Inaugural dissertation
for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of the
Ruprecht – Karls – University
Heidelberg

Presented by

M.Sc. Anna Jungblut
Born in: Emmendingen, Germany
Oral examination: 22.07.2024

Structural and biochemical studies of ATP-dependent chromatin remodelers from the INO80/SWR1 family

Referees: Prof. Dr. Irmgard Sinning
Dr. Julia Mahamid

Preamble

Abstract

The genome of a eukaryotic cell is packaged by histone proteins into nucleosomes. ATP-dependent chromatin remodeling enzymes shape this first level of chromatin organisation and play a fundamental role in regulating access of nuclear factors to DNA. Eukaryotes possess four evolutionary-conserved families of chromatin remodelers. The INO80/SWR1 family members assemble into mega-Dalton sized, multi-subunit complexes and configure the nucleosome landscape at gene promoters, origins of replications and sites of DNA damage. In *Saccharomyces cerevisisae* (*S. c.*), the 15-subunit INO80 complex translocates DNA through ATP hydrolysis along the histone core leading to positioning of +1 nucleosomes downstream of promoter DNA. Subsequently, the 14-subunit *S. c.* SWR1 complex exchanges the canonical histone H2A/H2B dimers with the variant H2A.Z/H2B through ATP hydrolysis. H2A.Z containing +1 nucleosomes are proposed to serve as a gateway for gene transcription. Despite of this importance, the precise molecular mechanisms underlying these nucleosome remodeling reactions are poorly understood. In addition, it remains unclear what defines the substrate specificity and the different reaction outcomes of INO80 and SWR1 remodelers which in principle share a common subunit composition including nuclear actin and Actin-related proteins (Arps). To address these questions, I characterized the structural basis of nucleosome remodeling by INO80 and SWR1 complexes by using a combination of cryo-electron microscopy (cryoEM) and biochemistry.

In the first part of this thesis, I present the mechanistic basis of hexasome remodeling by the INO80 complex. Hexasomes are non-canonical nucleosomes lacking one copy of H2A/H2B and are a hallmark of actively transcribed genes. My findings demonstrate how INO80 identifies non-canonical DNA and histone marks of hexasomes that result from the absence of H2A/H2B. INO80 is activated by directly sensing an exposed H3/H4 histone interface, operating independently of the H2A/H2B acidic patch. This is enabled through a spin-rotated binding mode of the Ino80 motor domain while its Arp module remains tethered to the now unwrapped linker DNA. These findings shed light on how the absence of H2A/H2B creates altered substrates for chromatin remodeling, thereby creating a different, however unexplored layer of energy-mediated chromatin regulation.

In the second part, I present the first structural and biochemical study of SWR1 from *Chaetomium thermophilum* (*C. t.*). I identified the subunit composition of *C. t.* SWR1 and implemented a recombinant expression and purification strategy for both *S. c.* and *C. t.* SWR1 that enabled subsequent cryoEM studies. In addition, I investigated the structure of the Arp module by using an integrative structural biology approach building on cryoEM, AI-powered AlphaFold structure predictions and crosslinking mass spectrometry. Thereby, I determined an almost complete model of the Apo SWR1 complex in which the Arp module is flexibly tethered to the core of SWR1. A kink in the Arp module enables histone tail recognition of the Yaf9 subunit, while recognition of linker DNA might occur by switching SWR1 into an open conformation. Histone tail recognition and tethering to linker DNA may specify SWR1 activity at +1 nucleosomes.

INO80 and SWR1 share similar architectures, yet drive distinct functional outcomes. My thesis provides a structural framework for understanding the underlying molecular mechanisms.

Zusammenfassung

Chromatin verpackt und reguliert das Genom einer eukaryotischen Zelle. Das Nukleosom ist die grundlegende Einheit des Chromatins, in der DNA um Histonproteine gewickelt ist. ATP-hydrolysierende Enzyme, sogenannte Chromatin Remodeler, katalysieren die Anordnung und Zusammensetzung von Nukleosomen und spielen damit eine Schlüsselrolle in der Genomregulation. Chromatin Remodeler werden in vier verschiedene Familien unterteilt. Die INO80/SWR1-Familie besteht aus mega-Dalton großen, mehrteiligen Komplexen und gestaltet die Nukleosomenlandschaft an Genpromotoren, Replikationsursprüngen und DNA-Doppelstrangbrüchen. Der aus 15 Untereinheiten bestehende INO80-Komplex des Pilzes *Saccharomyces cerevisisae* (*S. c.*) katalysiert die Translokation von Histonen entlang der DNA durch ATP-Hydrolyse, was zur Positionierung von +1 Nukleosomen neben Promoterensequenzen von Genen führt. Anschließend tauscht der *S. c.* SWR1 Komplex durch ATP-Hydrolyse in diesen +1 Nukleosomen die kanonischen Histondimere H2A/H2B durch die H2A.Z/H2B Variante aus, die ein Merkmal aktiver Genexpression sind. Trotz dieser fundamentalen Bedeutung von INO80 und SWR1, sind die genauen molekularen Mechanismen, die diesen Nukleosomenumbaureaktionen zugrunde liegen, unzureichend verstanden. Insbesondere da beide Chromatin Remodeler strukturelle Ähnlichkeiten aufweisen und sie eine ähnliche Zusammensetzung haben, wie zum Beispiel nukleares Aktin und Aktin-ähnliche Proteine (Arps). Daher habe ich in meiner Doktorarbeit die molekularen Grundlage des Nukleosomenumbaus durch INO80 und SWR1 Chromatin Remodeler untersucht, indem ich eine Kombination aus Cryo-Elektronenmikroskopie (CryoEM) und biochemischen Methoden angewendet habe.

