction and the DNA bend are indicated with squares.

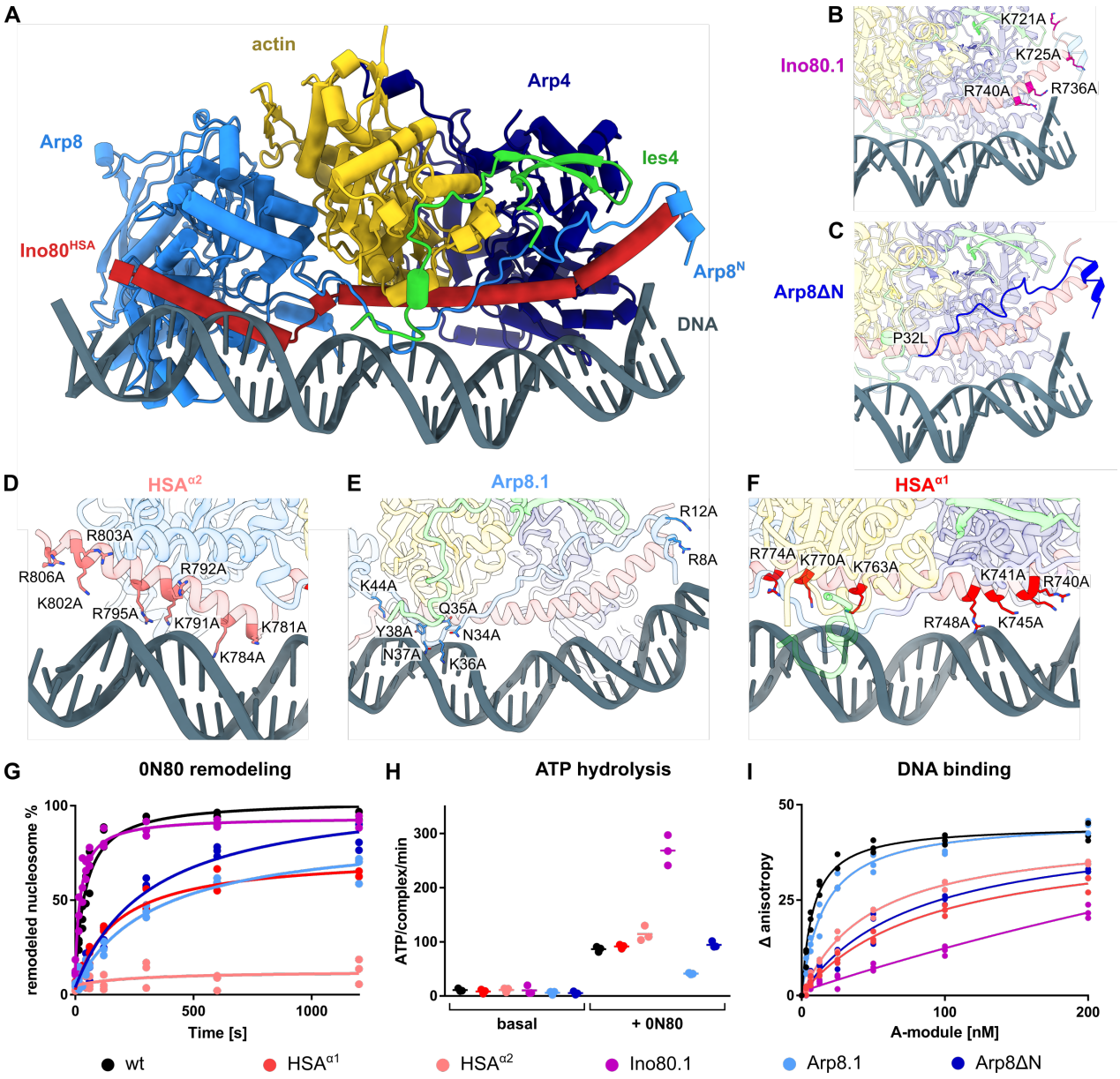


Fig. 4. Structural basis of DNA binding by the INO80 A-module. (**A**) Structural model of *C. thermophilum* A-module bound to DNA. (**B**) Ino80.1 mutation probes the distal region of HSA$^{\alpha1}$. (**C**) Illustration of the truncated portion of the Arp8 N-terminus; P32L truncation site. (**D**) HSA$^{\alpha2}$ mutations. (**E**) Arp8 mutations in hook and N-terminal helix. (**F**) HSA$^{\alpha1}$ mutations probe the central region of HSA$^{\alpha1}$. (**G**) Evaluation of the remodeling activity of *Ct*INO80$^{\Delta N}$ mutants. Band intensities of remodeled and unremodeled nucleosome species were quantified and the fraction of remodeled nucleosomes plotted against time. Data points were fitted using an exponential equation. Mean and individual data points (n = 3, technical replicates). (**H**) ATPase rate of *Ct*INO80$^{\Delta N}$ mutants with and without stimulation by nucleosomes. Rates were calculated

from the linear area of the raw data and were corrected by a buffer blank. Mean and individual data points (n = 3, technical replicates). (**I**) Fluorescence anisotropy assay to monitor the binding of *C. thermophilum* A-module and mutants to a 50bp DNA. The data were fitted to a non-linear non-cooperative 1:1 binding model. Individual data points of three independent experiments are plotted.

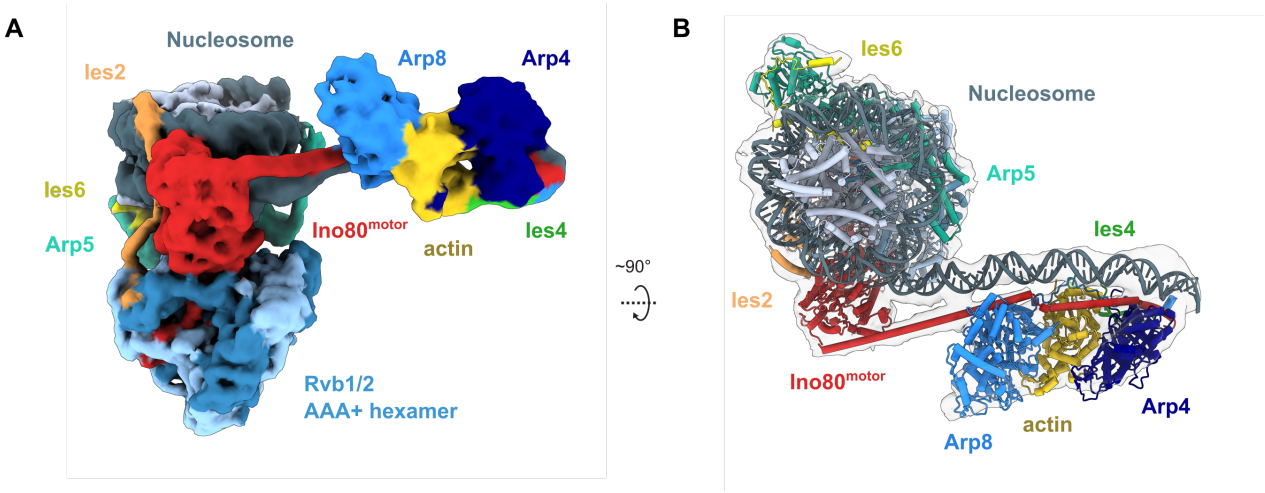


Fig. 5. The INO80 A-module interacts with extranucleosomal DNA. (**A**) Cryo-EM reconstitution (multi-body refined) of the A- and C-modules binding to nucleosome and extranucleosomal DNA. (**B**) Structural model of INO80$^{\Delta N}$ based on structures of C-module bound to the nucleosome and A-module.

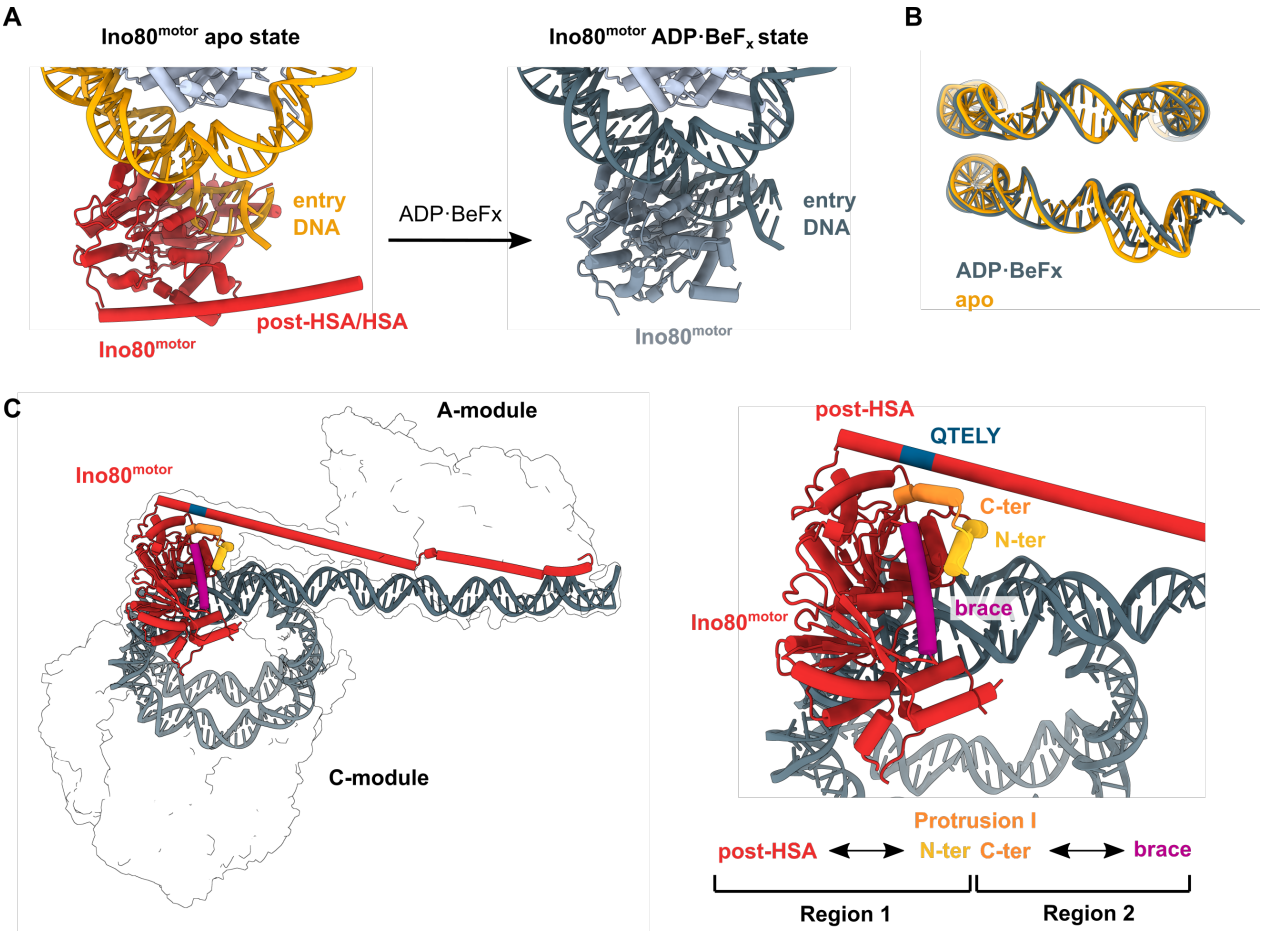


Fig. 6. Ino80^{motor} conformations in *apo* and ADP·BeF$_x$ states. (**A**) Structural model of the Ino80^{motor} interacting with the nucleosome at SHL-6 in *apo* state (left) and ADP·BeF$_x$ state (right). Structured post-HSA domain is visible in *apo* state. (**B**) Comparison of the nucleosomal DNA in *apo* state (orange) and ADP·BeF$_x$ state (grey). (**C**) Structural model of the A- and C-modules bound to nucleosome and extranucleosomal DNA. The Ino80^{motor} and post-HSA/HSA (red) and nucleosomal DNA (dark grey) are highlighted. The N-terminal helix of Protrusion I engages the post-HSA (region 1), whereas the C-terminal Protrusion I helix contacts the brace (region 2). The conserved QTELY motif forms the post-HSA contact site towards Protrusion I.

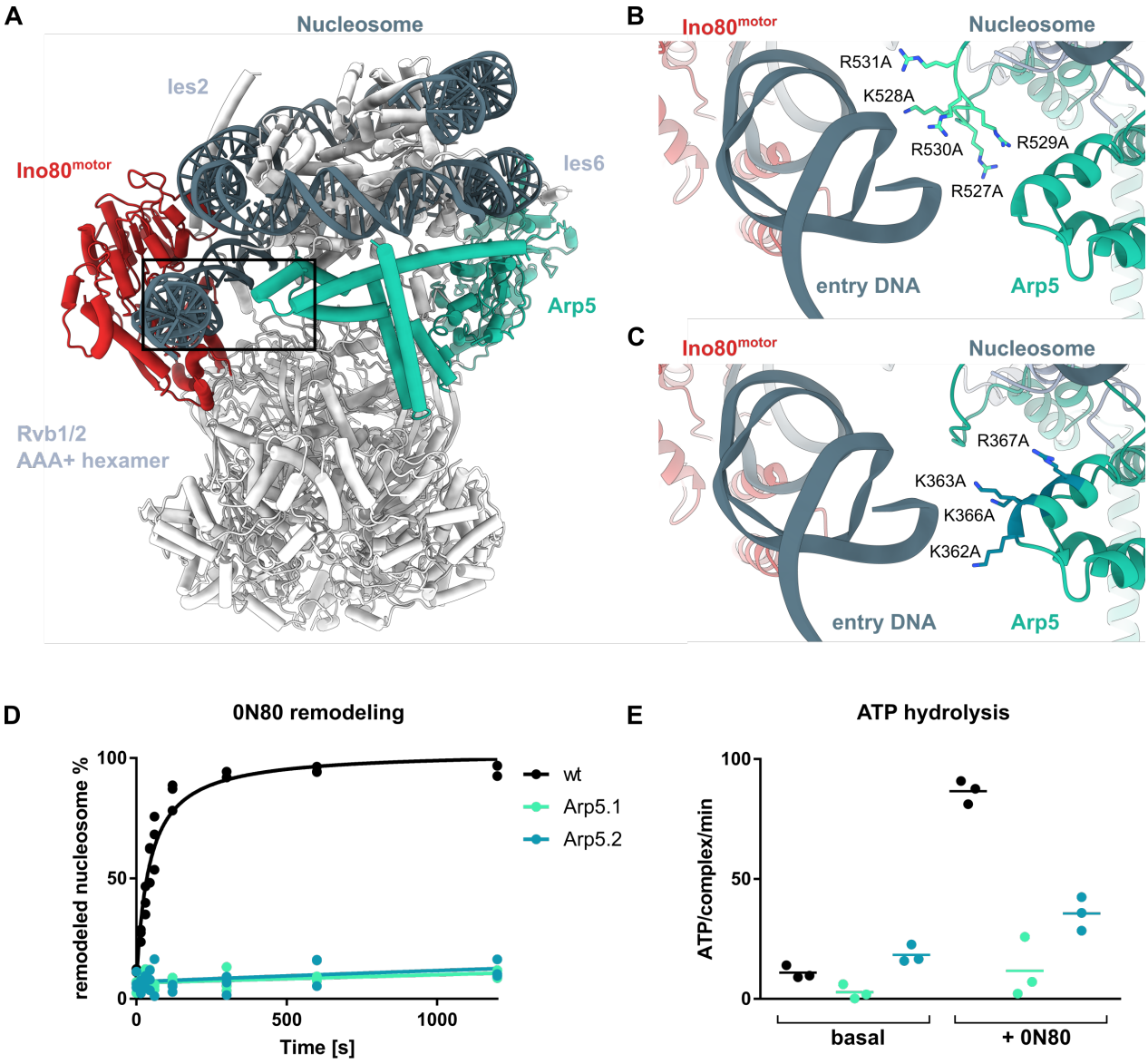


Fig. 7. The Arp5 grappler interacts with entry DNA and regulates the motor domain. (**A**) Structural model of the C-module, highlighting Arp5 (green, in "cross" configuration), Ino80^{motor} (red) and nucleosomal DNA (dark grey). (**B, C**) Docking model of two loop regions at or near entry DNA contain patches of Arg/Lys residues, suggesting they form entry DNA contact sites (**D**) Evaluation of the nucleosome sliding activity of *C. thermophilum* INO80 Arp5 grappler mutants. Band intensities of remodeled and unremodeled nucleosome species were quantified and the fraction of remodeled nucleosomes plotted against time. Data points were fitted using an exponential equation. Mean and individual data points (n = 3, technical replicates) (**E**) ATPase rate of *C. thermophilum* INO80$^{\Delta N}$ and mutants with and without stimulation by nucleosomes. Rates were calculated from the linear area of the raw data and were corrected by a buffer blank. Mean and individual data points (n = 3, technical replicates).

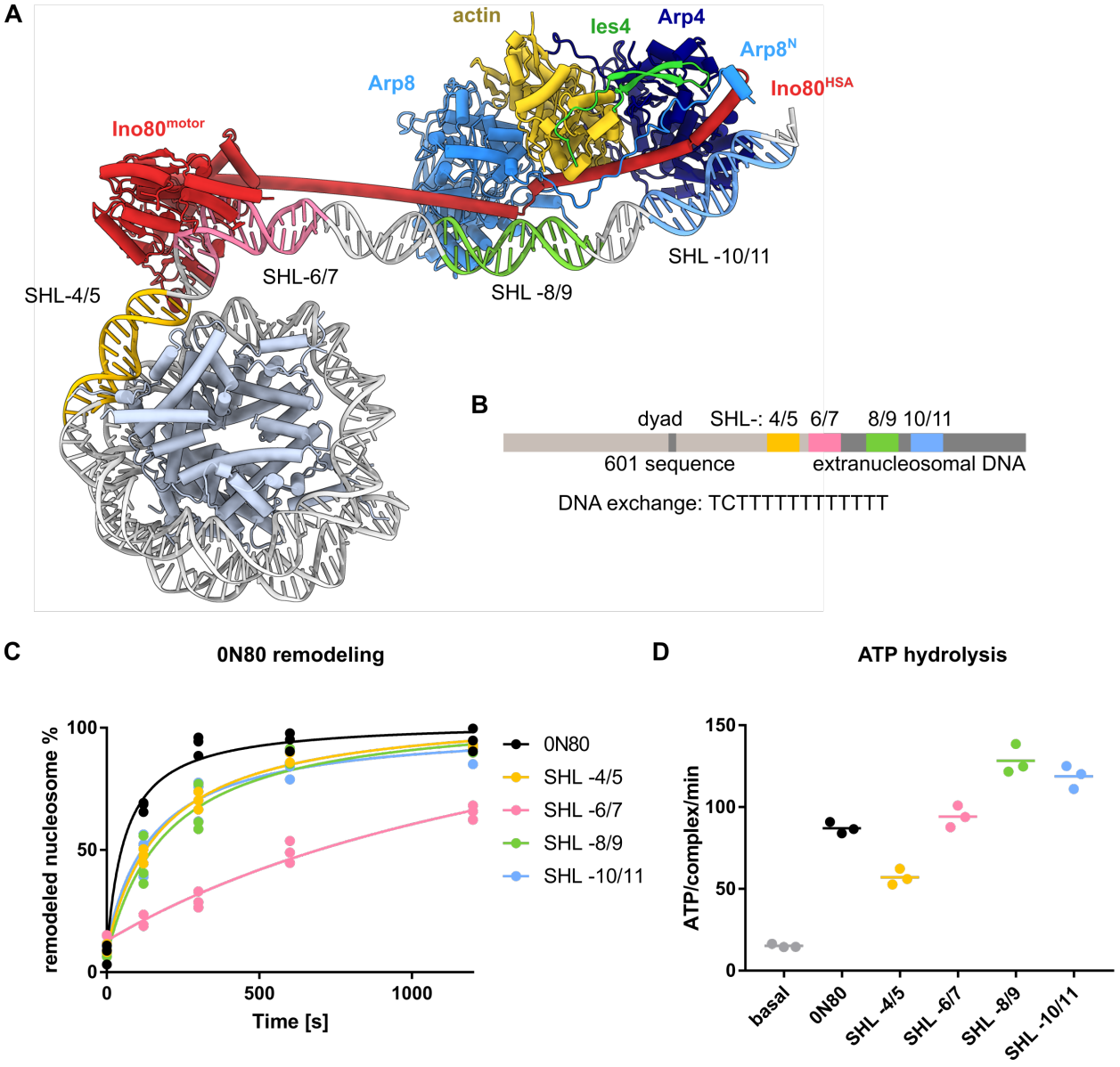


Fig. 8. Influence of A/T-rich DNA on *Ct*INO80$^{\Delta N}$ nucleosome remodeling. (**A**) Location of exchanged DNA cassettes on the structural docking model. (**B**) Schematic visualization of exchanged DNA sequence cassettes in distance to the dyad of the nucleosome. (**C**) Evaluation of the sliding activity of *Ct*INO80$^{\Delta N}$ with different nucleosomal substrates. Band intensities of remodeled and unremodeled nucleosome species were quantified and the fraction of remodeled nucleosomes plotted against time. Data points were fitted using an exponential equation. Mean and individual data points (n = 3, technical replicates). (**D**) ATPase rate of *Ct*INO80$^{\Delta N}$ with and without stimulation by different nucleosomes. Rates were calculated from the linear area of the raw data and were corrected by a buffer blank. Mean and individual data points (n = 3, technical replicates).

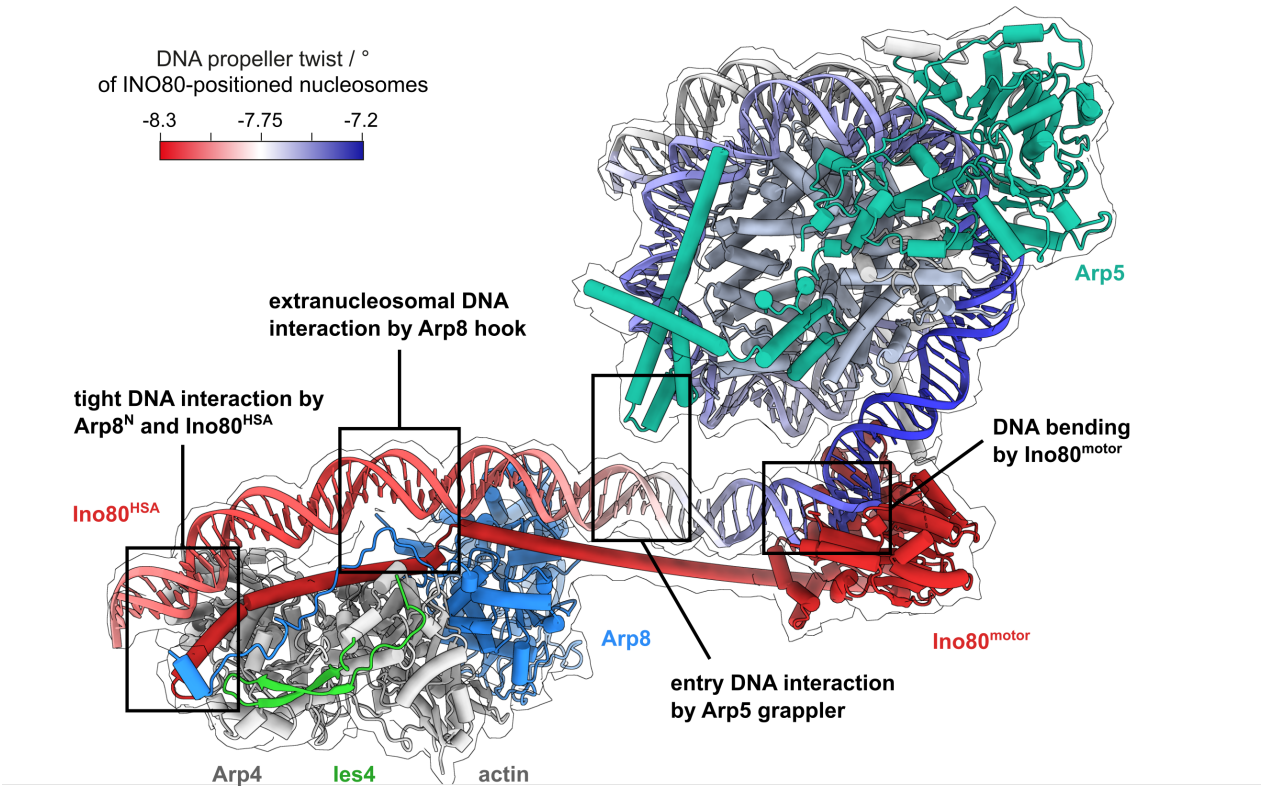


Fig. 9. Model of multivalent INO80$^{\Delta N}$-DNA interactions. The unified model integrates our structural and biochemical analysis. The Ino80^{motor} engages the nucleosome at SHL-6. The Arp5 grappler contacts entry DNA opposite of the Ino80^{motor}. The A-module binds extranucleosomal DNA and is linked to the Ino80^{motor} via the post-HSA/HSA domain. Propeller twist DNA shape data of INO80-positioned nucleosomes (*15*) were mapped onto model of linker and nucleosomal DNA by using red–white–blue color gradient.

Supplementary Materials for

Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Franziska Kunert1,†, Felix J. Metzner1,†, James Jung1,†, Markus Höpfler2,3,‡, Stephan Woike1,‡, Kevin Schall1,4, Dirk Kostrewa1, Manuela Moldt1, Jia-Xuan Chen5, Susanne Bantele2,6, Boris Pfander2,7,8, Sebastian Eustermann1,9,†, Karl-Peter Hopfner1*

*Corresponding author. Email: hopfner@genzentrum.lmu.de

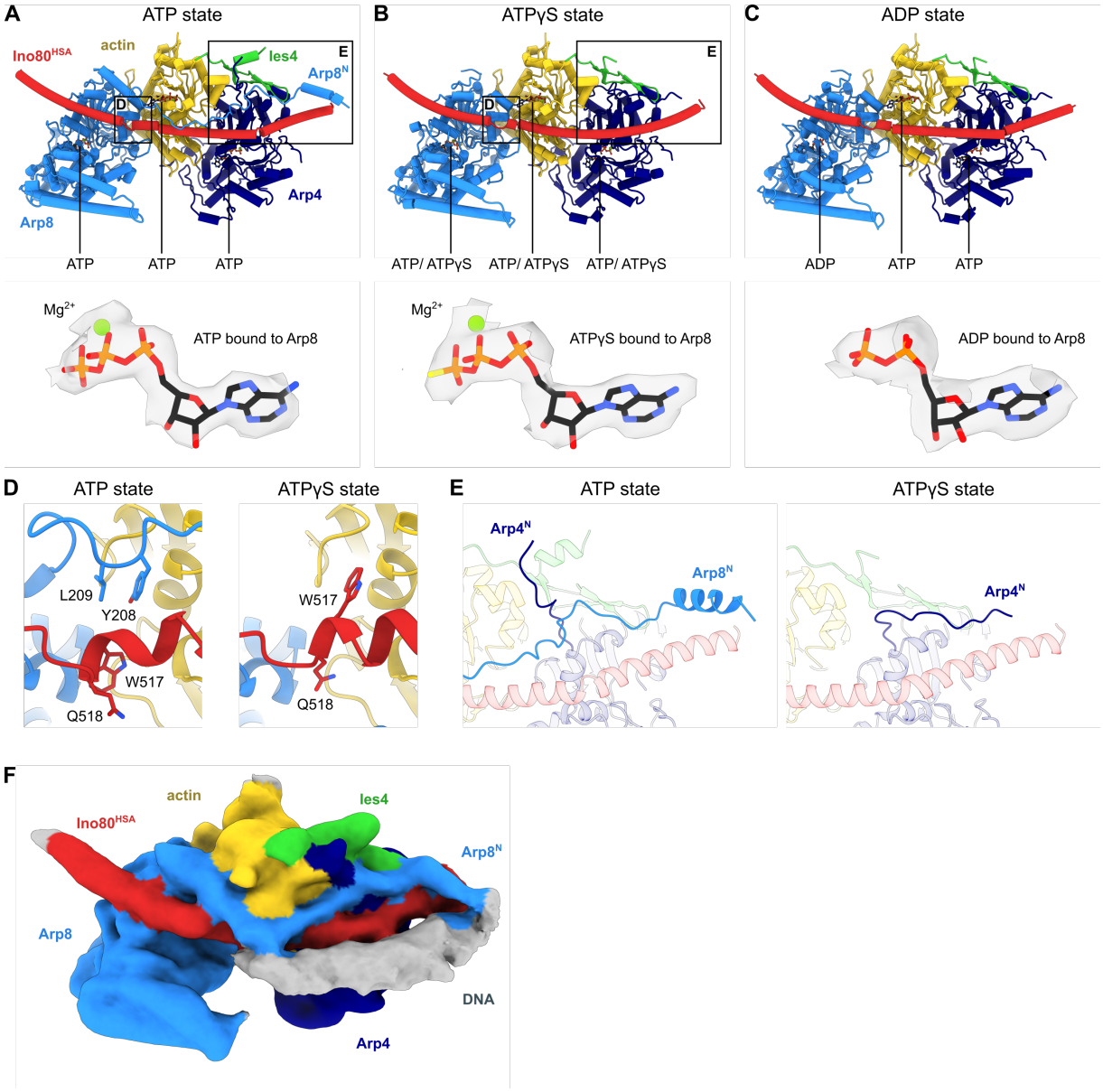


Fig. S1. The Arp8 N-terminus of the *S. cerevisiae* A-module binds DNA. (A-C) Structural model of the *S. cerevisiae (Sc)* A-module in (A) ATP state, (B) ATPγS state and (C) ADP state (ADP state is not a fully refined model, ATPγS state docked into ADP state reconstruction and ADP modelled into Arp8 nucleotide binding site). The protein subunits are color coded and annotated and the nucleotide states of Arp8, actin and Arp4 are indicated. (Bottom) Cryo-EM density maps of nucleotides bound to Arp8 (surface cutoff: 2 Å) *(62)* (D) Detailed view of the hydrophobic anchors of Ino80^{HSA} (W517) and Arp8^{N} (Y208) (left: ATP state, right: ATPγS state). Upon binding of Arp8^{N} along the A-module, Arp8 Y208 takes the position of Ino80 W517. (E) Detailed view of the N-termini of Arp4 and Arp8. In the ATP state (left), Arp8^{N} extends along Ino80^{HSA}. In the ATPγS state (right), Arp8^{N} is not resolved and Arp4^{N} takes the position of Arp8^{N}. (F) Cryo-EM reconstructions of A-module bound to DNA (ATP state). The protein subunits are color coded and annotated.

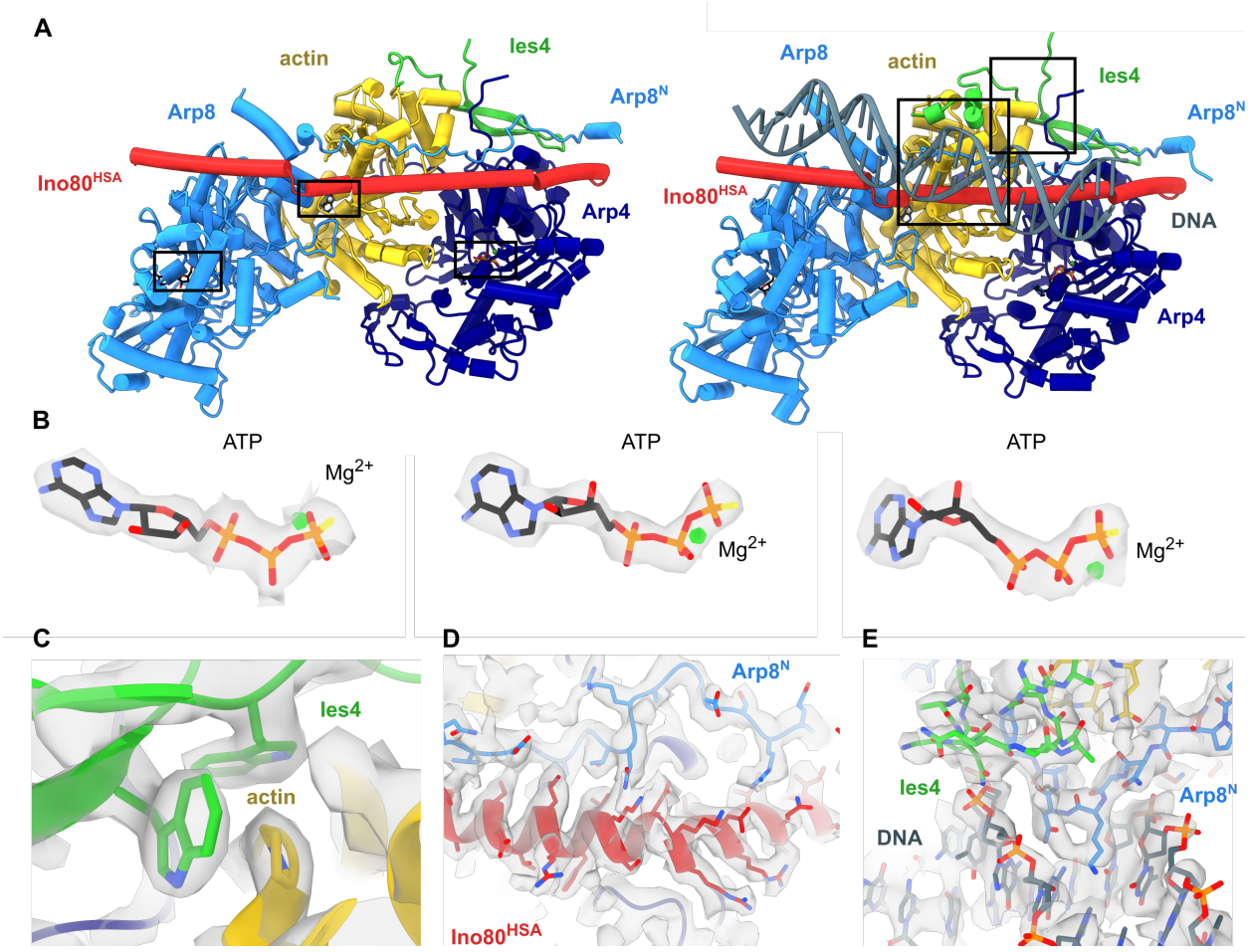


Fig S2: Examples of cryo-EM density map of the *C. thermophilum* A-module. (**A**) Structural model of *C. thermophilum (Ct)* A-module (left) without DNA or (right) with DNA bound. The protein subunits are color coded and annotated. (**B**) Cryo-EM density maps of nucleotides (ATP) bound to (left to right) Arp8, actin and Arp4 (surface cutoff: 2 Å) *(62)*. (**C**) Detailed view of the Ies4-actin interface in the cryo-EM density map of the *Ct*A-module. (**D**) Detailed view of the Ino80^{HSA}/Arp8^{N} interface in the cryo-EM density map of the *Ct*A-module. (**E**) Detailed view of the Ies4/Arp8^{N}-DNA interface in the cryo-EM density map of the *Ct*A-module bound to DNA.

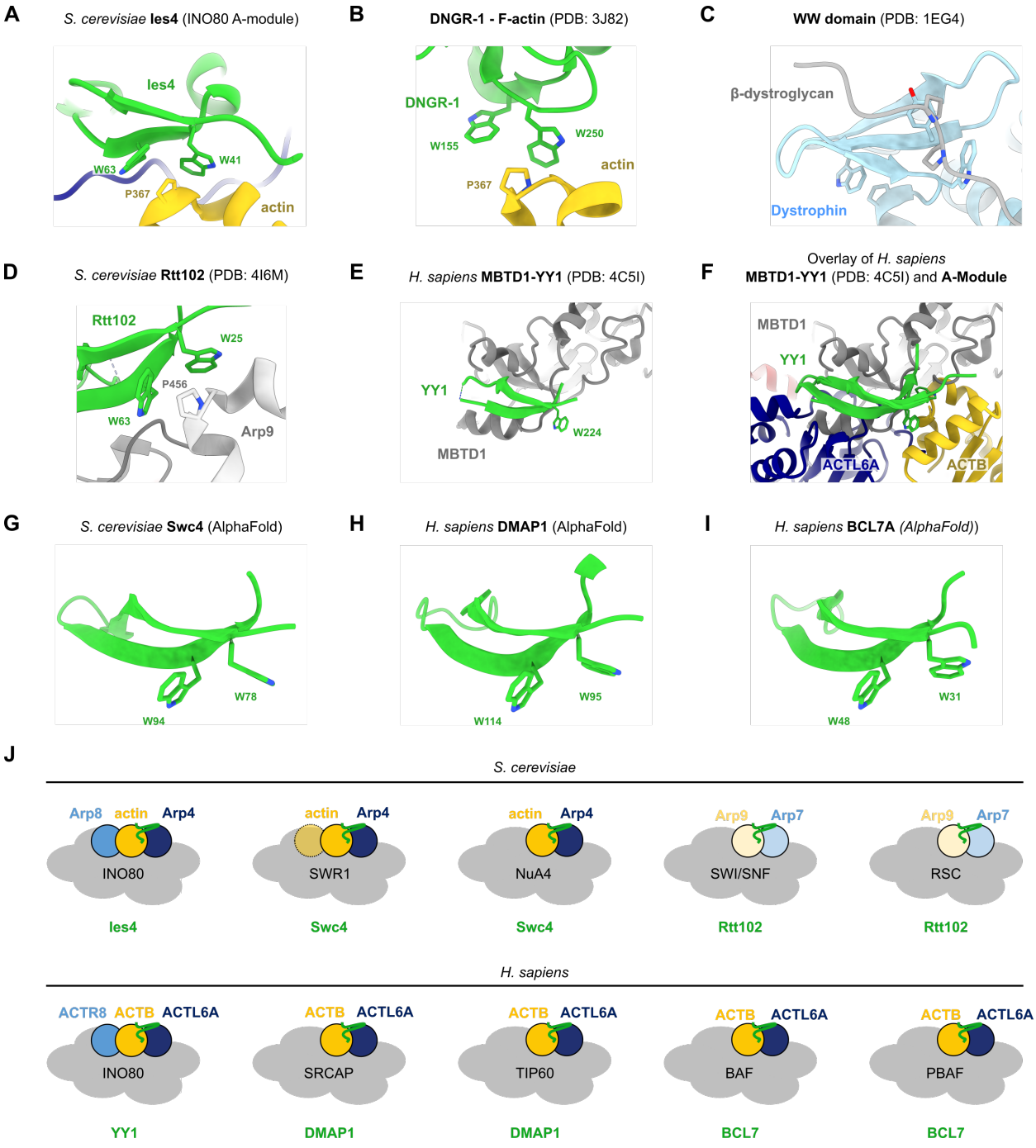


Fig. S3. The 2W-hairpin shows a conserved interaction mode with actin/ARPs. (**A**) Detailed view of the Ies4-actin interface in the *S. cerevisiae* A-module. The conserved tryptophan and proline residues are shown. (**B**) Detailed view of the DNGR-1-actin interface (PDB: 3J82). The conserved tryptophan and proline residues are shown. (**C**) Structure of the dysotrophin WW-domain in complex with a β-dystroglycan peptide (PDB: 1EG4). Conserved

tryptophan and proline residues are shown. (**D**) Detailed view of the Rtt102-Arp9 interface in *S. cerevisiae* SWI/SNF A-module (PDB: 4I6M). The conserved tryptophan and proline residues are shown. (**E**) Structure of *H. sapiens* YY1 bound to MBTD1 (PDB: 4C5I). The conserved tryptophan residue is shown. (**F**) Structural comparison of alternative YY1 binding modes. Structures of *H. sapiens* YY1 bound to MBTD1 (PDB: 4C5I) and *H. sapiens* YY1 bound to INO80 A-module are aligned onto the YY1 subunit. The conserved tryptophan and proline residues are shown. (**G**) AlphaFold-prediction of 2W-hairpin of *S. cerevisiae* Swc4. The conserved tryptophan residues are shown. (**H**) AlphaFold-prediction of 2W-hairpin of *H. sapiens* DMAP1. The conserved tryptophan residues are shown. (**I**) AlphaFold-prediction of 2W-hairpin of *H. sapiens* BCL7A. The conserved tryptophan residues are shown. (**J**) Illustration of the conservation of the Arp4-actin heterodimer in *S. cerevisiae* (Arp4-actin, Arp9-Arp7) and *H. sapiens* (ACTL6A-ACTB) INO80 and SWI/SNF family chromatin remodeling complexes. The respective 2W-hairpin containing complex subunits are indicated in green.