Im ersten Teil dieser Arbeit stelle ich meine Ergebnisse bezüglich der mechanistischen Grundlagen des Hexasom-Umbaus durch den INO80 Komplex vor. Hexasomen sind nicht-kanonische Nukleosomen, denen eine Kopie von H2A/H2B fehlt und die ein charakteristisches Merkmal aktiv transkribierter Gene sind. Ich zeige wie INO80 nicht-kanonische DNA- und Histonmerkmale von Hexasomen erkennt, die aus dem Verlust von H2A/H2B resultieren. INO80 rotiert um ca. 145° um Hexasomen zu binden, während sein Arp-Modul an der flexiblen Linker-DNA gebunden bleibt. Aufgrund des fehlenden H2A/H2B Dimers wird H3/H4 offen gelegt und nun stattdessen von INO80 und damit unabhängig vom Acidic Patch des H2A/H2B Histondimers erkannt. Diese Erkenntnisse zeigen, wie der Verlust von H2A/H2B nicht nur kompensiert wird, sondern auch zu einer unerwarteten Form der aktiven Chromatinregulation beitragen kann.

Der zweite Teil meiner Arbeit beschreibt die erste strukturelle und biochemische Studie von SWR1 aus dem thermophilen Fungus *Chaetomium thermophilum* (*C. t.*). Ich habe die Zusammensetzung von *C. t.* SWR1 identifiziert und eine rekombinante Expressions- und Reinigungsstrategie für *S. c.* als auch *C. t.* SWR1 etabliert, die anschließende CryoEM Studien ermöglichte. Mit einer integrativen Herangehensweise aus CryoEM, KI-unterstützen AlphaFold Strukturvorhersagen und Crosslinking Massenspektrometrie, habe ich außerdem die Struktur des Arp-Moduls untersucht. Dies erlaubte mir eine nahezu vollständige Struktur des Apo SWR1 Komplexes zu bestimmen, in der das Arp-Modul flexibel mit dem Hauptmodul von SWR1 verbunden ist. Ein 90° Knick im Arp-Modul ermöglicht die Erkennung von Histonenden durch die Yaf9 Untereinheit, während eine Konformationsänderung des SWR1 Komplexes in eine offene Konfiguration die Bindung der Linker-DNA ermöglicht. Die Erkennung der Histonenden in Kombination mit der Bindung an die Linker-DNA erklärt die Substratspezifität von SWR1.

Trotz der Ähnlichkeit von INO80 und SWR1 katalysieren diese Enzyme unterschiedliche Reaktionen in der Genomregulation. Meine Doktorarbeit etabliert die strukturellen Grundlagen, um die molekularen Mechanismen dieser Funktion zu verstehen.

Publications

1. Zhang M., **Jungblut A.***, Kunert F.*, Hauptmann L., Hoffmann T., Kolesnikova O., Metzner F., Moldt M., Weis F., DiMaio F., Hopfner K.-P., Eustermann S. Hexasome-INO80 complex reveals structural basis of noncanonical nucleosome remodeling. *Science* 381(6655), 313-319 (2023)

2. **Jungblut A.**, Hopfner K.-P., Eustermann S. Megadalton chromatin remodelers: common principles for versatile functions. *Current Opinion of Structural Biology* Volume 64, 134-144 (2020)

3. Vorländer M., **Jungblut A.**, Karius K., Baudin F., Kosinski J., Müller C.W. Structure of the TFIIIC subcomplex τA provides insights into RNA polymerase III pre-initiation complex formation *Nature Communications* 11, 4905 (2020)

Acknowledgements

The presented PhD work would not have been achievable without the support, advice and help from a wonderful group of people surrounding me. I want to take the opportunity and thank everyone of them.
I want to acknowledge all past and present members of the Eustermann lab (Sebastian Eustermann, Olga Kolesnikova, Min Zhang, Dana Grozavu, Thomas Dahlet, Lucia Alvarez, Luis Hauptmann, Lena Gottschalk, Jennifer Sauerland, Matteo Doudin and Olivier Gemin). Thank you for making the long days in the lab more fun as well as all the pizza nights we had. Most importantly, I am grateful to Sebastian for accepting me in his lab and trusting me to built up the lab with him and Olga Kolesnikova. I always appreciated the scientific discussions, your support and enthusiasm for even little achievements. Thanks for giving me the freedom to do the science I like.
I owe my deepest gratitude to our lab manager Olga. She is the good soul of the lab and motivated me even in difficult times. Olga prepared the nucleosomal substrates I used for my structural studies, gave me advice for cloning and protein productions and lastly continued the efforts of producing yeast SWR1. Next, I want to highlight Min Zhang who collaborated with me on the INO80 hexasome project. Min initiated the project, performed sample preparation, data acquisition and data processing. Our success of our *Science* publication would not have been possible without her. During my last year of my PhD, the master intern Lena Gottschalk supported me in studying the Arp module of SWR1. She cloned and purified the different Arp module constructs and pioneered the cryoEM work. Thank you for your hard and *beautiful* work and your catching enthusiasm for science. Another intern, Luis Hauptmann, developed a single-molecule mapping technique in the lab that he applied for the SWR1 dataset.
In addition, I acknowledge Thomas Hoffmann who keeps the IT infrastructure at SCB at the topmost level, keeps softwares up to date and is very patient with request on terminal usage and coding questions.
Next, I thank the past and current members of the cryoEM facility including Joseph Bartho, Félix Weis and Wim Hagen for keeping the microscopes in such a good state, for teaching me how to use them and for helping me out whenever the microscopes did not want to cooperate with me. I am beyond grateful to our protein expression facility and the Müller lab for sharing constructs for cloning or insect cells with me. A special thanks goes to Karine Lapouge for showing me how to use the mass photometer and how to perform MST measurements. Also a big thank you to Jonas Weidenhausen, Helga Grötsch and again Olga for the trouble shooting at the ÄKTA machines.
Furthermore, I thank my TAC members Prof. Dr. Irmgard Sinning, Dr. Julia Mahamid, Prof. Dr. Christian Häring and Prof. Dr. Beat Fierz for their feedback, advice and the fruitful discussions we had during the TAC meetings.
Finally, I want to thank the whole EMBL community and especially the 2019 predoc cohort for making my time at EMBL unforgettable – for all the fun we had, interesting seminars, coffee breaks, lunches and long beer hours. Only with your support, it was possible to go through the Covid pandemic. A special note goes to my predoc peers Léanne Strauß and Andi Brunner and to our *Basic Biology Mondays* that will for sure proof useful during my PhD defence. Lastly, I thank the most important person in my life – my husband Daniel. Thank you for your never ending support and for proof-reading my thesis even with your math and computer science background.