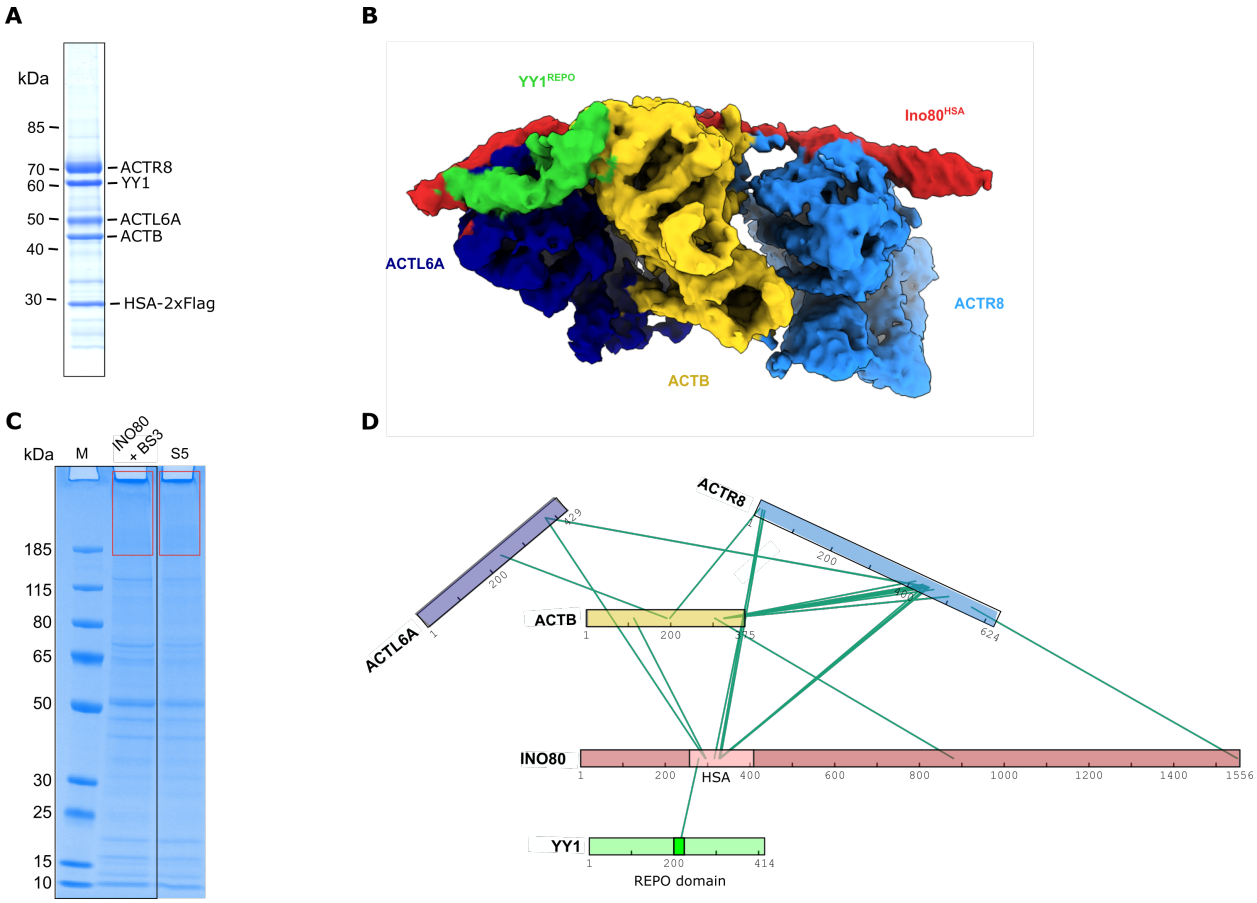


Fig. S4. YY1 interacts with the *H. sapiens* Ino80^{HSA}. (**A**) Coomassie-stained SDS-PAGE gel of the purified *H. sapiens (Hs)* A-module (**B**) Cryo-EM reconstruction of *Hs*A-module. The density is colored according to the underlying protein subunits. (**C**) Coomassie-stained SDS-PAGE gel showing the BS3 crosslinked INO80 complex. The red square comprising the high-molecular weight crosslinked species indicates the cut-out region for in gel digest used for mass spectrometry (MS) analysis. (**D**) Topological crosslink-MS scheme of *Hs*INO80 A-module subunits, showing inter-protein links (green).

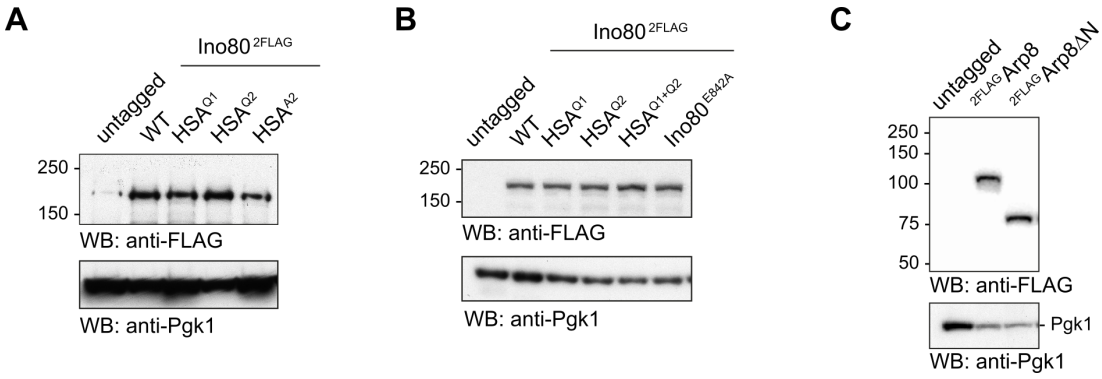


Fig. S5. HSA surface residues are critical for INO80 function in budding yeast. (**A**) Ino80 variants are expressed at levels similar to the Ino80 WT protein. Yeast cells as in Fig. 2A expressing the indicated 2FLAG-tagged Ino80 variants under the control of the endogenous *INO80* promoter were subjected to total protein analysis by western blotting using an anti-FLAG antibody. Pgk1 levels served as control. (**B**) Ino80 variants are expressed at levels similar to the Ino80 WT protein. Diploid yeast cells used for tetrad analysis in Fig. 2B expressing the indicated 2FLAG-tagged Ino80 variants under the control of the endogenous *INO80* promoter were subjected to total protein analysis by western blotting using an anti-FLAG antibody. Pgk1 levels served as control. (**C**) The Arp8∆N variant is expressed similar to Arp8 WT levels. Yeast cells as in Fig. 2C and D expressing the indicated 2FLAG-tagged Arp8 variants under the control of the endogenous *ARP8* promoter were subjected to total protein analysis by western blotting using an anti-FLAG antibody. Pgk1 levels served as control.

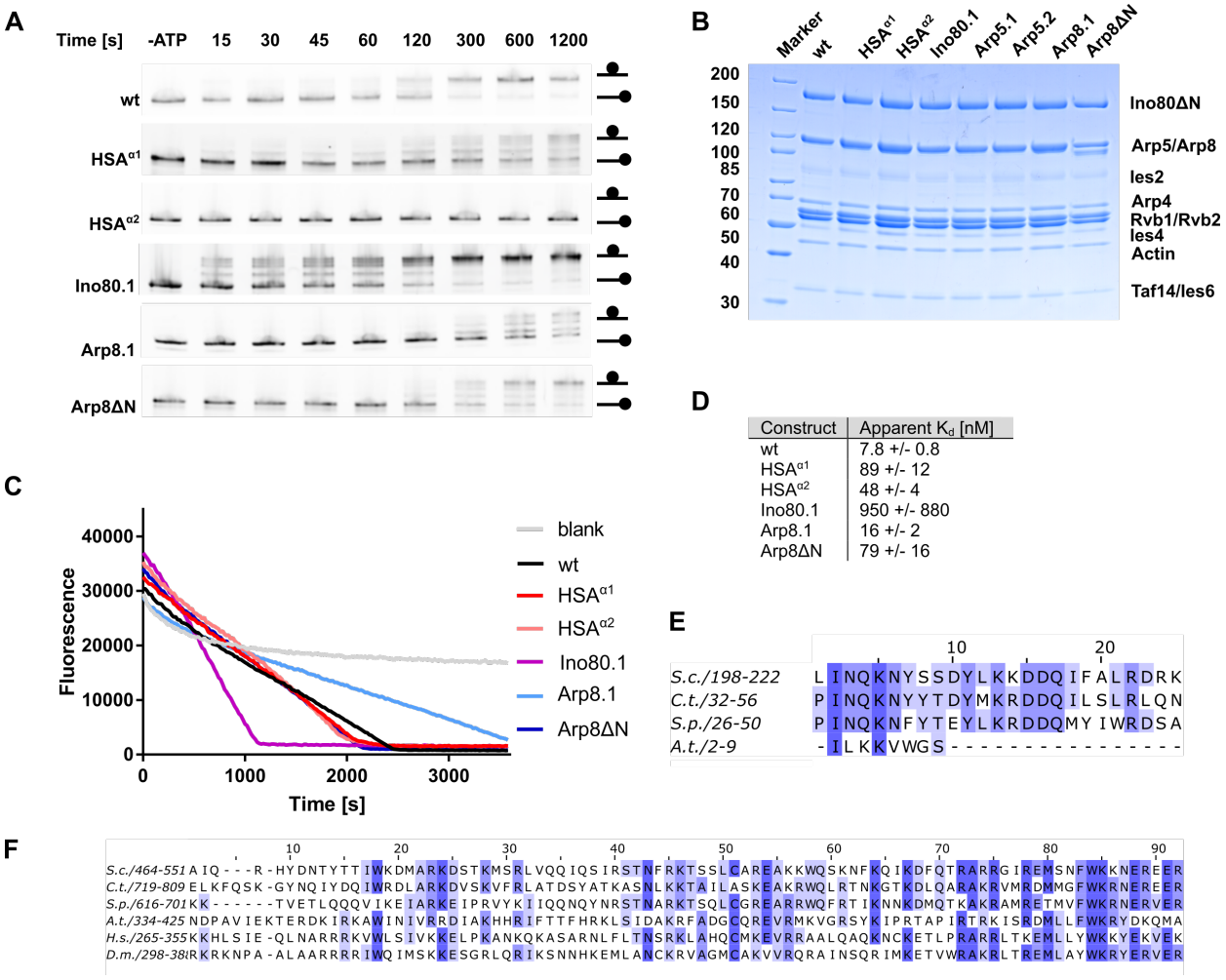


Fig. S6. *C. thermophilum* INO80 mutants influence sliding and ATPase activity. (**A**) Sliding of 0N80 nucleosomes by *Ct*INO80$^{\Delta N}$ and mutants analyzed by native PAGE. (**B**) Coomassie stained SDS-PAGE gel of *Ct*INO80$^{\Delta N}$ and mutants (**C**) Raw data of ATPase assays. ATPase rates were determined for *Ct*INO80$^{\Delta N}$ wild type (WT) and the mutants, along with nucleosome-stimulated rates. (**D**) Fluorescence anisotropy assay to monitor the binding of *Ct*A-module and mutants to a 50bp DNA. The data were fitted to a non-linear non-cooperative 1:1 binding model and the apparent K_d values were calculated. The mean +/- SEM of three independent experiments are shown. (**E**) Multiple sequence alignment *(75)* of Arp8 N-terminus (Arp8 hook). *S.c., S. cerevisiae; C.t., C. thermophilum; S.p., S. pombe; A.t., A. thaliana.* (**F**) Multiple sequence alignment *(75)* of the Ino80^{HSA}. *S.c., S. cerevisiae; C.t., C. thermophilum; S.p., S. pombe; A.t., A. thaliana; H.s., Homo sapiens; D.m., D. melanogaster.*

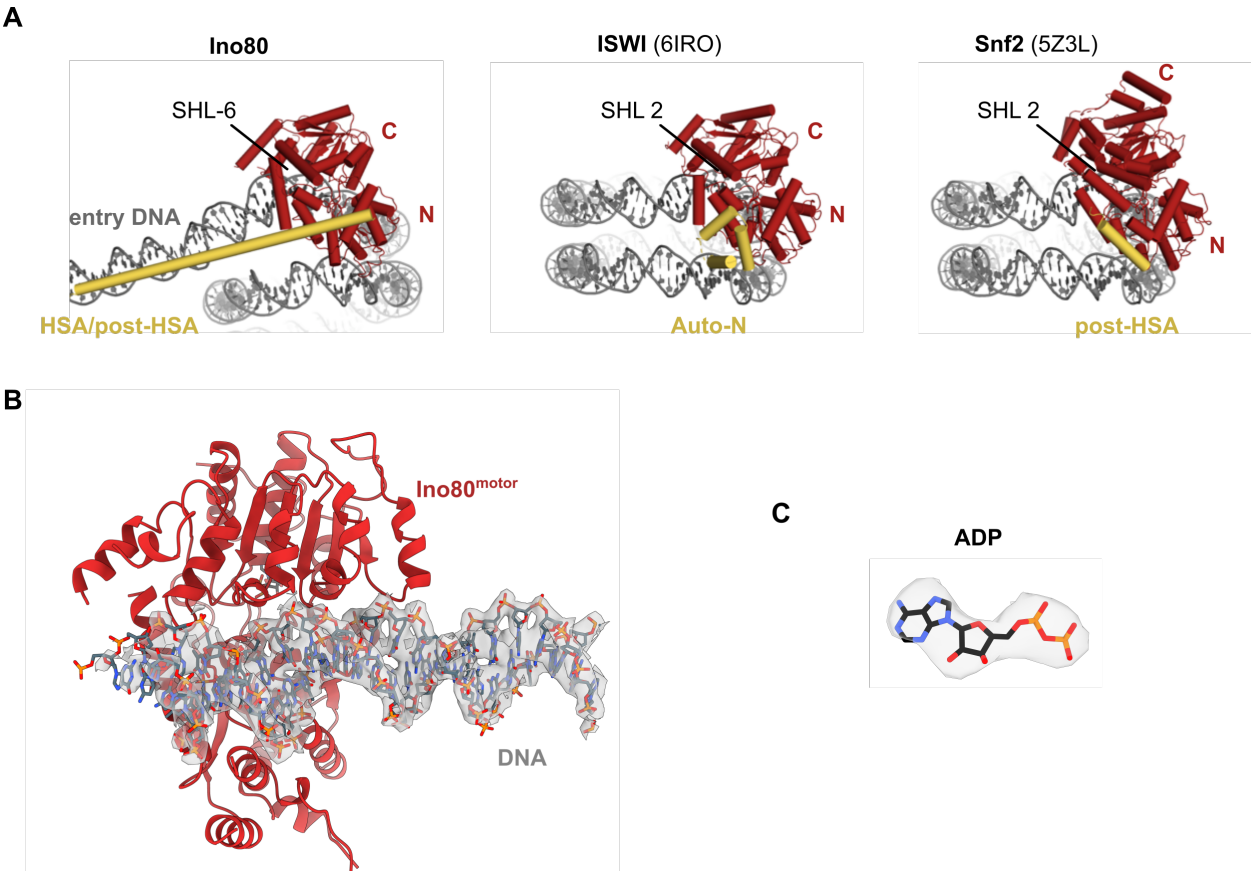


Fig. S7. Example of cryo-EM density map of the Ino80^{motor} (ADP·BeF$_x$ state**).** (**A**) Comparison of Ino80, ISWI and Snf2 interacting with nucleosomes and similarity of post-HSA and Auto-N. Ino80 binds the nucleosome at SHL-6 while Isw1 and Snf2 bind at SHL-2. (**B**) Detailed view of the cryo-EM density map of the Ino80^{motor}-bound DNA of the *C. thermophilum* C-module. The protein subunit is color coded and annotated (surface cutoff: 2 Å) *(62).* (**C**) Cryo-EM density map of ADP bound to the Ino80^{motor} (surface cutoff: 2 Å) *(62)*. Note, the light BeF$_x$ moiety is not visible in the density map, in line with other studies *(19).*