Abbreviations

ADP	Adenosine diphosphate
Arp	Actin-related protein
ATP	Adenosine triphosphate
AZ	nucleosomes with one copy of H2A and H2A.Z respectively
bp	Base pair
BS3	Bis(sulfosuccinimidyl)suberate
BTB	Brf1 and TBP
CHD	Chromodomain helicase DNA-binding
C. t.	Chaetomium thermophilum
CTF	Contrast transfer function
DBD	DNA binding domain
DNA	Desoxyribonucleic acid
DSB	Double-strand break
DTT	Dithiothreitol
eGFP	Enhanced green flourescent protein
EM	Electron microscopy
6-FAM	6-Carboxyfluorescein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMG	High mobility group
HSA	Helicase-SANT-associated
Ies	Ino eighty subunit
INO80	Inositol auxotrophy 80
IPTG	Isoprpyl-β-D-thiogalactopyranoside
ISWI	Imitation switch
kb	Kilobase pairs
kDa	KiloDalton
MST	Microscale thermophoresis
N. c.	Neurospora crassa
NFR	Nucleosome-free region
NP40	Nonidet P-40 (octylphenoxypolyethoxyethanol)
nt	Nucleotide
NuA4	Nucleosome Acetyltransferase of H4
LB	Lysogeny broth
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
Pol II	RNA Polymerase II
RNA	Ribonucleic acid
RSC	Remodeling the Structure of Chromatin
SANT	Swi3-Adaptor 2- Nuclear receptor corepressor-TFIIIB
S. c.	Saccharomyces cerevisiae
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SHL	Super Helical Location
SRCAP	Snf2-related CREBBP activator protein
SWI/SNF	Switch/sucrose non-fermentable
SWR1	SWI/SNF related 1

Taf14 TBP associated factor 14
Tip60 Tat-interactive protein, 60 kDa
TMT Tandem Mass Tag
TSS Trancription start site
wt Wildtype
XL crosslinking
XL-MS crosslinking mass spectrometry
YEATS Yaf9, ENL, AF9, Taf14, and Sas5
Z/B H2A.Z/H2B
ZF zinc finger
ZZ nucleosomes with two copies of H2A.Z

Contents

1 Introduction

1.1 Chromatin architecture

1.1.1 Nucleosome architecture and chromatin

The basic unit of chromatin is formed by nucleosomes that arise from two H2A/H2B histone dimers and one H3/H4 tetramer as well as DNA. Together the histone dimers and the tetramer give rise to the histone core also referred to as octamer. About 147 bp of DNA wrap 1.65-times around the histone core, thereby enabling tight packing of DNA into the cell nucleus [1]. The DNA-protein interactions are facilitated by positively charged histone residues that pack against the negatively charged DNA sugar backbone of the DNA minor groves. The DNA is bend around the histone core which results in a preference of GC base pairs for major grooves and AT pairs for minor grooves, thereby favoring bending of DNA altering from its B-form [2][3][4]. The approximately fourteen helical turns that wrap around the histone core are termed superhelical locations (SHL) from −7 to +7 (Figure 1.1a). They refer to the middle position of the minor groove that is solvent exposed and consequently accessible for chromatin factors. Half numbers (e.g. SHL +3.5) represent the minor groove contact points to the histone core. SHL 0 refers to the dyad position of the nucleosome that forms the symmetry point of the nucleosome that harbors a two-fold pseudo symmetry. *In vivo*, the symmetry is broken through non-symmetrical DNA sequences on both sides of the nucleosome.

Histone fold, modifications and variants

Histones comprise a common histone fold of three helices that are intersected by a loop respectively [1] (Figure 1.1b). The helices and loops are referred to as helices α1-3 and loops L1 and L2. In addition, histone specific features are present adjacent to the histone fold on the N-terminus and in some cases the C-terminus as well. The N-termini of the histones (and especially lysine residues) are unstructured and subjected to post-translational modifications including methylation, acetylation, phosphorylation, ubiquitination and SUMOylation (reviewed in [5][6]). These chemical modifications represent a hallmark of the epigenetic code and play pivotal roles in all DNA related processes including gene regulation. Some modifications like acetylation have a direct impact on chromatin architecture as this negative chemical group equalizes the positively charged lysines of the histone tails. As a consequence, the chromatin is more loose and more accessible for chromatin factors for binding [7][8]. Post-translational modifications are introduced by *writers* (e.g. acetylases), recognized by *readers* that can act as effector proteins (e.g. through bromo domains in case of acetylated histones) and removed by *erasers* (e.g. de-acetylases).
Besides nucleosome occupancy and histone tail modifications, histone variants are a third pillar of the epigenetic code [9]. Histone variants exist for all four canonical histones but are more common for H2A and H3 (reviewed in [10]). While the canonical histones are expressed and deposited during mitosis, histone variants are introduced in a replication independent manner into nucleosomes. Histone variants differ in their amino acid sequence from their canonical counterparts, but often only very subtle. The change in sequence results in their recognition of a specialized set of histone chaperons and chromatin factors. In addition, altered post-translational modifications can occur and the

altered amino acid sequence can have a direct impact on chromatin architecture like the human H3.3 histone variant that facilitates nucleosome ejection. As a consequence, histone variants have implications in all DNA-templated processes [10] including gene regulation, DNA repair, nucleosome turnover, telomere maintenance and centromere establishment. While some histone variants are continuously expressed, others are subjected to dynamic patterns or even tissue specific expression. In addition some histone variants are species-specific while others like H2A.Z are very conserved and thought to only have evolved once in evolution [11].

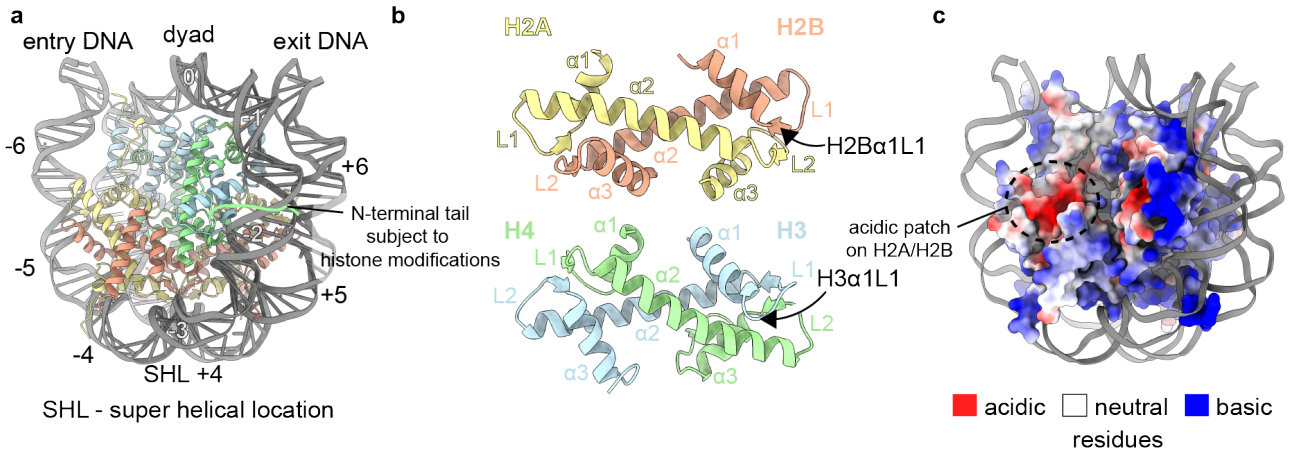


Figure 1.1: **a** Nucleoesome structure with labelled super helical locations (SHL) and N-terminal H4 tail highlighted as possible target for histone modifications. **b** Histone fold of H2A, H2B, H3 and H4. Labelling of helices and loops according to [12]. **c** Same nucleosome view as in **a** but with histones colored according to their respective charge. In dotted sphere, acidic patch formed by H2A and H2B residues. Representation of **a** adapted from [13] and based on PDB 1AOI.

Common features for nucleosome binding

The recent advancement in cryoEM enabled the structural characterization of chromatin factors bound to nucleosomes. Thereby, common principles for nucleosome recognition could be identified [14], that are: multivalency of nucleosome contacts, acidic patch binding on the H2A/H2B surface as well as protein-histone elbow interactions. In more detail, the acidic patch is formed by a combination of conserved H2A (E56, E61, E64, D90, E91, and E92; in *Xenopus*) and H2B residues (E105 and E113) [1]. The underlying aspartate and glutamine residues create a negatively charged surface on the nucleosome (Figure 1.1c) that is commonly recognized by chromatin factors [14] [15][16]. The interaction is faciliated through two or more positively charged arginine residues of the respective chromatin factor and is also referred to as arginine anchor nucleosome-binding motif [17].
Another common feature for nucleosome binding are the H3α1L1 and H2Bα1L1 elbows that sit at the junction of the α1 helix and the adjacent loop L1 respectively [14]. The interface of the alpha helices and loops resemble the shape of an elbow and are in case of H3α1L1 and H2Bα1L1 solvent exposed and hence accessible for chromatin factors.

First level of chromatin organization and chromatin organization around promoter sites

On a more global scale, two neighboring nucleosomes are connected by linker DNA – also referred to as extra-nucleosomal DNA. Nucleosomes are not randomly distributed on genomic DNA but are

evenly spaced through energy consumption and dedicated enzymes [18][19]. Regular assemblies of nucleosomes are called arrays. Their presence is anti-correlated with gene activity [20], meaning that nucleosomes in actively transcribed regions are less ordered as RNA Polymerase II leaves nucleosomes at random positions after transcribing through them. Nucleosomes can be positioned to certain genomic locations. Such a high degree of positioning is observed in yeast [21] while it is less prominent in metazoans and plants [22][23]. Furthermore, arrays can be phased in relation to certain loci, protein barriers or well positioned nucleosomes [24]. Most prominently, phased arrays are found at binding sites of transcription factors, insulators or around transcription start sites [9][25][26][27][28]. Promoter sites are enriched for poly dA:dT sequences [21][29] which disfavor intrinsically wrapping around histones. These sequences together with regulatory factors like Rap1, Reb1 and Abf1 recruit chromatin remodeling complexes (introduced in the next section) of the SWI/SNF family which in turn remove nucleosomes from promotors [30][24][31]. This creates a nucleosome free region (NFR) upstream of the transcription start site (TSS) which is required for recruitment of the RNA Polymerase machinery and hence for active gene transcription [32][33]. Following the creation of the NFR, +1 and −1 nucleosomes down and upstream of the NFR are positioned by chromatin remodelers [34][35][24][36][37]. Furthermore, correct positioning of the +1 nucleosome is linked with correct TSS selection and transcription initiation [38] (see Figure 1.2). The nucleosomes downstream of the +1 nucleosome are regularly spaced through the interplay of chromatin remodelers (ISWI, CHD, INO80) [39][40][24], thus forming a phased array in relation to the +1 nucleosome.