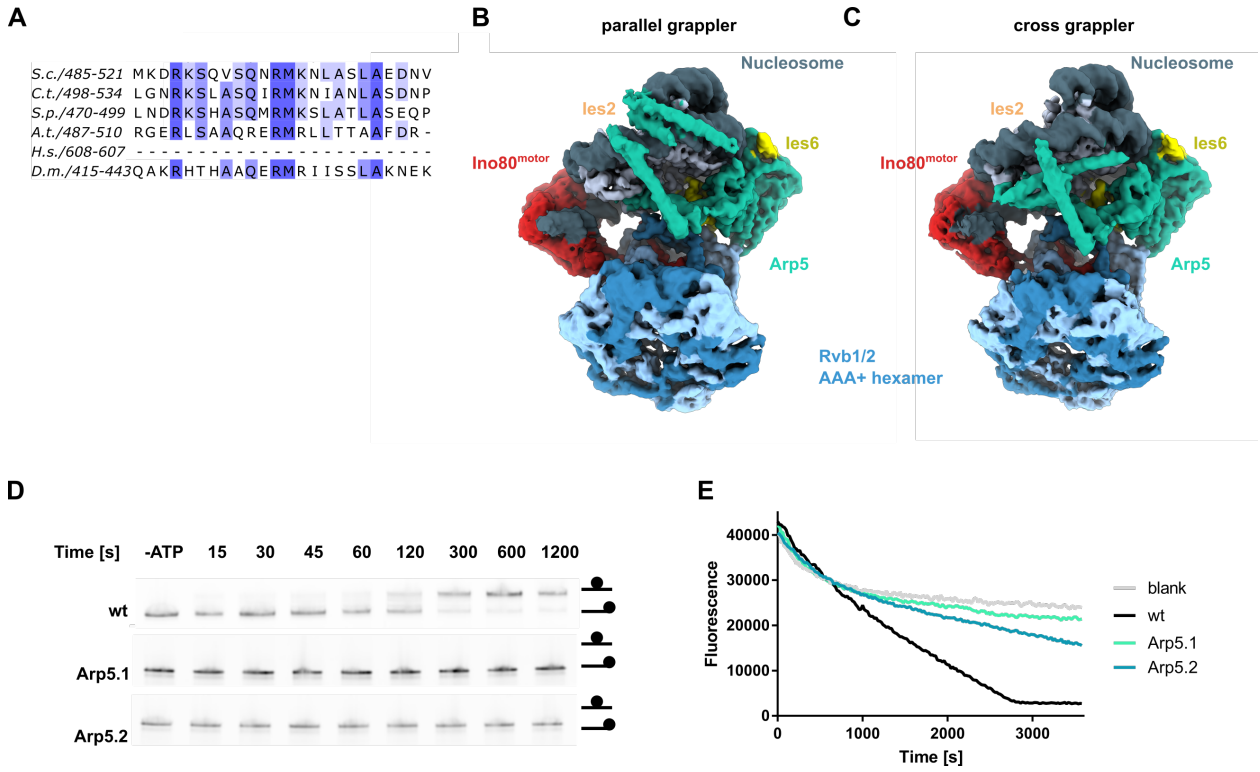


Fig. S8. Arp5 entry DNA interaction is important for *Ct*INO80$^{\Delta N}$ sliding and ATPase activity. (**A**) Multiple sequence alignment *(75)* of the grappler "foot". *S.c., S. cerevisiae; C.t., C. thermophilum; S.p., S. pombe; A.t., A. thaliana; H.s., Homo sapiens; D.m., D. melanogaster*. (**B**-**C**) Cryo-EM reconstitution of the nucleosome bound INO80 C-module complex with the Arp5 grappler in (**B**) parallel and (**C**) cross conformation. The protein subunits are color coded and annotated. (**D**) Sliding of 0N80 nucleosomes by *Ct*INO80$^{\Delta N}$ and mutants analyzed by native PAGE. (**E**) Raw data of ATPase assays. ATPase rates were determined for *Ct*INO80$^{\Delta N}$ wild type (wt) and the mutants, along with nucleosome-stimulated rates.

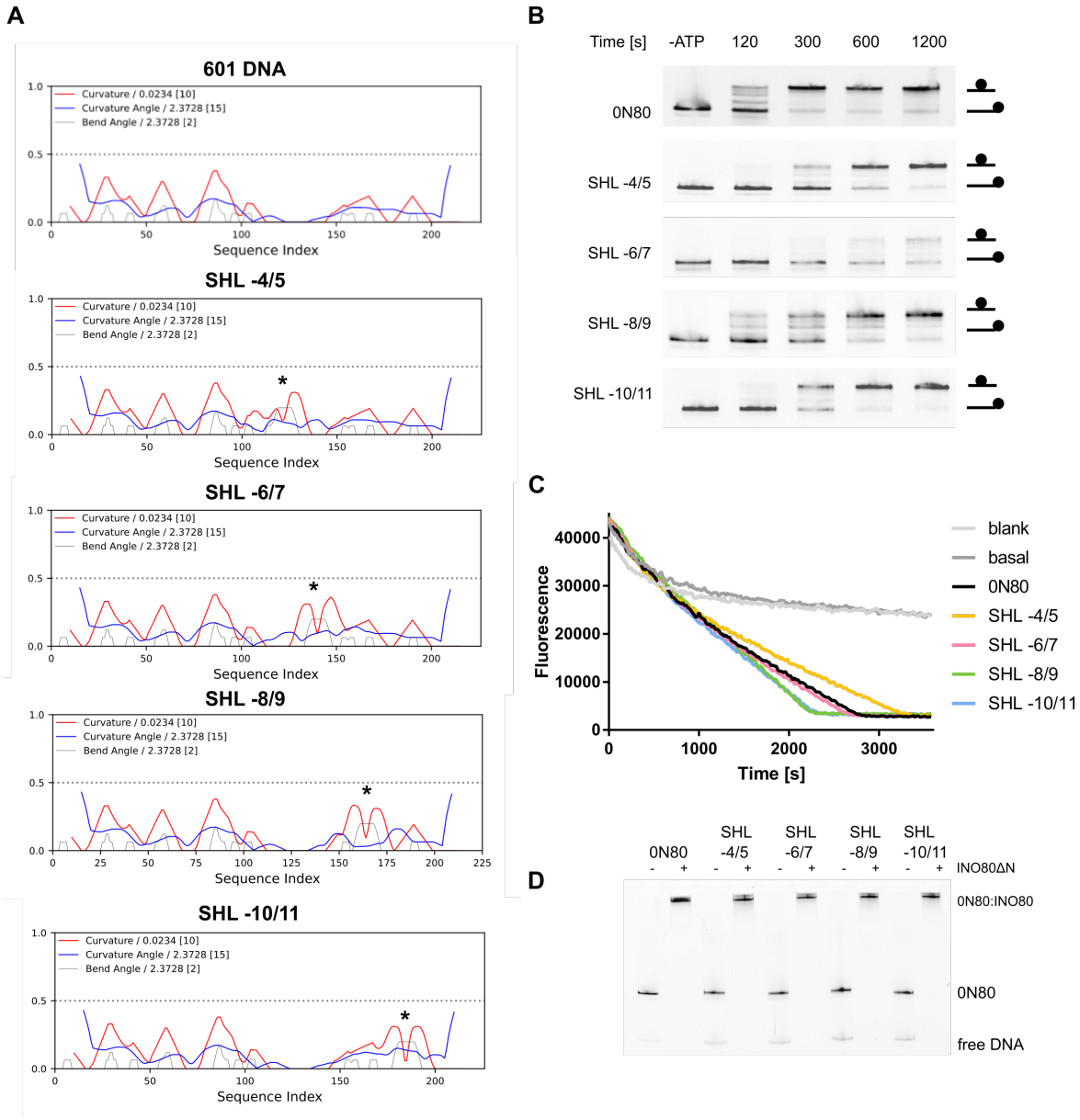


Fig. S9. Influence of A/T-rich DNA on *Ct*INO80$^{\Delta N}$ nucleosome remodeling. (**A**) DNA curvature analysis of the 601-based 0N80 DNA template sequence and sequences with replaced A/T-rich DNA cassettes (https://github.com/cgohlke/dnacurve). (**B**) Sliding of different 0N80 nucleosomes by *Ct*INO80$^{\Delta N}$ analyzed by native PAGE. (**C**) Raw data of ATPase assays in presence of different 0N80 nucleosomes. ATPase rates were determined for *Ct*INO80$^{\Delta N}$ wild type (wt) along with nucleosome-stimulated rates. (**D**) Electrophoretic mobility shift assay of different 0N80 nucleosome substrates bound to *Ct*INO80$^{\Delta N}$ analyzed by native page.

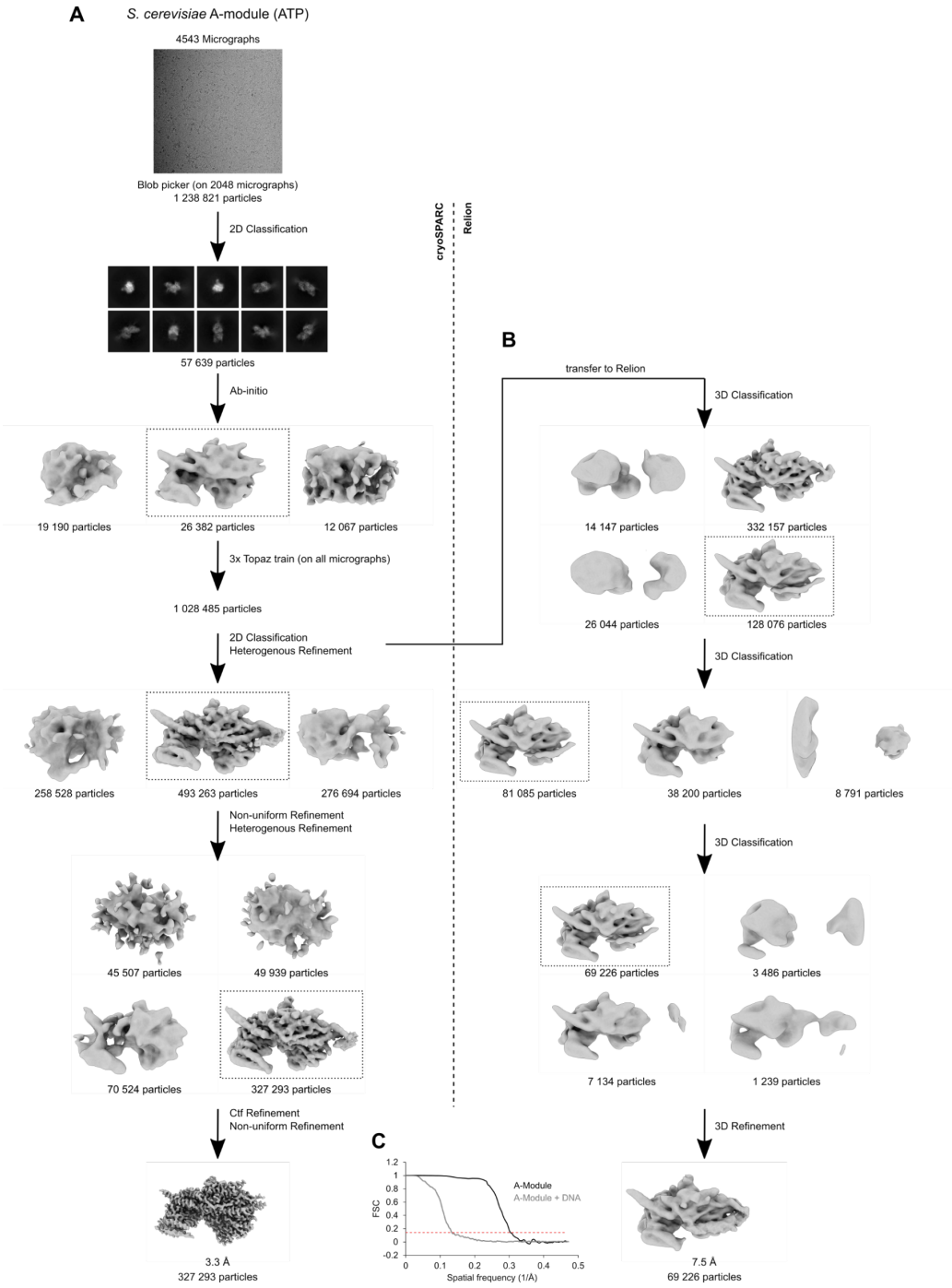


Fig. S10. Cryo-EM data analysis of *S. cerevisiae* A-module (ATP-state). (**A**) Cryo-EM data processing workflow of *S. cerevisiae* A-module using cryoSPARC v3.2.0 *(57)* and (**B**) A-module bound to DNA using cryoSPARC v3.2.0 *(57)* and Relion-3.0 *(58)*. A representative micrograph, representative 2D classes and the cryo-EM data processing workflow are shown. (**C**) Gold-standard Fourier shell correlation (FSC) curves of the final A-module and DNA bound A-module reconstructions. The red line indicates the 0.143 cutoff criterion.

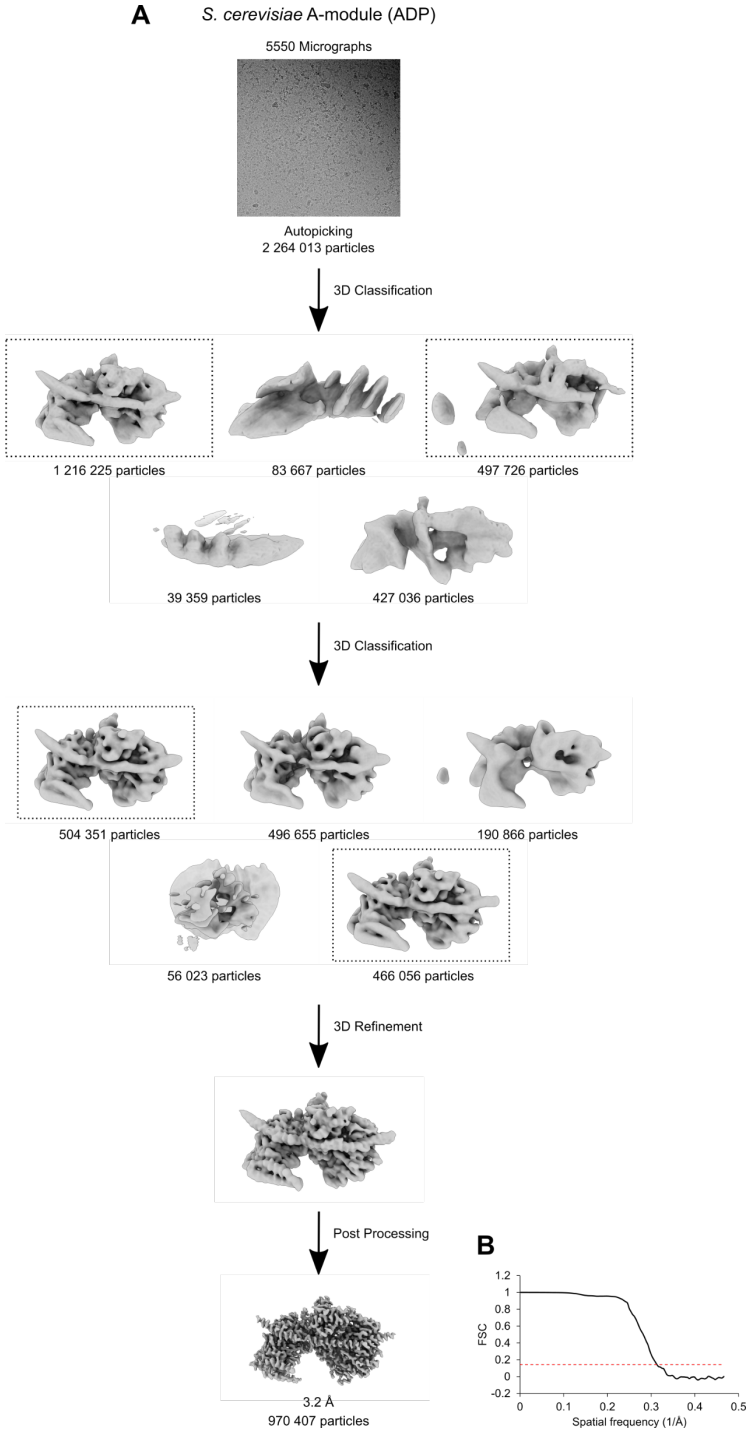


Fig. S11. Cryo-EM data analysis of *S. cerevisiae* A-module (ADP-state). (A) Cryo-EM data processing workflow of *S. cerevisiae* A-module using Relion-3.0 *(58)*. A representative micrograph and the cryo-EM data processing workflow are shown. (B) Gold-standard Fourier shell correlation (FSC) curve of the final A-module reconstruction. The red line indicates the 0.143 cutoff criterion.

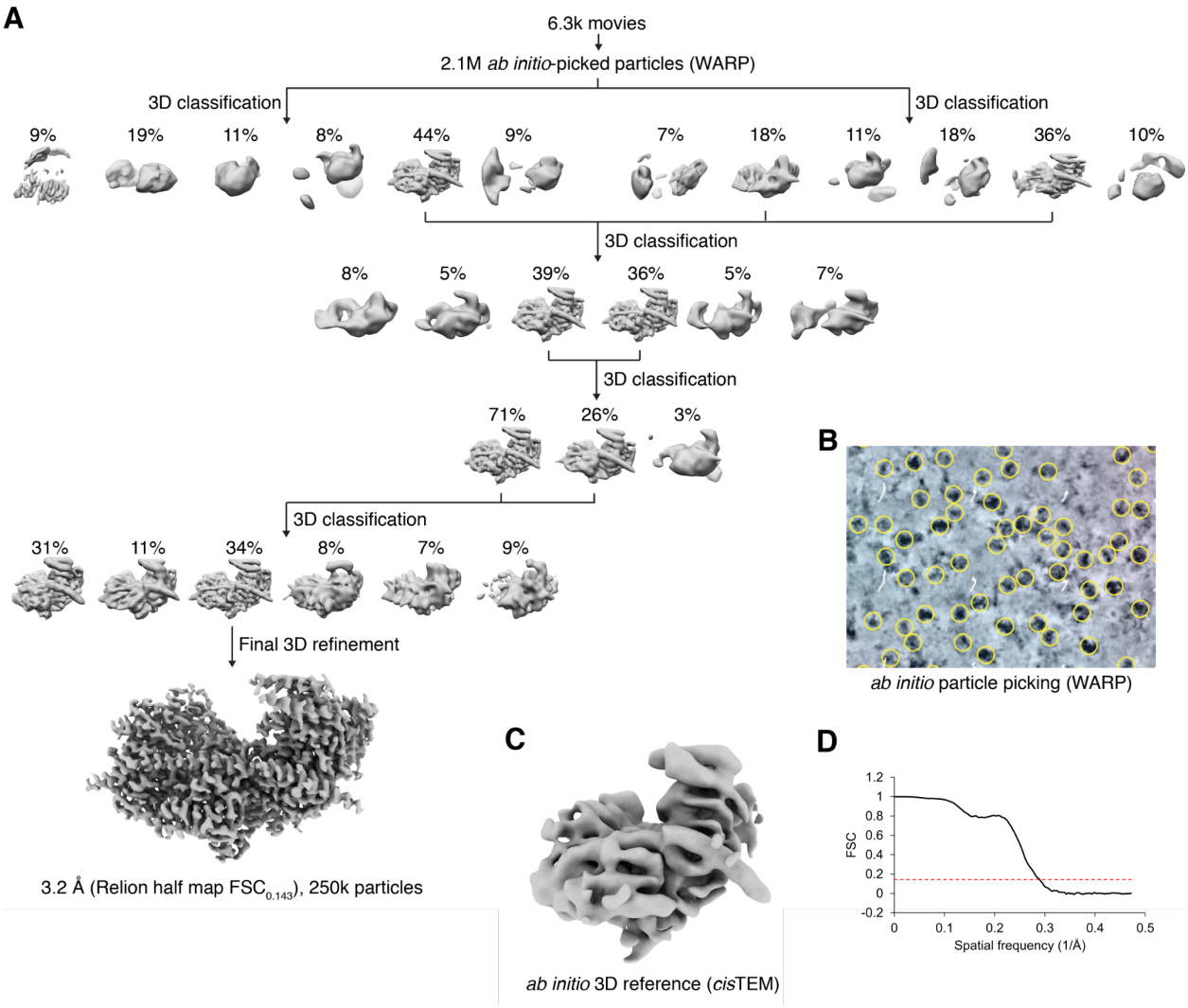


Fig. S12. Cryo-EM data analysis of *S. cerevisiae* A-module (ATPγS-state). (**A**) Good particles were selected through a series of focused 3D classifications and subsequently polished in Relion *(58)*. The final 3D reconstructions were generated and the resolution values calculated by Relion independent half map $FSC_{0.143}$ criterion. (**B**) The particles were picked *ab initio* and qualitatively filtered using WARP *(60)*. (**C**) The *ab initio* 3D model without DNA bound was generated in *cis*TEM *(61)* and used as the 3D reference for DNA-bound datasets to avoid bias in DNA presence and conformation. (**D**) Gold-standard Fourier shell correlation (FSC) curve of the final A-module reconstruction (no mask applied). The red line indicates the 0.143 cutoff criterion.