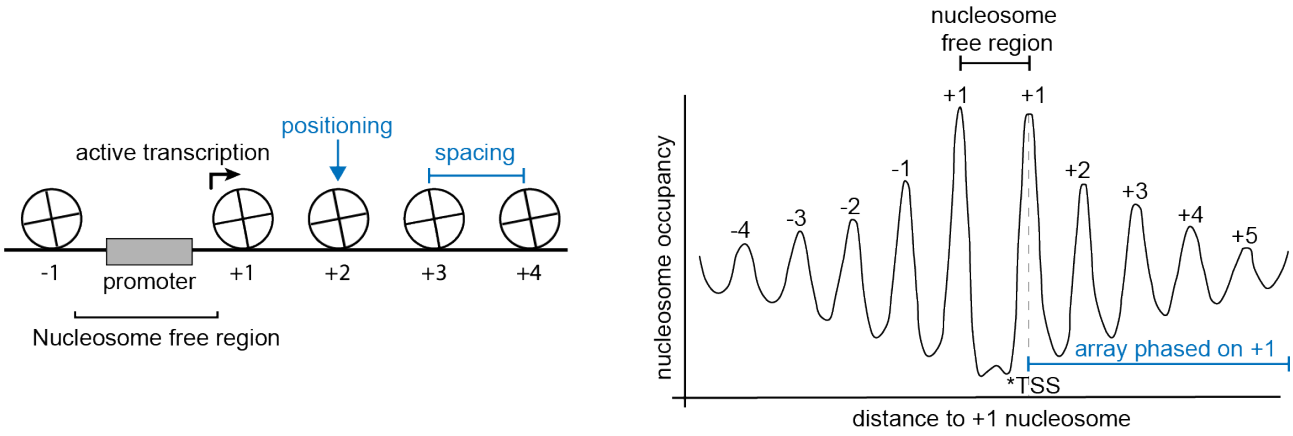


Figure 1.2: **Nucleosome positioning, spacing and phasing around promoter sites** *Left*: Chromatin organization downstream of promoter sites. The +1 nucleosome is well positioned and nucleosomes are equally spaced. *Right*: Schematic graph of nucleosome occupancy around promoter sites. Nucleosomes downstream of the nucleosome free region are phased in relation to the +1 nucleosome.

Higher-order chromatin structure

The linker length connecting two neighboring nucleosomes varies among different species, cell types, developmental stage and active versus repressed chromatin [41][42] which has a direct impact in chromatin organization. Higher-order chromatin structures are enabled through nucleosome-nucleosome interactions. Here, the main contacts are formed of one nucleosome and the one two nucleosomes away (N with N+2) [43]. The two nucleosomes are stacking on top of each other which is facilitated by the interaction of a basic region on the H4 N-terminus and the acidic patch

of the N+2 nucleosome [42]. In addition, the presence of H1 linker histones that can be positioned at the entry and exit sites of nucleosomes, can further stabilize the packaging of chromatin [44]. This gives rise to an overall zigzag architecture enabling the formation of chromatin fibres. Until recently, those fibres were thought to have a static diameter of 30 nm [45] but they arose artificially from non-physiological *in vitro* reconstitutions and present cation ions. Recent advances in *in situ* electron tomography uncovered dynamic fiber diameters ranging from 8 to 24 nm [18][46]. Further, clusters of 2-8 nucleosomes are formed as visualized by super resolution microscopy [47]. These features enable the packaging of about 2 m of DNA into each of our cell nuclei.

1.1.2 The histone variant H2A.Z

Besides the known four canonical histones H2A, H2B, H3 and H4, a zoo of histone variants exists [10] presenting another layer of genome regulation. One among them is the H2A variant H2A.Z (Htz1 in yeast) which shares about 60 % sequence identity with H2A [48]. At the same time H2A.Z is highly conserved among different species. The presence of H2A.Z is essential for eukaryotic embryo development [49][50]. Overall, the 3D architecture and protein-DNA contacts of H2A.Z containing nucleosomes resemble nucleosomes containing the canonical H2A as shown by X-ray crystallography [48]. However, on a molecular level the differences in amino acid sequence result in a destabilization of the dimer-tetramer and dimer-dimer interfaces making the H2A.Z-containing histone core less stable. In more detail, three hydrogen bonds in the interface of the H3αN and the C-terminus of H2A.Z are absent in comparison to canonical H2A. Also, the L1 loop of H2A.Z that is self-interacting with the second H2A.Z copy is subjected to amino acid difference leading to a destabilization of the dimer-dimer interaction. On top of that, the acidic patch of H2A.Z is extended by two residues increasing the size of the acidic patch from six to eight residues. Thereby, the nucleosomal interface that is recognized by chromatin binding factors is significantly changed. H2A.Z is present on active gene sites [51][52][53] promoting transcriptional activity through the less stable histone core and a higher degree of flexibility between the DNA entry/exit sites and the histone core [54]. However, the impact of H2A.Z is context dependent and the higher flexibility of the DNA entry/exit sites can also lead to chromatin compaction and heterochromatin formation [54]. The functions of H2A.Z and how its energy required incorporation is facilitated, is presented in Sections 1.5.1 and 1.5.3.

1.1.3 Hexasomes

Nucleosome particles are not static but rather dynamic in regard to their composition and structure. In more detail, nucleosomes can undergo a breathing motion by altering the DNA-histone interactions. Alteration in histone composition leads to the formation of sub-nucleosomal particles that can be also referred to as non-canonical nucleosomes[55][56]. Among these are tetrasomes that contain only the H3/H4 tetramer or hexasomes [57]. Hexasomes are composed of six instead of eight histones, namely the H3/H4 tetramer and only one copy of H2A/H2B [58]. The loss of one H2A/H2B results in a smaller DNA footprint compared to canonical nucleosomes [56][59]. Hexasomes are not only a theoretical construct, but were found *in vivo* in context of transcription by RNA Polymerase II (Pol II) as well as DNA replication [58][55][60]. Both events require a transient disturbance of histone-DNA contacts to allow the polymerases to pass through nucleosomes [61]. Furthermore, structural methods [62][63] in combination with *in vivo* mapping [55] suggest that Pol II causes the eviction of H2A/H2B thereby creating hexasomes. In addition, a study in *Drosophila melongaster* showed that it is either the proximal or the distal H2A/H2B copy that is missing in +1 nucleosomes in relation to promoter DNA [55]. While the loss of the proximal H2A/H2B copy is

linked to transcription elongation and Pol II stalling, [55][64] the loss of the distal one enhances the passage of Pol II through the nucleosome eventually leading to the eviction of the histone core [55]. Removal of one H2A/H2B copy during transcription is facilitated by the histone chaperone FACT [65][66] but can also have a stabilizing effect on the histone core. Also other chromatin factors interact with hexasomes and can even re-position them [67][68][69]. One class of chromatin factors that remodel hexasomes are the so called ATP-dependent chromatin remodelers.

1.2 ATP-dependent chromatin remodelers

The packaging of DNA into nucleosomes controls the access to the DNA, thereby regulating fundamental cellular processes. While areas in between nucleosomes are accessible for chromatin factors, DNA becomes inaccessible when wrapped around nucleosomes. As a consequence, binding sites of transcription factors for example can be buried in the nucleosome preventing binding events [9][19]. Nucleosome positioning and re-organization are energy-driven processes that ensure stability of chromatin while allowing plasticity to adopt to environmental signals. To facilitate the (re-)organization of chromatin a family of enzymes evolved in eukaryotes that act on nucleosomal substrates. This enzyme family is called ATP-dependent chromatin remodelers – hereafter referred to as chromatin remodelers. All chromatin remodelers contain a Snf2-like protein that belongs to the SF2 superfamily of ATPases [70]. While the Snf2-like proteins differ in their N-terminal and C-terminal features, they all contain two RecA-like lobes – the N-lobe and C-lobe – forming together the motor domain (Figure 1.3b). In more detail, the two lobes enclose around nucleosomal DNA at remodeler specific super helical locations. While both DNA strands are enclosed, one is referred to as tracking strand which resembles the strand on which helicases translocate [19]. The second strand is denoted as guide strand [71][72][25]. The motor domain contains one ATP binding pocket that lays at the interface of the two RecA-like lobes. One of the lobes harbors a Walker A and Walker B motif that enable ATP binding and hydrolysis respectively. The second lobe has lost these motifs but contributes with an arginine finger motif to ATP hydrolysis [2]. To accommodate ATP hydrolysis, the two lobes are closed and reopened upon binding of the next ATP molecule. The caused movement of the motor lobes is translated into the translocation of the bound DNA. In more detail, ATP hydrolysis and more specifically the release of inorganic phosphate, causes a movement of the C-lobe in respect to the N-lobe leading to the insertion of the gating helix and displacing the tracking strand by one base pair [71]. This creates a transiently A-form of the DNA [26]. Upon binding of the next ATP molecule the B-form helix is reinstalled and the guide strand follows the translocation of the tracking strand. Thereby, during each ATP hydrolysis cycle one base pair of DNA is translocated alongside the histone core. The DNA entering the nucleosomal core, is referred to as entry DNA while the opposite end is the exit site or exit DNA. As a result, the DNA-histone interactions are broken and the nucleosomal substrate gets remodeled. The outcome of the remodeling reaction depends on the accessory subunits of the respective chromatin remodeler. They assemble onto the Snf2-like ATPase enabling thereby substrate specificity and facilitating only a certain reaction outcome. The accessory subunits can be either regulatory proteins or scaffolding ones that enable the assembly of even more proteins. Chromatin remodelers catalyze at least one of the following reactions: nucleosome assembly, histone editing, nucleosome sliding and positioning or the eviction of nucleosomes. Through these functions, chromatin remodelers facilitate in an energy-dependent manner the organization of chromatin and regulate the access of nuclear factors to DNA and histones. Thereby, chromatin remodelers are involved in fundamental processes like transcription activation and repression, DNA replication and DNA repair. Based on their Snf2-like motor domain, chromatin remodelers are subdivided into four families: 1) CHD (chromodomain he-

licase DNA-binding), 2) ISWI (imitation switch) 3) SWI/SNF (switch/sucrose non-fermentable) and 4) INO80/SWR1 (Inositol auxotrophy 80) [73] (Figure 1.3a). While the ISWI and CHD remodelers can act as single subunits but also form small complexes, SWI/SNF and INO80 complexes are of mega-Dalton size harboring a multitude of subunits. Each family harbors a set of chromatin remodelers while higher eukaryotes even have subtypes of the same remodeler which are either cell or tissue-specific or are only expressed during a specific developmental phase [74][27]. Usually these subtypes differ in their subunit composition as for example the human SWI/SNF complex called BAF. Despite the different reaction outcomes, it was suggested that chromatin remodelers act based on the same mechanism that is dictated by the Snf2-like motor domain [75][19]. The translocation of one base pair of DNA leads to a twist defect that diffuses along the nucleosomal DNA by rotation of the DNA. This explains how an individual Snf2-type motor can facilitate nucleosome sliding without the need of accessory subunits. Yet, also larger step sizes of nucleosome remodeling were observed through biochemical and single-molecule experiments [76]. In the latter case, delays between DNA entry and exit were observed speaking against a simple diffusion mechanism. This would be in line with a proposed motor-rotor-stator-grip mechanism were a grip element downstream of the motor domain prevents continuous twist diffusion [77]. Instead, multiple steps of DNA translocation are accumulated building up a DNA strain. Once the strain is sufficient, the twist diffusion continues past the grip element leading to nucleosome sliding and a net DNA translocation. Grip elements are the Arp5-Ies6 dimer of INO80 and Snf5 of the SWI/SNF complex. Alternatively, the DNA strain can result in the destabilization of the H2A/H2B H3/H4 interface resulting in histone eviction or exchange. Rotor and stator elements are the DNA and protein scaffolds of the multi-subunit chromatin remodelers respectively. Recent advances in cryoEM enabled the visualization of chromatin remodelers from all four families bound to nucleosomal substrates which inspired the rotor-stator-grip mechanism [77][13].