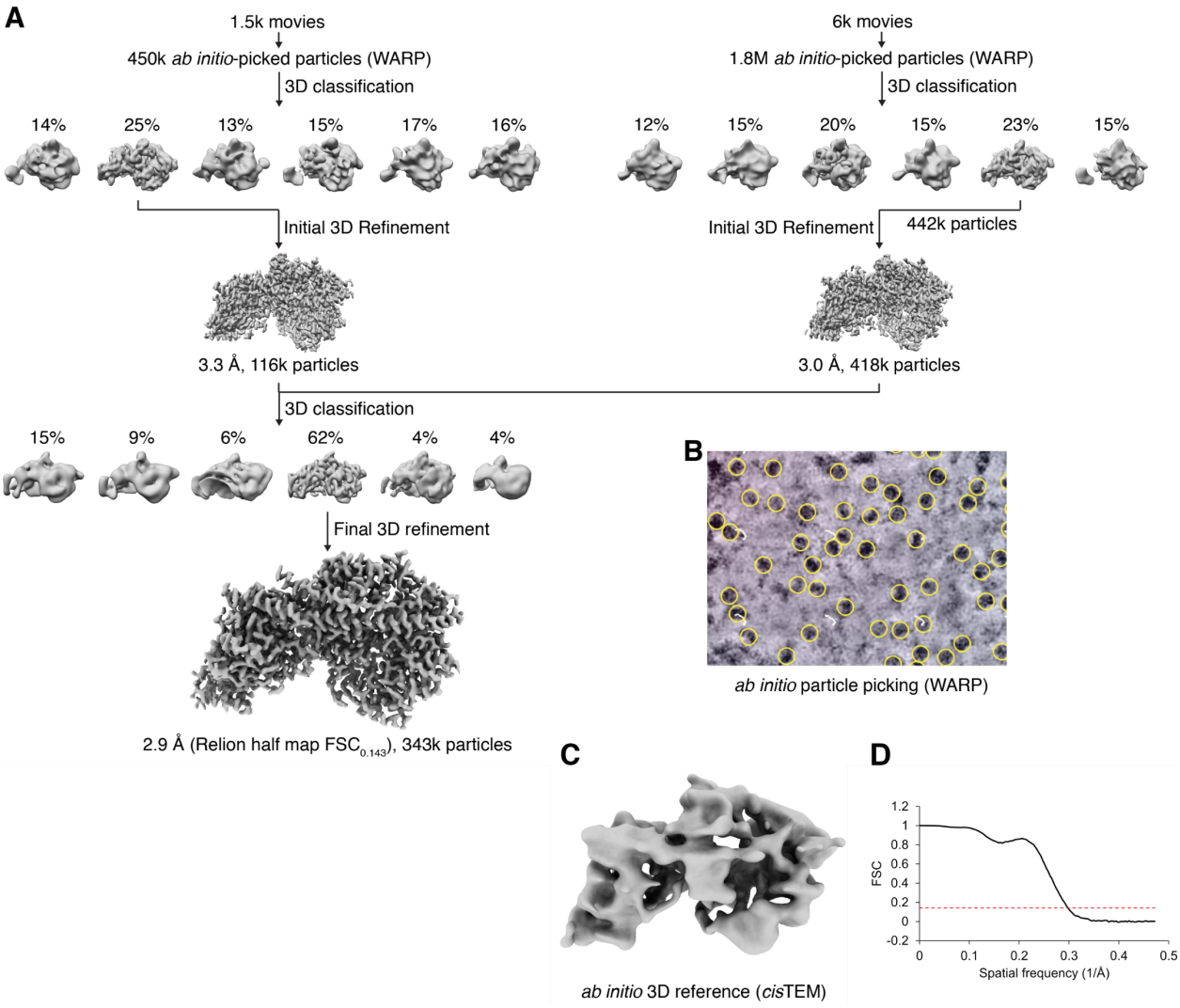


Fig. S13. Cryo-EM data analysis of *C. thermophilum* A-module (ATPγS-state). (**A**) Good particles were selected through a series of focused 3D classifications and subsequently polished in Relion *(58)*. The final 3D reconstructions were generated and the resolution values calculated by Relion independent half map $FSC_{0.143}$ criterion. (**B**) The particles were picked *ab initio* and qualitatively filtered using WARP *(60)*. (**C**) The *ab initio* 3D model without DNA bound was generated in *cis*TEM *(61)* and used as the 3D reference for DNA-bound datasets to avoid bias in DNA presence and conformation. (**D**) Gold-standard Fourier shell correlation (FSC) curve of the final A-module reconstruction (no mask applied). The red line indicates the 0.143 cutoff criterion.

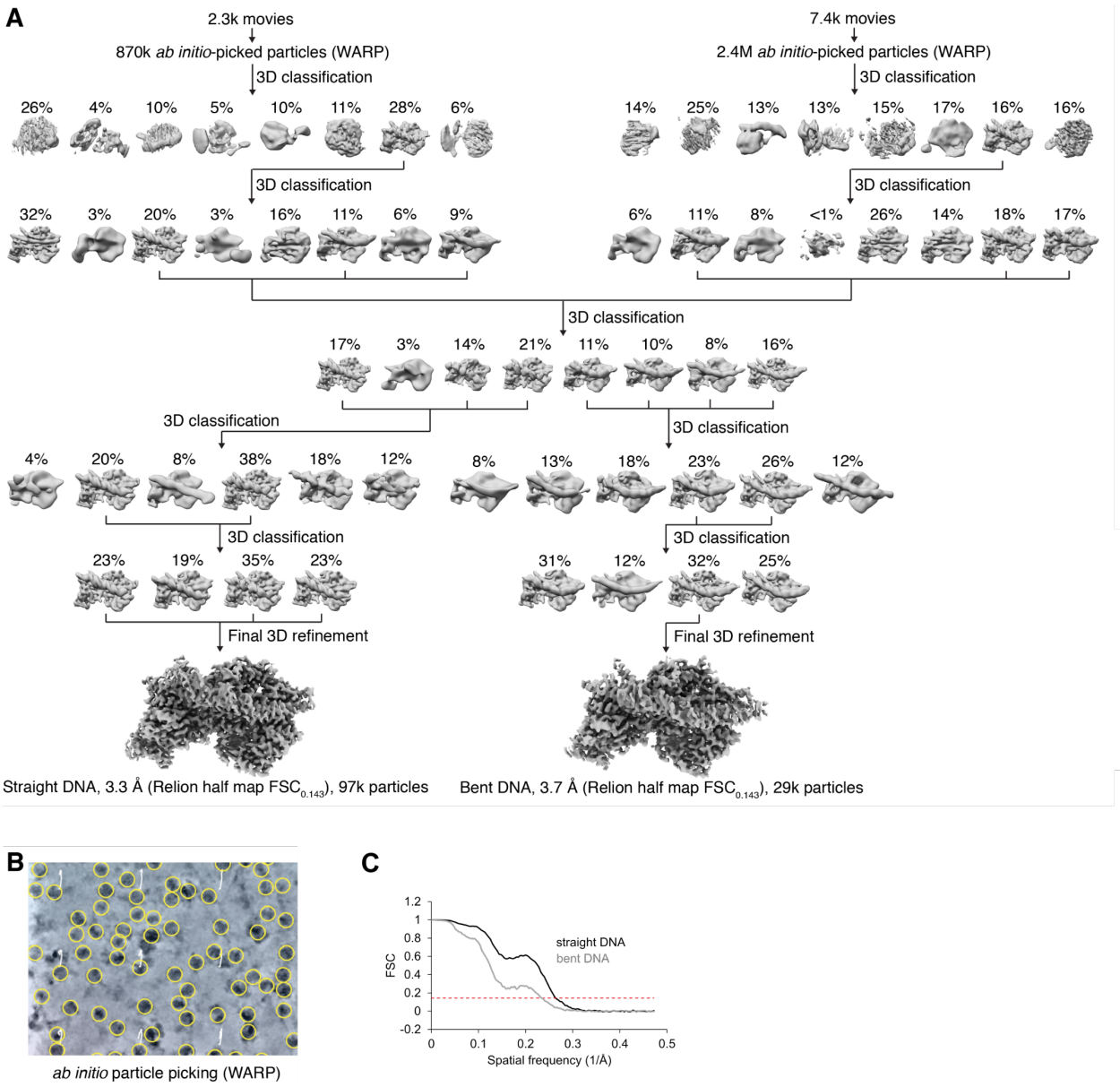


Fig. S14. Cryo-EM data analysis of *C. thermophilum* A-module and DNA (ADP·AlF$_x$). The *ab initio* 3D model without DNA bound was generated in *cis*TEM *(61)* and used to avoid bias in DNA presence and conformation. (**A**) DNA-bound A-module species were isolated from DNA-free species, followed by discernment and isolation of A-modules with straight and curved DNA bound through a series of focused 3D classifications in Relion *(58)*. The final 3D reconstructions were generated and the resolution values calculated by Relion independent half map FSC$_{0.143}$ criterion. (**B**) The particles were picked *ab initio* and qualitatively filtered using WARP *(60)*. (**C**) Gold-standard Fourier shell correlation (FSC) curves of the final A-module reconstructions (no mask applied). The red line indicates the 0.143 cutoff criterion.

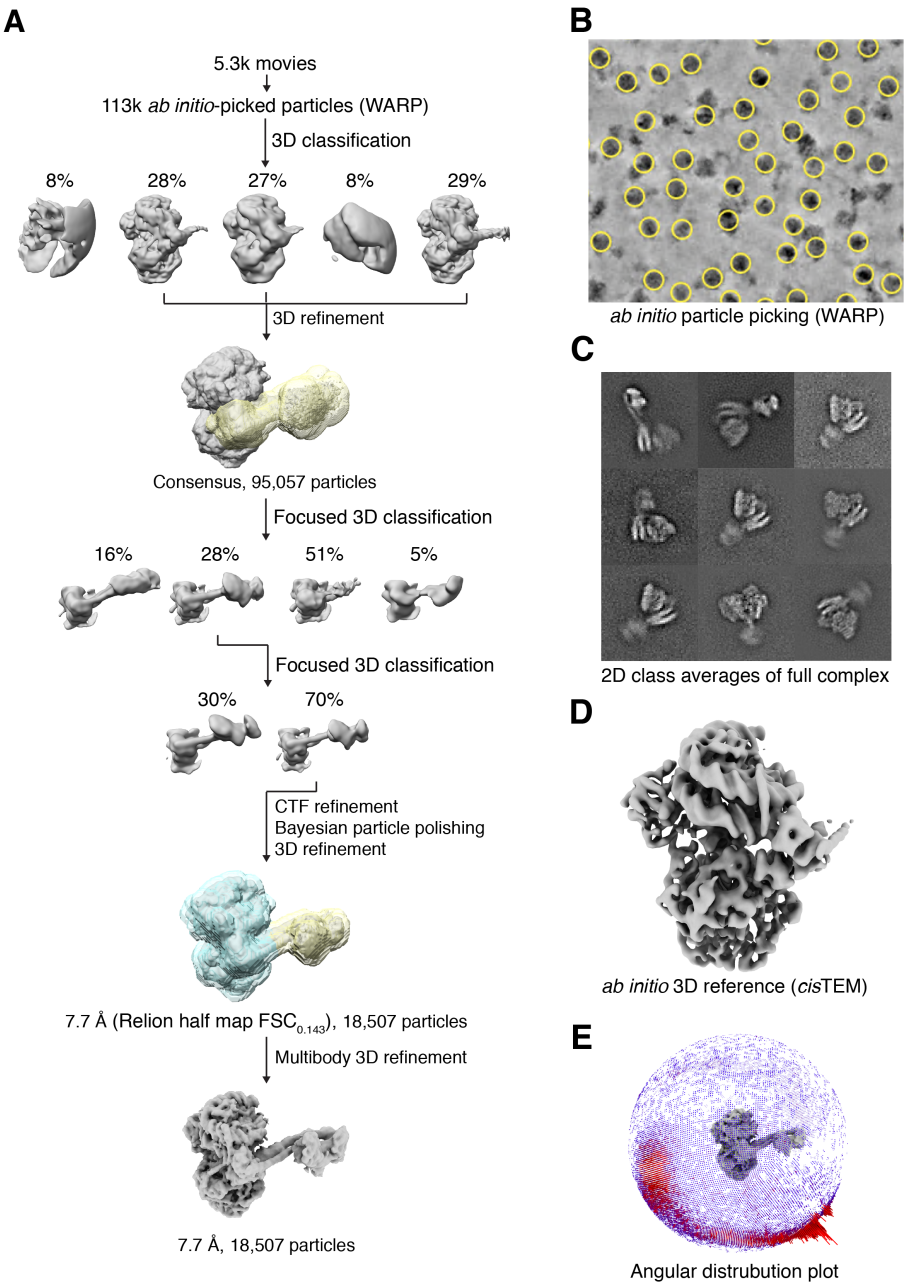


Fig. S15. Cryo-EM data analysis of *C. thermophilum* INO80$^{\Delta N}$. (**A**) A small subpopulation with the C- and A-modules more coherently placed were isolated through a series of focused 3D classifications in Relion and the final 3D reconstruction was generated by Relion multibody 3D refinement *(58)*. (**B**) The particles were picked *ab initio* and qualitatively filtered using WARP *(60)*. (**C**) 2D class averages of isolated full INO80 complex (**D**) The *ab initio* 3D model was generated in cisTEM *(61)*. (**E**) Angular distribution plot of the full INO80 complex 3D reconstruction.

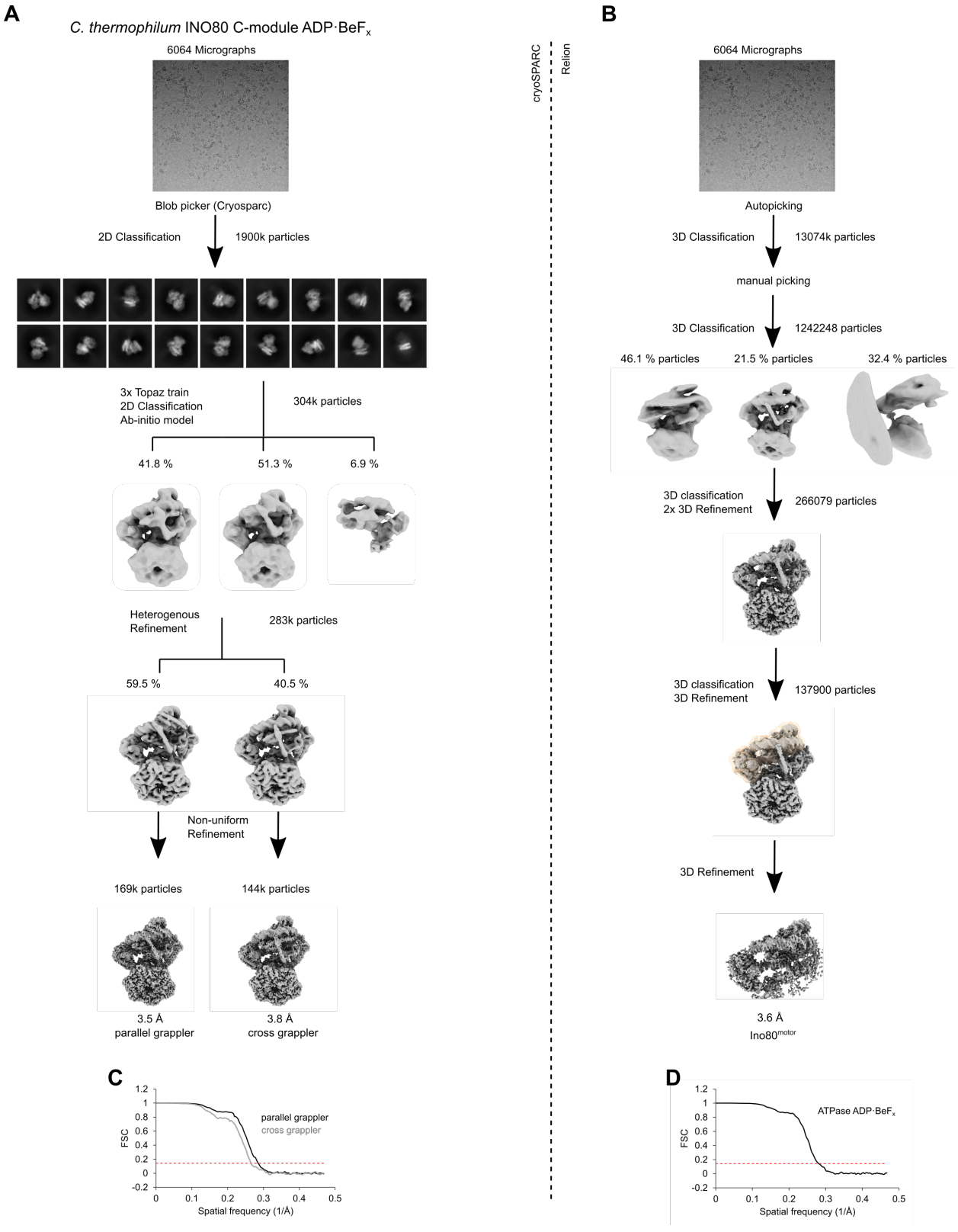

A
C. thermophilum INO80 C-module ADP·BeFx
6064 Micrographs
Blob picker (Cryosparc)
2D Classification
1900k particles
3x Topaz train
2D Classification
Ab-initio model
304k particles
41.8 %
51.3 %
6.9 %
Heterogenous
Refinement
283k particles
59.5 %
40.5 %
Non-uniform
Refinement
169k particles
144k particles
3.5 Å
parallel grappler
3.8 Å
cross grappler
C
FSC
Spatial frequency (1/Å)
parallel grappler
cross grappler
cryoSPARC
Relion
B
6064 Micrographs
Autopicking
3D Classification
13074k particles
manual picking
3D Classification
1242248 particles
46.1 % particles
21.5 % particles
32.4 % particles
3D classification
2x 3D Refinement
266079 particles
3D classification
3D Refinement
137900 particles
3D Refinement
3.6 Å
Ino80motor
D
FSC
Spatial frequency (1/Å)
ATPase ADP·BeFx

Fig. S16. Cryo-EM data analysis of *C. thermophilum* INO80 C-module with 0N80 nucleosomes. (**A**) Cryo-EM data processing workflow o*f C. thermophilum* INO80 C-module with ADP·BeF_x using cryoSPARC v3.2.0 *(57)*. Representative micrograph of INO80 and representative classes of a 2D classification of the particles used for the final INO80 C-module reconstruction. (**B**) Local Refinement of the $Ino80^{motor}$ bound to nucleosome (**C-D**) Gold-standard Fourier shell correlation (FSC) curves of the final (**C**) INO80 C-module bound to the nucleosome with Arp5 grappler in parallel and cross conformation and (**D**) $Ino80^{motor}$ bound to the nucleosome (ADP·BeF_x). The red line indicates the 0.143 cutoff criterion.