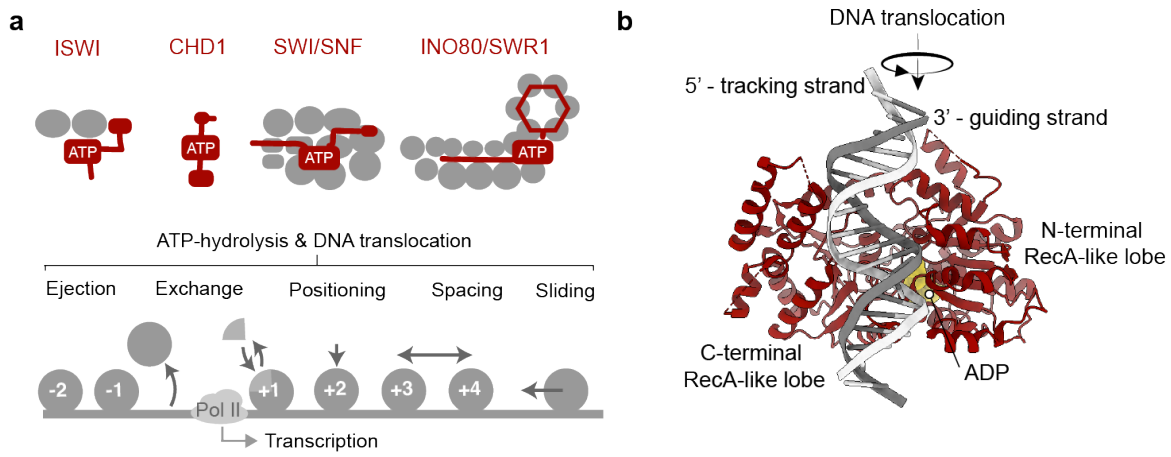


Figure 1.3: **ATP-dependent chromatin remodelers.** **a** Overview of the four families of ATP-dependent chromatin remodelers (top). Possible reaction outcomes of ATP-hydrolysis and DNA translocation (bottom). **b** Chromatin remodelers comprise a common motor domain of an C-terminal and N-terminal RecA-like lobe. ADP (in yellow) sits in the nucleotide binding pocket.

1.2.1 CHD and ISWI remodelers

Chromatin remodelers of the CHD and ISWI family are small compared to the SWI/SNF and INO80 complexes. They consist either only of the Snf2-type motor domain or a limited number of accessory subunits that modulate the motor's ATPase activity or facilitate specific substrate binding. CHD and ISWI remodelers are involved in nucleosome assembly and nucleosome sliding. During nucleosome assembly, histone chaperons deliver H2A/H2B dimers and H3/H4 tetramers to nascent DNA after DNA replication [78]. CHD and ISWI remodelers assist with the assembly of histone dimers and tetramers into mature nucleosomes [79][80]. Next, they implement regular spacing between nucleosomes to form a regular array [81][82][83][84] through nucleosome translocation. Thereby, CHD and ISWI remodelers are mostly involved in gene repression by aiding the formation of heterochromatin [84][84]. Apart from DNA replication, nucleosome positions are reinstalled after transcription as Pol II transcription can lead to either the loss of nucleosomes or change in their position [85][86][87].

ISWI family of chromatin remodelers

In the ISWI family the main activity arises from the motor subunit itself. However, the ISWI complexes contain up to four accessory subunits [88][81] that modulate nucleosome binding, linker DNA binding and sliding activity [89]. The two main complexes from yeast are ISW1 and ISW2 [90][91] while other prominent family members are the human NURF complex and the ACF complex that exists in human and flies. ISWI complexes are mainly associated with gene repression through its activity of creating nucleosomal arrays of evenly spaced nucleosomes. For example, yeast ISW1a creates a spacing of 30 bp of DNA between two neighboring nucleosomes [24][91][92][93][94] and places nucleosomes onto NFRs which blocks the assembly of the transcription machinery. In contrast to the yeast complexes, the human NURF complex is associated with the promotion of transcription [95]. ISWI complexes slide nucleosomes away from DNA ends by their directional sliding activity [88][96][93]. When the amount of DNA is limited, ISWI has a centering activity as for example shown for the ACF complex, leading to heterochromatin formation [84][97][98]. In addition, ISWI remodelers ISW1a and ISW2 position +1 nucleosomes [99][100]. On a molecular and mechanistic level, ISWI remodelers bind their nucleosomal substrates at SHL +/−2 [90] through the RecA-like lobes. The interaction is additionally driven by the sensing of linker DNA and binding of the N-terminal tail of H4 [101][102][103][90][93][89]. Binding of the linker DNA is facilitated through the C-terminal HAND-SANT-SLIDE (HSS) domain of the ISWI ATPase [104] that undergoes a conformational change during ATP hydrolysis [2]. ATPase activity is regulated by the autoinhibitory N-terminal region (AutoN) and the negative regulator of coupling (NegC) which are located next to the RecA-like lobes. NegC senses the length of linker DNA and couples ATPase activity to sliding [105] while the AutoN domain competes with the H4 tail for binding at the RecA-like lobes [102]. CryoEM enabled structural studies of ISWI motor domains bound to mono-nucleosomes [106][107] and the yeast ISWI1a bound to a di-nucleosome [108].