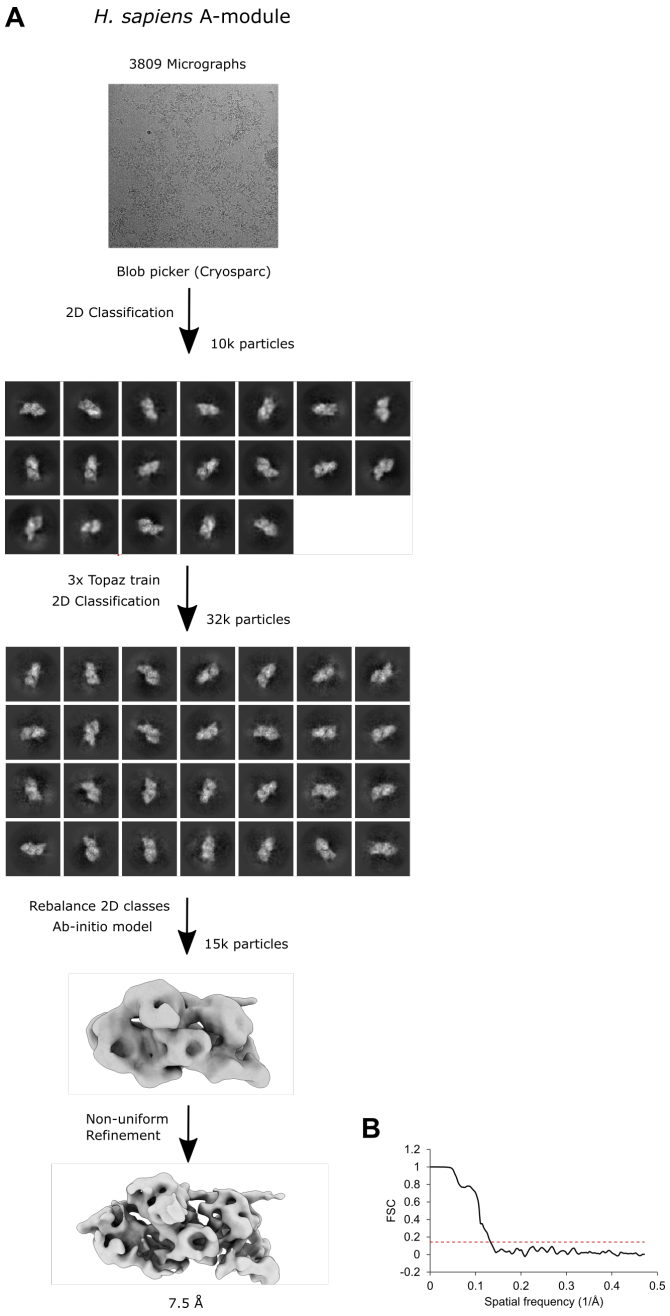


Fig. S17. Cryo-EM data analysis of *H. sapiens* A-module. (**A**) Cryo-EM data processing workflow of *Hs*A-module using cryoSPARC v3.2.0 *(57)*. Representative micrograph of *Hs*A-module and representative classes of a 2D classification of the particles used for the final A-module reconstruction after an *ab initio* model. (**B**) Gold-standard Fourier shell correlation (FSC) curve of the final A-module reconstruction. The red line indicates the 0.143 cut-off criterion

Table S2. *In vivo* yeast strains.

name	gene	mutation
HSA^{Q1}	INO80	R482Q, K483Q, K487Q, R490Q, R504Q, K505Q, R512Q, K515Q, K516Q
HSA^{Q2}	INO80	K523Q, R533Q, R534Q, R537Q, K544Q, K545Q, R551Q, K554Q, K555Q
HSA^{Q1+Q2}	INO80	R482Q, K483Q, K487Q, R490Q, R504Q, K505Q, R512Q, K515Q, K516Q, K523Q, R533Q, R534Q, R537Q, K544Q, K545Q, R551Q, K554Q, K555Q
HSA^{A2}	INO80	K523A, K526A, R533A, R534A, R537A, K544A, K545A, R548A
ino80ΔN	INO80	Delta of 1-492
ino80^{E842A}	INO80	E842A
arp8ΔN	ARP8	Delta of 1-197

Table S3. *In vivo* yeast mutants.

BP5148	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, URA3::Ylplac211-empty*
BP5150	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::INO80-2FLAG*
BP5155	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::ino80-HSA-Q1-2FLAG*
BP5159	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::ino80-HSA-Q2-2FLAG*
BP5415	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::ino80-HSA-A2-2FLAG*
BP5137	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, ∆ura3/URA3::ino80-HSA-Q1-2FLAG*
BP5138	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, ∆ura3/URA3::ino80-HSA-Q2-2FLAG*
BP5139	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, ∆ura3/URA3::ino80-HSA-Q1+Q2-2FLAG*
BP5140	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, ∆ura3/URA3::ino80-E842A-2FLAG*
BP5128	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, arp8::natNT2, URA3::2FLAG-ARP8*
BP5125	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ura3∆, arp8::natNT2*

BP5133	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, arp8::natNT2, URA3::2FLAG-arp8-aa1-197*Δ

6 Discussion

In this cumulative thesis, I summarize three publications about two Swi2/Snf2 family members, which – although belonging to the same protein family – differ markedly in their topology and biochemical activity. Two publications focus on Mot1, a single-subunit ATPase that acts as a transcriptional regulator by removal of the TATA-box binding protein (TBP) from its genic promoter site. The crystal structure of full-length nucleotide-free Mot1 (Chapter 5.1) together with five cryo-EM structures, each capturing a conformational snapshot of stepwise TBP disruption from DNA, can be combined into a molecular movie of this essential biological process (Chapter 5.2). Thus, these findings reveal insights into an unusual non-processive Swi2/Snf2 enzyme activity.

The third publication focuses on the INO80 A-module, presenting high-resolution cryo-EM structures from three species allowing for comparative structural analysis (Chapter 5.3).

Below, I discuss important findings from these publications, ultimately aiming at general and enzyme-specific concepts of Swi2/Snf2 ATPase regulation

6.1 Empty, nucleotide-free Mot1 resides in an autoinhibited resting state.

The Mot1 N-terminal domain (Mot1^{NTD}) with its characteristic 16 HEAT repeat arch is published in two conformations – as a post-dissociated product state with TBP bound and a latch occupying the TBP concave site and bound to TBP:DNA in complex with its cofactor NC2, both from *Enzephalozoon cuniculi* (*Ec*). However, it took some time until also the near full-length Mot1 including its C-terminal domain (Mot1^{CTD}) could be crystallized, from the fungus *Chaetomium thermophilum* (*Ct*). The C-terminal ATPase is in an open conformation, with the RecA2-like lobe 2 swiveled back, thereby separating the nucleotide binding motifs. A comparable state is seen for other Swi2/Snf2 structures e.g., Rad54[45] and Chd1[53]. It is believed that it represents a resting state and nucleotide and substrate-binding stabilize a closed ATPase conformation poised for ATP-binding and hydrolysis. Consistently, it was observed earlier that for Mot1 alone, deleting Mot1^{NTD} led to an increase of ATPase rate, thus Mot1^{NTD} inhibited the ATPase. On the other hand, binding of TBP stimulated the ATPase rate over 10-fold[145]. This can be explained by our structural data, showing that a spiral-like form of Mot1^{NTD} apparently stabilizes the misaligned open RecA-like lobes by embedding them in the open site of the HEAT repeat arch. Further, we showed that disruptive point mutations between lobe 2 and HEAT repeat 1 led to increased ATPase rates, apparently supporting a closed active ATPase conformation. However, the TBP:DNA dissociation activity (TBP remodeling) was not detectably enhanced but remained in the wildtype regime. It

appears that the autoinhibition prevents futile cycles of ATP-hydrolysis, being activated only in the presence of TBP or TBP:DNA. The outermost C-terminal 60 residues were almost completely omitted in the crystallization construct, but interestingly, a C-terminal deletion mutant (Mot1$^{\Delta 50C}$), showed an inhibition of TBP remodeling. This points to a regulatory role of the so-called C-terminal bridge element, analogous to the allosteric regulation of the bridge in some nucleosome remodelers[146,147]. The fact that the ATPase rate of Mot1 bridge-deletion mutant exceeded the rate of wildtype Mot1 accompanied by impaired TBP remodeling led to the conclusion that the bridge positively couples ATPase and remodeling activity. But how the bridge facilitates its regulatory role was still unclear at this point.

6.2 The Mot1 bridge acts as hook and anchor

We continued using *Ct*Mot1 due to its thermal and structural stability for cryo-EM studies in complex with TBP bound to TATA-box DNA. Capturing diverse nucleotide-dependent conformations, we arrived at two states that revealed Mot1^{CTD} bound to upstream DNA but with TBP dislodged from the TATA-box and bound to Mot1^{NTD}. In the Mot1 pre- and posthydrolysis states an extra density at the TBP concave site continues at the inner side of the HEAT-repeat ring. By crosslink-mass spectrometry and *de novo* model building of Mot1^{NTD} tracing the unassigned density (denoted “anchor”), we identified it as part of the C-terminal bridge that also occupies the DNA-binding site of TBP (denoted “hook”), taking on the role of the latch as seen in *Ec*Mot1. The hook competitively interferes with TBP:DNA rebinding and deletion of the hook strongly diminished TBP dissociation from DNA. N-terminal of the bridge, a Swi2/Snf2 characteristic brace folds from RecA-like lobe 2 back to lobe 1 and has been proposed as lobe-to-lobe communicator[31] and more recently identified as part of a regulatory hub in nucleosome remodelers[33]. Curiously, the brace extends into a turn followed by another shorter helix (denoted “wedge”, see below). Structurally, the brace-wedge formation resembles the pincer1-pincer2 arrangement of the distantly related viral RNA receptor RIG-I. The pincer domains extend into the C-terminal RD-domain that recognizes Rig-I’s preferred substrate, 5’-triphosphate dsRNA, therefore acting as a relay that regulates short-range translocation on dsRNA[41]. Although this regulatory mode could be seen as blueprint for Mot1 at first, when comparing pre- and posthydrolysis states, an extensive spatial rearrangement becomes visible.

6.3 A wedge element acts as a lever for a pivot movement of Mot1^{CTD}

We used the unhydrolyzable nucleotide analogue ADP-BeF_3^- to mimic the ATP ground state. Compared to the fully bound Mot1:TBP:DNA complex (substrate recognition state), the TBP has undergone an induced-fit-like engulfment by Mot1^{NTD} and the HEAT repeat array has transformed from its near spiral form of the autoinhibited resting state into an ordered ring, closed between HEAT repeat (HR) 1 and 16. This ring closure releases Mot1^{CTD}, which remains flexibly attached to an extension of HR 16. The resulting flexibility is leveraged by the wedge element following the brace, that is in proximity to Mot1^{NTD} in the prehydrolysis state. Upon ATP-hydrolysis, the wedge protrudes between HR 16 and lobe 1. This correlates with a 30° pivot movement of Mot1^{CTD} relative to Mot1^{NTD} and corresponding bending of upstream DNA bound by the ATPase. The DNA bending is accompanied by minor groove widening by a conserved phenylalanine at RecA2B. This gripping and bending of upstream DNA correlates with a 1-2 base pair rotational DNA translocation. Remarkably, while the TATA-motif still faces the dislodged TBP in the prehydrolysis state, it becomes rotated away by the short-range translocation step, making a TBP-TATA rebinding even more unlikely. Another ATP binding event – here we used ATPγS as an ATP-analogue – results in the Mot1 product state after complete TBP:DNA disruption.

Our structure-based results of TBP remodeling by Mot1 agree with biochemical and biophysical studies investigating the chemo-mechanical basis of this process[104,105] and offer a specialized and remarkable example of unsuspected Swi2/Snf2 enzymatic activity. Mot1 employs an intricate mechanism, not just translocating DNA, but also bending it and competitively occluding its former TBP binding site.

6.4 Dimerization of Mot1:TBP:DNA complexes – physiologically relevant or artefact?

For all nucleotide- and DNA-bound Mot1:TBP complexes we observed a small fraction of homo-dimerized complexes besides the expected Mot1:TBP:DNA monomers in size-exclusion chromatography (SEC) and in cryo-EM. Remarkably, the bridge-deletion Mot1 complex appeared predominantly as a stable dimer and could be reconstructed at high-resolution. Since its first recombinant purification, it has been a matter of discussion if Mot1 appears in oligomerized form, as it was eluting from a size-exclusion column with double the size as expected for the monomer. However, sucrose-gradient sedimentation and *in vivo* expression of differently tagged Mot1 still confirmed the monomeric state[145]. It was only years later, that two groups put the dimerization of Mot1 back on the agenda. Their use of flow cell FRET assays enabled a single-molecule perspective of Mot1:TBP:DNA complex formation and nucleotide-dependent

dissociation. After washing out excessive Mot1 in solution after complex formation, the dissociation reaction proved to be surprisingly inefficient, leading to the conclusion that Mot1 in solution is needed for efficient TBP:DNA disruption[105,148]. First, based on the highly resolved dimer structure, we introduced perturbatory mutations at the three dimer-dimer interfaces to disrupt Mot1 dimer formation. However, these mutated Mot1 complexes did not show a significantly reduced TBP displacement from DNA compared to wildtype Mot1 in our native gel shift assays (data not shown). Thus, if a dimerization influencing catalysis occurs at all, it is below the detection limit of our ensemble measurement approach. Second, in our structure dimerization occurs between Mot1:TBP:DNA complexes in nucleotide-bound states, and not as proposed in the aforementioned publications between a nucleotide-bound Mot1:TBP:DNA complex and a nucleotide-free Mot1. How this should increase remodeling activity remains obscure considering the structural arrangement of the dimer, although a potential *in vivo* role at e.g., in topologically-associated domain (TADs) proximal promoters cannot be excluded. Moreover, it is conceivable that certain chemical conditions such as pH and ionic strength could support dimer formation and imply that the observed dimerized complexes may be an artifact of our *in vitro* reconstitution.

But how can the apparent necessity of additional Mot1 in solution be explained? It is not surprising that in a reductive *in vitro* setup Mot1 is not as efficient as in a complex *in vivo* context. As suggested, one or more catalytic events, most probably by different Mot1 molecules might be needed for an efficient TBP displacement, and the likelihood of a disruptive event increases with the concentration of Mot1 in solution. It remains speculative that Mot1 dimerization or even oligomerization underlies efficient TBP remodeling, because measuring direct Mot1:Mot1 interactions was beyond the scope of each experimental setup so far. Further, in our structural work, the TBP remodeling process could be captured as a series of monomeric Mot1:TBP:DNA complexes in nucleotide-dependent conformations, which explains stepwise TBP displacement in unprecedent structural detail – without apparent necessity of dimerization.

6.5 Non-processive TBP remodeling by Mot1

Based on earlier DNA footprinting assays[103] and our structural studies, we can confidently propose a non-processive activity of the Mot1 ATPase. This implies that a short-range DNA translocation (1-2 base pairs) would suffice to disrupt the TBP:DNA complex instead of a processive long-range translocation. This translocation mode has also been suggested for related superfamily 2 members such as RLRs (RIG-I like receptors) and DEAD box helicases. To example, short-range translocation by RIG-I on the viral dsRNA substrate might act as a proofreading

mechanism recognizing the preferred short RNA duplexes[41]. But the closest homology to Mot1 enzymatic activity can be derived from its Swi2/Snf2 relative, the CSB ATPase. CSB shifts polymerase II upon sensing a bulky DNA adduct one base pair over the lesion site. Remarkably, it uses an intercalating phenylalanine in a lever-like motion to "push" the complex along the DNA[47,48], which is reminiscent of the minor groove intercalating Mot1 phenylalanine.

How does this finding translate to nucleosome remodelers? Although the twist-diffusion model for DNA translocation underlying nucleosome sliding has been established for yeast Snf2[149], there remains a lack of understanding, how remodelers like INO80 positions a nucleosome in longer, approximately 10 - 20 bp steps. In the "rotor-stator-grip" model, continuous pumping of entry DNA accumulates as a DNA bulge against the "counter-grip" subunits Arp5 and Ies6, whose release manifests as a long-range DNA shift around the nucleosome core[62,67]. The A-module that binds extranucleosomal DNA would act as a ratchet that prevents back-slippage of the entry DNA during bulge formation. Despite this intriguing model assigning a function to each subunit, recent findings must be continuously integrated, especially new data showing a more intricate role of the A-module in the allosteric regulation of the remodeling process. This includes the role of DNA shape[142], the regulatory hub first proposed for the RSC complex ATPase [33] and our novel structural insights presented in Chapter 5.3.

6.6 The INO80 A-module dynamically interacts with extranucleosomal DNA

While the last years showed a remarkable progress in structural elucidation of large, multisubunit chromatin remodelers bound to their nucleosomal substrate, the intricate factors that influence and regulate the remodeling reaction are now beginning to be discovered. In genomic research, the focus for DNA-binding proteins relied on classic DNA sequence motifs, that could be identified via sequence alignments and evolutionary analysis. Widening the research scope for INO80 in order to integrate the genomic context, it became clear that DNA shape features intrinsic to certain sequences are one determinant of strong nucleosome positioning besides barrier factors, nucleosome density and histone modifications and variants.

Still, how remodelers act in a complex genomic environment with different DNA sequences, neighboring nucleosomes and protein obstacles is far from completely understood. The A-module comprising Arp8, N-actin and Arp4 bound to the HSA-helix, forming the middle part of the elongated motor ATPase N-terminus, is an essential allosteric regulator of remodeling in response to extranucleosomal DNA[125,138,142] and known to act as "molecular ruler" in conjunction with the N-module, at least in yeast[63,141]. In Chapter 5.3, we build on previous work on the structure of the yeast A-module[125] and cryo-EM structures of the INO80 C-module[62], aiming at a

contextual structural analysis of INO80 C- and A-modules, nucleosome bound and with an entry DNA overhang. High-resolution structures of *S. cerevisiae* and *C. thermophilum* A-modules alone and bound to DNA gained a more detailed and species-dependent characterization of the A-module. The proposed two-sided communication between A-module and ATPase could be underpinned structurally. The nucleotide state of the ATPase influenced the orientation of the HSA/post-HSA helix and extranucleosomal DNA bendability regulated coupling between ATPase activity and nucleosome sliding. The biggest impact on DNA translocation was observed upon insertion of relatively stiff DNA at the ATPase, which clearly reduced remodeling efficiency. These findings extend the functional biochemistry on DNA shape and regulatory interactions between ATPase and A-module[142] and undermine it with a chemo-mechanical basis. This shows, how integration of biochemical and structural data and – more specific – the semantic mapping of functional data to structures enables a deeper understanding of a complex cellular process.

Also, we found a plausible explanation for the location of INO80 at superhelical location (SHL) -6 at the nucleosome instead of SHL 2 for most chromatin remodelers. The Arp5 in a "cross-grappler" conformation contacts the entry DNA and seems to ensure productive feeding of the DNA into the ATPase. The unusual nucleosome position of INO80 near the dyad enables monitoring of entry DNA shape features, which could conduct a conformational counteraction of the A-module and the motor ATPase itself. This multifactorial sensibility towards extranucleosomal DNA features might be the reason that INO80 is the only chromatin remodeling complex, that solely can establish ordered *in vitro* nucleosome arrays on genome-wide libraries comparable to native yeast arrays[121].

6.7 Client proteins anchor at the A-module via a conserved 2W-domain

It is known that A-modules feature additional subunits over the defining N-actin and Arps. In fungal INO80, Ies4 is part of the A-module[139] and for human INO80, the abundant transcription factor YY1 was early identified as a constitutive component of the A-module[140,150]. These client proteins likely support integrity of the complex, as is the case for Rtt102 in the RSC A-module[135], or are involved in recruitment to genomic sites. YY1, with its DNA-binding zinc finger domains recruits INO80 to promoter regions[150], but the structural basis for the YY1:A-module interaction remained unsolved. A first hint was the structure of the RSC complex A-module, which also features Rtt102 as an additional subunit[135]. Curiously, Rtt102 resides on a ridge tracing along Arp7 and Arp9 and featuring a β-hairpin. We noticed this β-hairpin motif also in Ies4 and YY1 and via an alignment with homologs from different species we identified two sequentially and structurally conserved tryptophan residues, which seem to accommodate on a miniaturized

version of the abundant WW-domain[151–153]. For Ies4 and YY1, we observed this double-tryptophan motif, denoted "2W motif", which forms a hydrophobic pocket that incorporates a proline residue – consistent with the binding mode already seen for Rtt102. Apparently, the 2W β-hairpin emerged as a common protein anchor, that allows versatile binding to different complexes. It's established role in anchoring of client proteins to A-modules was only recently confirmed by the structure of the acetyltransferase NuA4, where it facilitates Epl1 association with Arp4:Actin[154]. Also, for example, YY1 uses the same motif to bind to the polycomb group protein 4MBTD1[155,156]. Impressively, the 2W β-hairpin exemplifies how versatility of a relatively simple structural motif enables divergent biological function dependent on the binding partner.

6.8 Deep learning algorithms accelerate structural biology

Only recently, a small revolution in the determination process of protein structures took place with the publication of the deep-learning based structure prediction software Alphafold2[157] and RosettaFold[158]. Consequently, the availability of all known protein structures listed in a database already changes the interpretation of experimentally gained protein density maps. This progress is illustrated in Chapter 5.3, where we had to rely on a low-resolved human A-module density and still arrived at a plausible structural model. First of all, it allowed unambiguous fitting of the independently reconstructed subunits Arp4:Actin[159] and Arp8[160]. No experimental structure existed for the human HSA domain but relying on structural training data in the Protein Data Bank (PDB), a highly confident human HSA structure could be calculated by Alphafold2. Building in the YY1 β-hairpin manually would have been impossible due to the low resolution, but after an AlphaFold2 calculation of human Arp4:Actin in complex with the respective YY1 region, we arrived at a model where the YY1:A-module interaction was almost homologous to the Ies4 and Rtt102 binding mode. Also, deducing prior knowledge from high-resolution experimental structures of fungal A-modules enables a full-featured comparison with human A-module subunits. For example, all Arp8 subunits from the three species comprise an elongated N-terminus, which in the fungal Arp8 aligns with the DNA, comprising an "AT-hook" adding to the HSA domain as a DNA interactor, as shown in Chapter 5.3. Human Arp8 has a much shorter N-terminus, but instead a larger intermitted helix-loop-helix apical domain than the fungal homologues.

The function of this extended region is still unknown, but it could be another DNA-binding component or an eventual binding platform for interactors. These comparisons display the species-dependent variety in the structures of those evolutionary conserved protein building block (Fig. 8.8). This demonstrates that besides high-resolution structures from different contexts

and conformational space, extended parts of the A-module subunits itself still need to be elucidated.

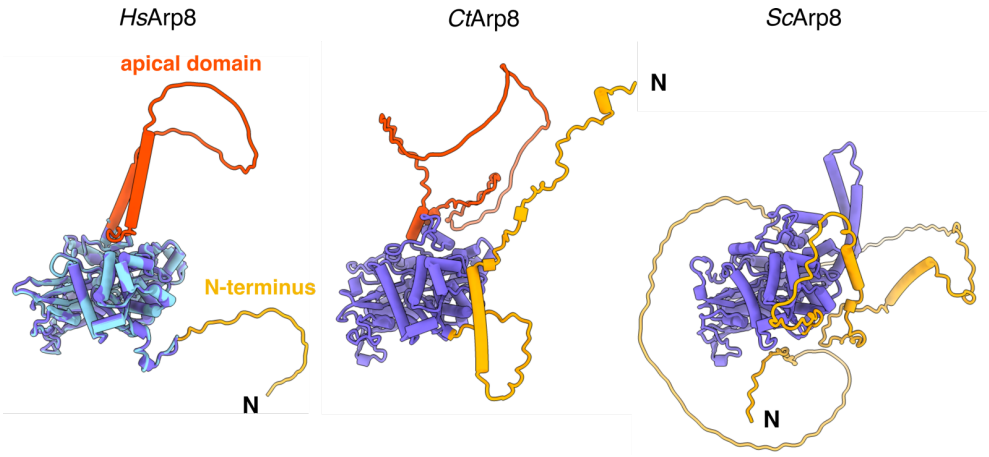


Fig. 6.8: Structural comparison of Arp8 from three species as predicted by Alphafold2.
From left to right: *Homo sapiens* (*Hs*) Arp8 structure prediction (violet) aligned with the Arp8 crystal structure (PDB 4FO0) (light blue), *Chatomium thermophilum* (*Ct*) and *Saccharomyces cerevisiae* (*Sc*) Arp8 structure predictions.

These examples already show that deep learning algorithms are incorporated as a matter of course in the structural prediction process and in the future will be much more used, also for the design of expression constructs and to foresee the structural impact of mutants. It is only a matter of time, until multimeric protein complexes will be reliably predicted, which enable structural biologists to go on and widen the scope, to characterize molecular machines in the context of the whole cell by cryogenic electron tomography, for example.

References

1. Singleton, M. R. & Wigley, D. B. Modularity and specialization in superfamily 1 and 2 helicases. *Journal of Bacteriology* vol. 184 1819–1826 Preprint at https://doi.org/10.1128/JB.184.7.1819-1826.2002 (2002).

2. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75–84 (1999).

3. Mcglynn, P., Mahdi, A. A. & Lloyd, R. G. Characterisation of the catalytically active form of RecG helicase. *Nucleic Acids Research* vol. 28 (2000).

4. Lee, J. Y. & Yang, W. UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke. *Cell* **127**, 1349–1360 (2006).

5. Gu, M. & Rice, C. M. Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proceedings of the National Academy of Sciences* (2010) doi:10.1073/pnas.0913380107.

6. Singleton, M. R., Dillingham, M. S. & Wigley, D. B. Structure and mechanism of helicases and nucleic acid translocases. *Annual Review of Biochemistry* vol. 76 23–50 Preprint at https://doi.org/10.1146/annurev.biochem.76.052305.115300 (2007).

7. Sedman, J. & Stenlund, A. The Papillomavirus E1 Protein Forms a DNA-Dependent Hexameric Complex with ATPase and DNA Helicase Activities. *J Virol* **72**, 6893–6897 (1998).

8. Ma, J. B. *et al.* Dynamic structural insights into the molecular mechanism of DNA unwinding by the bacteriophage T7 helicase. *Nucleic Acids Res* **48**, 3156–3164 (2020).

9. Steinmetz, E. J. & Platt, T. Evidence supporting a tethered tracking model for helicase activity of Escherichia coli Rho factor. *Proceedings of the National Academy of Sciences* **91**, 1401–1405 (1994).

10. Li, N. *et al.* Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).

11. Fairman-Williams, M. E., Guenther, U. P. & Jankowsky, E. SF1 and SF2 helicases: Family matters. *Curr Opin Struct Biol* **20**, 313–324 (2010).

12. Duerr, H., Flaus, A., Owen-Hughes, T. & Hopfner, K. P. Snf2 family ATPases and DExx box helicases: Differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res* **34**, 4160–4167 (2006).

13. Flaus, A. & Owen-Hughes, T. Mechanisms for ATP-dependent chromatin remodelling: The means to the end. doi:10.1111/j.1742-4658.2011.08281.x.

14. He, Y., Andersen, G. R. & Nielsen, K. H. Structural basis for the function of DEAH helicases. *EMBO Rep* **11**, 180–186 (2010).

15. Chen, M. C. & Ferré-D'Amaré, A. R. Structural Basis of DEAH/RHA Helicase Activity. *Crystals (Basel)* **7**, 253 (2017).

16. Linder, P. & Jankowsky, E. From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol* **12**, 505–516 (2011).

17. Mcglynn, P., Mahdi, A. A. & Lloyd, R. G. Characterisation of the catalytically active form of RecG helicase. **28**, 2324–2332 (2000).

18. Loenen, W. A. M., Dryden, D. T. F., Raleigh, E. A. & Wilson, G. G. Type I restriction enzymes and their relatives. *Nucleic Acids Res* **42**, 20–44 (2014).

19. Butterer, A. *et al.* Type III restriction endonucleases are heterotrimeric: Comprising one helicase-nuclease subunit and a dimeric methyltransferase that binds only one specific DNA. *Nucleic Acids Res* **42**, 5139–5150 (2014).

20. Ye, J., Osborne, A. R., Groll, M. & Rapoport, T. A. RecA-like motor ATPases - Lessons from structures. *Biochim Biophys Acta Bioenerg* **1659**, 1–18 (2004).

21. Neigeborn, L. & Carlson, M. Genes affecting the regulation of SUC2 gene expression by glucose repression in saccharomyces cerevisiae. *Genetics* **108**, 845–858 (1984).

22. Stern, M., Jensen, R. & Herskowitz, I. Five SWI genes are required for expression of the HO gene in yeast. *J. Mol. Biol.* **178**, 853–868 (1984).

23. Richmond, E. & Peterson, C. L. *Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2*. *Nucleic Acids Research* vol. 24 (1996).

24. Laurent, B. C., Yang, X. & Carlson, M. An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family. *Mol Cell Biol* **12**, 1893–1902 (1992).

25. Eisen, J. A., Sweder, K. S. & Hanawalt, P. C. Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res* **23**, 2715–2723 (1995).

26. Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* **34**, 2887–2905 (2006).

27. Hoelper, D., Huang, H., Jain, A. Y., Patel, D. J. & Lewis, P. W. Structural and mechanistic insights into ATRX-dependent and -independent functions of the histone chaperone DAXX. *Nat Commun* **8**, (2017).

28. Goyal, N. *et al.* RAD54 N-terminal domain is a DNA sensor that couples ATP hydrolysis with branch migration of Holliday junctions. *Nat Commun* **9**, 1–10 (2018).

29. Shen, M. *et al.* Structural basis for the multi-activity factor Rad5 in replication stress tolerance. *Nat Commun* **12**, (2021).

30. Willhoft, O. *et al.* Structure and dynamics of the yeast SWR1 : nucleosome complex. **199**, (2018).

31. Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* (2017) doi:10.1038/nature22036.

32. Baker, R. W. *et al.* Structural insights into assembly and function of the RSC chromatin remodeling complex. *Nat Struct Mol Biol* **28**, 71–80 (2021).

33. Clapier, C. R., Verma, N., Parnell, T. J. & Cairns, B. R. Cancer-Associated Gain-of-Function Mutations Activate a SWI/SNF-Family Regulatory Hub. *Mol Cell* 1–14 (2020) doi:10.1016/j.molcel.2020.09.024.

34. Farnung, L., Vos, S. M., Wigge, C. & Cramer, P. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* (2017) doi:10.1038/nature24046.

35. McKnight, J. N., Jenkins, K. R., Nodelman, I. M., Escobar, T. & Bowman, G. D. Extranucleosomal DNA Binding Directs Nucleosome Sliding by Chd1. *Mol Cell Biol* **31**, 4746–4759 (2011).

36. Clapier, C. R. & Cairns, B. R. Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature* **492**, 280–284 (2012).

37. Mueller-Planitz, F., Klinker, H., Ludwigsen, J. & Becker, P. B. The ATPase domain of ISWI is an autonomous nucleosome remodeling machine. *Nat Struct Mol Biol* **20**, 82–89 (2013).

38. Rehwinkel, J. & Gack, M. U. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nat Rev Immunol* **20**, 537–551 (2020).

39. Kowalinski, E. *et al.* Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* **147**, 423–435 (2011).

40. Luo, D. *et al.* Structural insights into RNA recognition by RIG-I. *Cell* **147**, 409–422 (2011).

41. Schmidt, A., Rothenfusser, S. & Hopfner, K. P. Sensing of viral nucleic acids by RIG-I: From translocation to translation. *Eur J Cell Biol* **91**, 78–85 (2012).

42. Rawling, D. C., Kohlway, A. S., Luo, D., Ding, S. C. & Pyle, A. M. The RIG-I ATPase core has evolved a functional requirement for allosteric stabilization by the Pincer domain. *Nucleic Acids Res* **42**, 11601–11611 (2014).

43. Hopfner, K. P., Gerhold, C. B., Lakomek, K. & Wollmann, P. Swi2/Snf2 remodelers: Hybrid views on hybrid molecular machines. *Current Opinion in Structural Biology* Preprint at https://doi.org/10.1016/j.sbi.2012.02.007 (2012).

44. Lewis, R., Dürr, H., Hopfner, K. P. & Michaelis, J. Conformational changes of a Swi2/Snf2 ATPase during its mechano-chemical cycle. *Nucleic Acids Res* **36**, 1881–1890 (2008).

45. Dürr, H., Körner, C., Müller, M., Hickmann, V. & Hopfner, K. P. X-Ray structures of the sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell* **121**, 363–373 (2005).

46. Thomä, N. H. *et al.* Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nat Struct Mol Biol* **12**, 350–356 (2005).

47. Kokic, G., Wagner, F. R., Chernev, A., Urlaub, H. & Cramer, P. Structural basis of human transcription–DNA repair coupling. *Nature* **598**, 368–372 (2021).

48. Yan, C. *et al.* Mechanism of Rad26-assisted rescue of stalled RNA polymerase II in transcription-coupled repair. *Nat Commun* **12**, 1–12 (2021).

49. van der Weegen, Y. *et al.* The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nat Commun* **11**, (2020).

50. Nathan Gamarra, Stephanie L Johnson, Michael J Trnka, Alma L Burlingame, G. J. N. The nucleosomal acidic patch relieves auto-inhibition by the ISWI remodeler SNF2h. *Elife*.

51. Farnung, L., Vos, S. M., Wigge, C. & Cramer, P. Nucleosome–Chd1 structure and implications for chromatin remodelling. *Nature 2017 550:7677* **550**, 539–542 (2017).

52. Farnung, L., Ochmann, M., Engeholm, M. & Cramer, P. Structural basis of nucleosome transcription mediated by Chd1 and FACT. *Nat Struct Mol Biol* **28**, 382–387 (2021).

53. Hauk, G., McKnight, J. N., Nodelman, I. M. & Bowman, G. D. The Chromodomains of the Chd1 Chromatin Remodeler Regulate DNA Access to the ATPase Motor. *Mol Cell* **39**, 711–723 (2010).

54. Sundaramoorthy, R. *et al.* Structure of the chromatin remodelling enzyme Chd1 bound to a ubiquitinylated nucleosome. *Elife* (2018) doi:10.7554/eLife.35720.001.

55. Nodelman, I. M. *et al.* Nucleosome recognition and DNA distortion by the Chd1 remodeler in a nucleotide-free state. **29**, (2022).

56. Chakrabarti, S., Jarzynski, C. & Thirumalai, D. Processivity, Velocity, and Universal Characteristics of Nucleic Acid Unwinding by Helicases. *Biophys J* **117**, 867–879 (2019).

57. Mueller-Planitz, F., Klinker, H. & Becker, P. B. Nucleosome sliding mechanisms: New twists in a looped history. *Nat Struct Mol Biol* **20**, 1026–1032 (2013).

58. Luger, K., Mä Der, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. *Crystal structure of the nucleosome core particle at 2.8 A ˚ resolution*. *Nature © Macmillan Publishers Ltd* vol. 389 (1997).

59. Peng, Y., Li, S., Landsman, D. & Panchenko, A. R. Histone tails as signaling antennas of chromatin. *Curr Opin Struct Biol* **67**, 153-<60 (2021).

60. Cutter, A. R. & Hayes, J. J. A brief review of nucleosome structure. *FEBS Lett* **589**, 2914–2922 (2015).

61. Liu, G., Xing, Y., Zhao, H., Cai, L. & Wang, J. The implication of DNA bending energy for nucleosome positioning and sliding. *Sci Rep* 1–12 (2018) doi:10.1038/s41598-018-27247-x.

62. Eustermann, S. *et al.* Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* (2018) doi:10.1038/s41586-018-0029-y.

63. Zhou, C. Y. *et al.* The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding. *Mol Cell* **69**, 677-688.e9 (2018).

64. Li, M. *et al.* Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature 2019 567:7748* **567**, 409–413 (2019).

65. Winger, J., Nodelman, I. M., Levendosky, R. F., Bowman, G. D. & Jenkins, T. C. A twist defect mechanism for ATP-dependent translocation of nucleosomal DNA. (2018) doi:10.7554/eLife.34100.001.

66. Li, M. *et al.* Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature* 1 (2019) doi:10.1038/s41586-019-1029-2.

67. Jungblut, A., Hopfner, K. P. & Eustermann, S. Megadalton chromatin remodelers: common principles for versatile functions. *Curr Opin Struct Biol* **64**, 134–144 (2020).

68. Huertas, J. & Cojocaru, V. Breaths, Twists, and Turns of Atomistic Nucleosomes. *Journal of Molecular Biology* vol. 433 Preprint at https://doi.org/10.1016/j.jmb.2020.166744 (2021).

69. Brandani, G. B., Niina, T., Tan, C. & Takada, S. DNA sliding in nucleosomes via twist defect propagation revealed by molecular simulations. *Nucleic Acids Res* **46**, 2788–2801 (2018).

70. Yan, L. & Chen, Z. A Unifying Mechanism of DNA Translocation Underlying Chromatin Remodeling. *Trends Biochem Sci* **45**, 217–227 (2020).

71. Karnitz, L., Morrison, M. & Young, E. T. Identification and characterization of three genes that affect expression of ADH2 in Saccharomyces cerevisiae. *Genetics* **132**, 351–359 (1992).

72. Davis, J. L., Kunisawa, R. & Thorner, J. A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast Saccharomyces cerevisiae. *Mol Cell Biol* **12**, 1879–1892 (1992).

73. Buratowski, S., Hahn, S., Sharp, P. A. & Guarente, L. *Function of a yeast TATA element-binding protein in a mammalian transcription system*. (1988).

74. van Dyke, M. W., Roeder, R. G. & Sawadogo, M. Physical Analysis of Transcription Preinitiation Complex Assembly on a Class II Gene Promoter. *Science (1979)* **241**, 1335–1338 (1988).

75. Auble, D. T., Wang, D., Post, K. a I. W. & Hahn, S. Molecular Analysis of the SNF2 / SWI2 Protein Family Member MOT1 , an ATP-Driven Enzyme That Dissociates TATA- Binding Protein from DNA. **17**, 4842–4851 (1997).

76. Auble, D. T. & Hahn, S. An ATP-dependent inhibitor of TBP binding to DNA. *Genes Dev* **7**, 844–856 (1993).

77. Cote, J., Quinn, J., Workman, J. L. & Peterson, C. L. Stimulation of GAL4 Derivative Binding to Nucleosomal DNA by the Yeast SWI/SNF Complex. *Science (1979)* **265**, 53–60 (1994).

78. Andrau, J. *et al.* Mot1p is essential for TBP recruitment to selected promoters during in vivo gene activation. **21**, 5173–5183 (2002).

79. Dasgupta, A., Darst, R. P., Martin, K. J., Afshari, C. A. & Auble, D. T. Mot1 activates and represses transcription by direct , ATPase-dependent mechanisms. 1–6 (2002) doi:10.1073/pnas.052397899.

80. Geisberg, J. v. & Struhl, K. Cellular stress alters the transcriptional properties of promoter-bound Mot1-TBP complexes. *Mol Cell* **14**, 479–489 (2004).

81. Zentner, G. E. & Henikoff, S. Mot1 Redistributes TBP from TATA-Containing to TATA-Less Promoters. *Mol Cell Biol* **33**, 4996–5004 (2013).

82. Sprouse, R. O. *et al.* Function and structural organization of Mot1 bound to a natural target promoter. *Journal of Biological Chemistry* **283**, 24935–24948 (2008).

83. Cramer, P. Organization and regulation of gene transcription. *Nature* vol. 573 45–54 Preprint at https://doi.org/10.1038/s41586-019-1517-4 (2019).

84. Roeder, R. G. 50+ Years of Eukaryotic Transcription: an Expanding Universe of Factors and Mechanisms. *Nat Struct Mol Biol* (2019) doi:10.1038/s41594-019-0287-x.

85. Kawakami, E., Adachi, N., Senda, T. & Horikoshi, M. Leading role of TBP in the Establishment of Complexity in Eukaryotic Transcription Initiation Systems. *Cell Rep* **21**, 3941–3956 (2017).

86. Denissov, S. *et al.* Identification of novel functional TBP-binding sites and general factor repertoires. *EMBO Journal* **26**, 944–954 (2007).

87. Coleman, R. A., Taggart, A. K. P., Benjamin, L. R. & Pugh, B. F. Dimerization of the TATA binding protein. *Journal of Biological Chemistry* vol. 270 13842–13849 Preprint at https://doi.org/10.1074/jbc.270.23.13842 (1995).

88. Coleman, R. A. & Pugh, B. F. Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. *Journal of Biological Chemistry* **270**, 13850–13859 (1995).

89. Hoopes, B. C., LeBlanc, J. F. & Hawley, D. K. Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. *Journal of Biological Chemistry* **267**, 11539–11547 (1992).

90. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. Crystal structure of a yeast TBP/TATA-box complex. *Nature 1993 365:6446* **365**, 512–520 (1993).

91. Burley, Stephen K., Nikolov, D. B. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. Preprint at (1995).

92. Ravarani, C. N. J., Chalancon, G., Breker, M., de Groot, N. S. & Babu, M. M. Affinity and competition for TBP are molecular determinants of gene expression noise. *Nat Commun* **7**, 1–13 (2016).

93. Ravarani, C. N. J. *et al.* Molecular determinants underlying functional innovations of TBP and their impact on transcription initiation. *Nat Commun* **11**, 1–16 (2020).

94. Anandapadamanaban, M. *et al.* High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nat Struct Mol Biol* **20**, 1008–1014 (2013).

95. Juo, Z. S., Kassavetis, G. A., Wang, J., Geiduschek, E. P. & Sigler, P. B. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* **422**, 534–539 (2003).

96. Butryn, A. *et al.* Structural basis for recognition and remodeling of the TBP:DNA:NC2 complex by Mot1. *Elife* (2015) doi:10.7554/eLife.07432.001.

97. Wollmann, P. *et al.* Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. *Nature* (2011) doi:10.1038/nature10215.

98. Spedale, G. *et al.* Tight cooperation between Mot1p and NC2β in regulating genome-wide transcription, repression of transcription following heat shock induction and genetic interaction with SAGA. *Nucleic Acids Res* **40**, 996–1008 (2012).

99. Kim, Y., Ebright, Y. W., Goodman, A. R., Reinberg, D. & Ebright, R. H. Nonradioactive, ultrasensitive site-specific protein-protein photocrosslinking: Interactions of α-helix 2 of TATA-binding protein with general transcription factor TFIIA and transcriptional repressor NC2. *Nucleic Acids Res* **36**, 6143–6154 (2008).

100. Schluesche, P., Stelzer, G., Piaia, E., Lamb, D. C. & Meisterernst, M. NC2 mobilizes TBP on core promoter TATA boxes. *Nat Struct Mol Biol* **14**, 1196–1201 (2007).

101. Butryn, A., Woike, S., Shetty, S. J., Auble, D. T. & Hopfner, K.-P. Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state. *Elife* **7**, 1–12 (2018).

102. Ramya Viswanathan and David T. Auble. One small step for Mot1; one giant leap for other Swi2/Snf2 enzymes? *Biochim Biophys Acta* **1809**, 488–496 (2011).

103. Auble, D. T. & Steggerda, S. M. Testing for DNA Tracking by MOT1, a SNF2/SWI2 Protein Family Member. *Mol Cell Biol* **19**, 412–423 (2015).

104. Moyle-Heyrman, G., Viswanathan, R., Widom, J. & Auble, D. T. Two-step mechanism for modifier of transcription 1 (Mot1) enzyme-catalyzed displacement of TATA-binding protein (TBP) from DNA. *Journal of Biological Chemistry* **287**, 9002–9012 (2012).

105. Heiss, G. *et al.* Conformational changes and catalytic inefficiency associated with Mot1-mediated TBP–DNA dissociation. *Nucleic Acids Res* (2019) doi:10.1093/nar/gky1322.

106. Maeshima, K., Imai, R., Tamura, S. & Nozaki, T. Chromatin as dynamic 10-nm fibers. *Chromosoma* **123**, 225–237 (2014).

107. Finch, J. T. & Klug, a. Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci U S A* **73**, 1897–1901 (1976).

108. Li, Y. *et al.* Analysis of three-dimensional chromatin packing domains by chromatin scanning transmission electron microscopy (ChromSTEM). *Sci Rep* **12**, (2022).

109. Ou, H. D. *et al.* ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science (1979)* **357**, (2017).

110. Maeshima, K., Tamura, S., Hansen, J. C. & Itoh, Y. Fluid-like chromatin: Toward understanding the real chromatin organization present in the cell. *Current Opinion in Cell Biology* vol. 64 77–89 Preprint at https://doi.org/10.1016/j.ceb.2020.02.016 (2020).

111. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

112. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics* vol. 20 207–220 Preprint at https://doi.org/10.1038/s41576-018-0089-8 (2019).

113. Zaret, K. S. Pioneer Transcription Factors Initiating Gene Network Changes. *Annual Review of Genetics* vol. 54 367–385 Preprint at https://doi.org/10.1146/annurev-genet-030220-015007 (2020).

114. Musselman, C. A., Lalonde, M. E., Côté, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nature Structural and Molecular Biology* vol. 19 1218–1227 Preprint at https://doi.org/10.1038/nsmb.2436 (2012).

115. Zhang, T., Cooper, S. & Brockdorff, N. The interplay of histone modifications – writers that read. *EMBO Rep* **16**, 1467–1481 (2015).

116. Poli, J., Gasser, S. M. & Papamichos-Chronakis, M. The INO80 remodeller in transcription, replication and repair. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**, (2017).

117. Gerhold, C. B., Hauer, M. H. & Gasser, S. M. INO80-C and SWR-C: Guardians of the genome. *Journal of Molecular Biology* Preprint at https://doi.org/10.1016/j.jmb.2014.10.015 (2015).

118. Gaku Mizuguchi, Xuetong Shen, Joe Landry, Wei-Hua Wu & Carl Wu. ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. *Science (1979)* **303**, (2004).

119. Yu, H. *et al.* INO80 promotes H2A.Z occupancy to regulate cell fate transition in pluripotent stem cells. *Nucleic Acids Res* **49**, 6739–6755 (2021).

120. Brahma, S. *et al.* INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. (2017) doi:10.1038/ncomms15616.

121. Krietenstein, N. *et al.* Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709-721.e12 (2016).

122. Dang, W. & Bartholomew, B. Domain Architecture of the Catalytic Subunit in the ISW2-Nucleosome Complex. *Mol Cell Biol* **27**, 8306–8317 (2007).

123. Hodges, C., Kirkland, J. G. & Crabtree, G. R. The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb Perspect Med* **6**, (2016).

124. Gerhold, C. B. & Gasser, S. M. INO80 and SWR complexes: Relating structure to function in chromatin remodeling. *Trends in Cell Biology* Preprint at https://doi.org/10.1016/j.tcb.2014.06.004 (2014).

125. Knoll, K. R. *et al.* The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nat Struct Mol Biol* **25**, 823–832 (2018).

126. Aramayo, R. J. *et al.* Cryo-EM structures of the human INO80 chromatin-remodeling complex. *Nat Struct Mol Biol* (2017) doi:10.1038/s41594-017-0003-7.

127. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. Atomic structure of the actin: DNase I complex. *Nature* **347**, 37–44 (1990).

128. Otterbein Ludovic R, Graceffa Phillip & Dominguez, R. The Crystal Structure of Uncomplexed Actin in the ADP State. *Science (1979)* **293**, 708–710 (2001).

129. Klages-Mundt, N. L., Kumar, A., Zhang, Y., Kapoor, P. & Shen, X. The Nature of Actin-Family Proteins in Chromatin-Modifying Complexes. *Front Genet* **9**, 1–16 (2018).

130. F. Merino, S. Pospich, J. Funk, F. Küllmer, H.-D. Arndt, P. Bieling, S. R. Structural transitions of F-actin upon ATP hydrolysis revealed at near-atomic resolution by electron cryo microscopy. *Nat Struct Mol Biol* **25**, (2018).

131. Cao, T. *et al.* Crystal structure of a nuclear actin ternary complex. *Proc Natl Acad Sci U S A* **113**, 8985–8990 (2016).

132. von der Ecken, J. *et al.* Structure of the F-actin-tropomyosin complex. *Nature* **519**, 114–117 (2015).

133. Szerlong, H. *et al.* The HSA domain binds nuclear actin-related proteins to regulate chromatin-remodeling ATPases. *Nat Struct Mol Biol* **15**, 469–476 (2008).

134. Fenn, S. *et al.* Structural biochemistry of nuclear actin-related proteins 4 and 8 reveals their interaction with actin. *EMBO Journal* **30**, 2153–2166 (2011).

135. Schubert, H. L. *et al.* Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler. *Proc Natl Acad Sci U S A* **110**, 3345–3350 (2013).

136. Dion, V., Shimada, K. & Gasser, S. M. Actin-related proteins in the nucleus: Life beyond chromatin remodelers. *Current Opinion in Cell Biology* vol. 22 383–391 Preprint at https://doi.org/10.1016/j.ceb.2010.02.006 (2010).

137. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* (2017) doi:10.1038/nrm.2017.26.

138. Brahma, S., Ngubo, M., Paul, S., Udugama, M. & Bartholomew, B. The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling. *Nat Commun* **9**, 1–10 (2018).

139. Tosi, A. *et al.* Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* **154**, 1207 (2013).

140. Chen, L. *et al.* Subunit organization of the human INO80 chromatin remodeling complex: An evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *Journal of Biological Chemistry* **286**, 11283–11289 (2011).

141. Oberbeckmann, E. *et al.* Ruler elements in chromatin remodelers set nucleosome array spacing and phasing. *Nat Commun* **12**, (2021).

142. Oberbeckmann, E. *et al.* Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nat Commun* **12**, (2021).

143. Clapier, C. R. *et al.* Regulation of DNA Translocation Efficiency Within the Chromatin Remodeler RSC/Sth1 Potentiates Nucleosome Sliding and Ejection. *Mol Cell* **62**, 453–461 (2016).

144. Turegun, B., Baker, R. W., Leschziner, A. E. & Dominguez, R. Actin-related proteins regulate the RSC chromatin remodeler by weakening intramolecular interactions of the Sth1 ATPase. *Commun Biol* **1**, (2018).

145. Adamkewicz, J. I., Mueller, C. G. F., Hansen, K. E., Prud'homme, W. A. & Thorner, J. Purification and enzymic properties of Mot1 ATPase, a regulator of basal transcription in the yeast saccharomyces cerevisiae. *Journal of Biological Chemistry* **275**, 21158–21168 (2000).

146. Clapier, C. R. Sophisticated Conversations between Chromatin and Chromatin Remodelers , and Dissonances in Cancer. (2021).

147. Zhou, C. Y., Johnson, S. L., Gamarra, N. I. & Narlikar, G. J. Mechanisms of ATP-Dependent Chromatin Remodeling Motors. *Annu Rev Biophys* (2016) doi:10.1146/annurev-biophys-051013-022819.

148. Viswanathan, R., True, J. D. & Auble, D. T. Molecular mechanism of Mot1, a TATA-binding protein (TBP)-DNA dissociating enzyme. *Journal of Biological Chemistry* (2016) doi:10.1074/jbc.M116.730366.

149. Yan, L. & Chen, Z. A Unifying Mechanism of DNA Translocation Underlying Chromatin Remodeling. *Trends Biochem Sci* 1–11 (2019) doi:10.1016/j.tibs.2019.09.002.

150. Cai, Y. *et al.* YY1 functions with INO80 to activate transcription. *Nat Struct Mol Biol* **14**, 872–874 (2007).

151. Zarrinpar, A. & Lim, W. A. Converging on proline: The mechanism of WW domain peptide recognition. *Nat Struct Biol* **7**, 611–613 (2000).

152. Verdecia, M. A., Bowman, M. E., Ping, K., Hunter, T. & Noel, J. P. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* **7**, 639–643 (2000).

153. Xin, H. *et al.* Structure of a WW domain containing fragment of dystrophin in complex with β-dystroglycan. *Nat Struct Biol* **7**, 634–638 (2000).

154. Qu, K., Chen, K., Wang, H., Li, X. & Chen, Z. Structure of the NuA4 acetyltransferase complex bound to the nucleosome. *Nature* **610**, 569–574 (2022).

155. Wilkinson, F. H., Park, K. & Atchison, M. L. Polycomb recruitment to DNA in vivo by the YY1 REPO domain. *Proc Natl Acad Sci U S A* **103**, 19296–19301 (2006).

156. Alfieri, C. *et al.* Structural basis for targeting the chromatin repressor Sfmbt to Polycomb response elements. *Genes Dev* **27**, 2367–2379 (2013).

157. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* (2021) doi:10.1038/s41586-021-03819-2.

158. Baek, M. *et al. Accurate prediction of protein structures and interactions using a three-track neural network*. *Science* vol. 373 https://predictioncenter.org/casp14/ (2021).

159. He, S. *et al.* Structure of nucleosome-bound human BAF complex. *Science (1979)* **367**, 875–881 (2020).

160. Gerhold, C. B. *et al.* Structure of Actin-related protein 8 and its contribution to nucleosome binding. *Nucleic Acids Res* **40**, 11036–11046 (2012).

Abbreviations

ADP	adenosine diphosphate
Arp	actin-related protein
ATP	adenosine triphosphate
ATPγS	adenosine 5'-(γ-thio) triphosphate
bp	base pair
cryo-EM	cryogenic electron microscopy
cryo-ET	cryogenic electron tomography
Ct	*Chaetomium thermophilum*
C-terminal	carboxy-terminal
DNA	desoxyribonucleic acid
ds	double strand
Ec	*Encephalitozoon cuniculi*
EMSA	electrophoretic mobility shift assay
F-actin	filamentous actin
FRET	Foerster Resonance Energy Transfer
G-actin	globular actin
GRF	general regulatory factor
GTF	general transcription factor
HEAT	Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR
Hs	*Homo sapiens*
HAS	Helicase-SANT-associated
HSS	HAND-SANT-SLIDE
MNase	micrococcal nuclease
N-actin	nuclear actin
NC2	negative cofactor 2
NDR	nucleosome-depleted region
N-terminal	amino-terminal (NH_2-terminal)
NTP	nucleotide triophosphate
PDB	Protein Data Bank
RNA	ribonucleic acid
Sc	*Saccharomyces cerevisiae*
SEC	size-exclusion chromatography
SGD	salt gradient dialysis
SHL	superhelical location
SF	superfamily
ss	single strand
TSS	transcription start site

Acknowledgements

First of all, I want to thank Karl-Peter Hopfner for the great opportunity to work in his laboratory. The curiosity driven research in his group, applying state-of-the-art technology and his never-ending fascination for structural biology deeply inspired me and my research. I thank him for putting the trust in me and the autonomy that he granted me.

Special thanks to Sebastian Eustermann, whose scientific excellence and creativity I always admired. Thanks for being a companion during this research project, especially at the beginning!
I want to thank our remodeler subgroup for countless hours of discussing and having fun. Thank you, Felix, Franzi, Kevin, Vanessa and Kilian! And thanks to Manu for accompanying me during the exhausting cloning procedures, sharing her expertise and supporting me with protein expression and purification!

Thanks to the whole Hopfner group, who really felt like home during the last years. Being part of a group of such ambitious scientists was a privilege. A big thank you to the people running the lab, especially Brigitte and Gregor, also for keeping the spirit high and being reliable colleagues in every situation.

Thanks, Katja and Joe for running the electron microscopy and for granting me their time and that they were always ready for the next Mot1 conformation to collect data on!
Thank you to all the people additionally involved in the publications: James Jung, Simon Wenzl, Götz Hagemann, Franz Herzog. A big thanks to Agata Butryn, my predecessor in the Mot1 project, who documented her experiments in great detail.