CHD family of chromatin remodelers

CHD mostly shares its architecture and function with the ISWI family [109] but comprises two chromodomains at its N-terminus that give rise to the name of the remodeler family [110]. It functions as a one-subunit remodeler in yeast [110] while the Chd1 remodelers form complexes in higher eukaryotes [111][112]. The different CHD subfamilies are involved in various types of reactions: in nucleosome assembly and spacing [83], in promotion of transcription by making promoter sites accessible [113] as well as nucleosome editing as for example the incorporation of the H3 variant

H3.3 [114]. One prominent member is the NuRD complex in metazoans which enables binding of repressors resulting in gene inactivation [112][115]. However, yeast Chd1 functions mainly in nucleosome assembly. Chd1 slides nucleosomes in a bi-directional manner [116]. It was also shown to slide hexasomes, however only in a one-directional manner indicating the involvement of H2A/H2B recognition for nucleosome sliding [68].

1.2.2 SWI/SNF remodelers

The SWI/SNF family covers complexes of mega-Dalton size containing a multitude of subunits besides the Snf2-type motor. Thereby, these complexes are larger compared to the ISWI and CHD families. SWI/SNF remodelers are the only chromatin remodeler complexes that have the ability to eject nucleosomes which facilitates access to DNA. As a consequence, they aid in the formation of nucleosome free regions present at gene promoter sites as described in the previous section [117]. In addition, SWI/SNF complexes can slide nucleosomes towards DNA ends but do not evenly space them [118][119]. Furthermore, these mega-Dalton sized complexes can make DNA binding sites accessible for transcription factors through loop formation that were before buried within the nucleosome [2]. Apart from the two typical RecA-like lobes, The Snf2-type subunit contains an HSA domain (helical/SANT-associated) with an adjacent post-HSA region, AT-hooks as well as a C-terminal bromodomain. The HSA helix is bound by the actin-related proteins Arp7 and Arp9 [120] in yeast or actin and Arp4 in the human complexes. The motor lobes engage SHL +/-2 likewise to the CHD and ISWI remodelers [121][106][72]. The SWI/SNF complexes comprise a modular architecture as visualized by cryoEM [122][121][123][124] of 1) the motor module 2) the Arp module and 3) the substrate recruitment model (SRM) (reviewed in [125]). The auxiliary subunits are part of the SRM module and are remodelers as well as species-specific enabling targeting to different sets of genes [126]. The Arp module bridges the motor and SRM modules but also functions in modulating motor activity through the interaction of the post-HSA helix with the motor lobes [127].
Prominent family members of the SWI/SNF family are the two yeast remodelers SWI/SNF and RSC that are paralogues and assemble around the Snf2 and Sth1 motor proteins respectively. The yeast RSC complex is not only larger as the SWI/SNF complex (17 versus 12 subunits) but also at least ten times more abundant [128]. In addition, SWI/SNF does not control cell cycle dependent genes and focuses on Pol II dependent promoters. Despite their overlap in function, the SWI/SNF and RSC remodelers do not act on the same set of genes. Furthermore, RSC has a sliding activity on +1 nucleosomes thereby moving it away from the TSS [129]. Comparable remodelers exist in human where the BAF complex resembles SWI/SNF and and pBAF shows similarity to the RSC complex [130]. Further, SWI/SNF complexes in higher eukaryotes underlie a combinatorial subunit composition resulting in tissue and developmental stage specific complexes [131]. Finally, the pivotal role of SWI/SNF remodelers in gene transcription makes them the most abundantly dysregulated targets in cancer [132].

1.3 INO80 family of chromatin remodelers

As the INO80 family of chromatin remodelers will be the focus of this thesis, I am introducing them in a separate chapter. First, focusing on their similarities followed by dedicated chapters on the INO80 and SWR1 complexes respectively.

1.3.1 Members of the INO80 family and their function

In yeast, the INO80 family of chromatin remodelers comprises two members, INO80 itself as well as SWR1. Despite their similar architecture, INO80 and SWR1 perform distinct nucleosome remodeling reactions. INO80 has an established function in sliding and positioning of nucleosomes as well as a debated role in exchanging the histone variant H2A.Z by its canonical counterpart H2A [76][133][134]. SWR1 incorporates the histone variant H2A.Z into nucleosomes and catalyses thereby the counter-reaction of INO80 [135][136][137]. It is of note that yeast possess in addition to SWR1 the histone acetyltransferase complex NuA4, while higher eukaryotes including humans possess the Tip60/p400 complex. The latter is a 'fusion' complex of the two orthologous yeast complexes SWR1 and NuA4. Yet, the human SRCAP complex, which resembles the yeast SWR1 complex, can catalyse histone exchange independently of the Tip60 histone acetyltransferase module. Together Tip60/p400 as well as SRCAP contribute to the pool of H2A.Z containing nucleosomes *in vivo*[138][135][139][51]. It remains however unclear how their activity is distributed in different cell types and developmental stages [140] and whether there are additional variants of INO80-family remodeler complexes in mammals. Recent advances in cryo electron microscopy enabled the determination of various chromatin remodelers of the INO80 family in complex with mono-nucleosomes and in different nucleotide states: INO80 (human) [141], INO80 (*C. t.*) [77], SWR1 (yeast) [142] SRCAP (human) [143] and also NuA4 (yeast) [144][145][146]. The structures provide a first structural framework for the common features of the INO80 family.

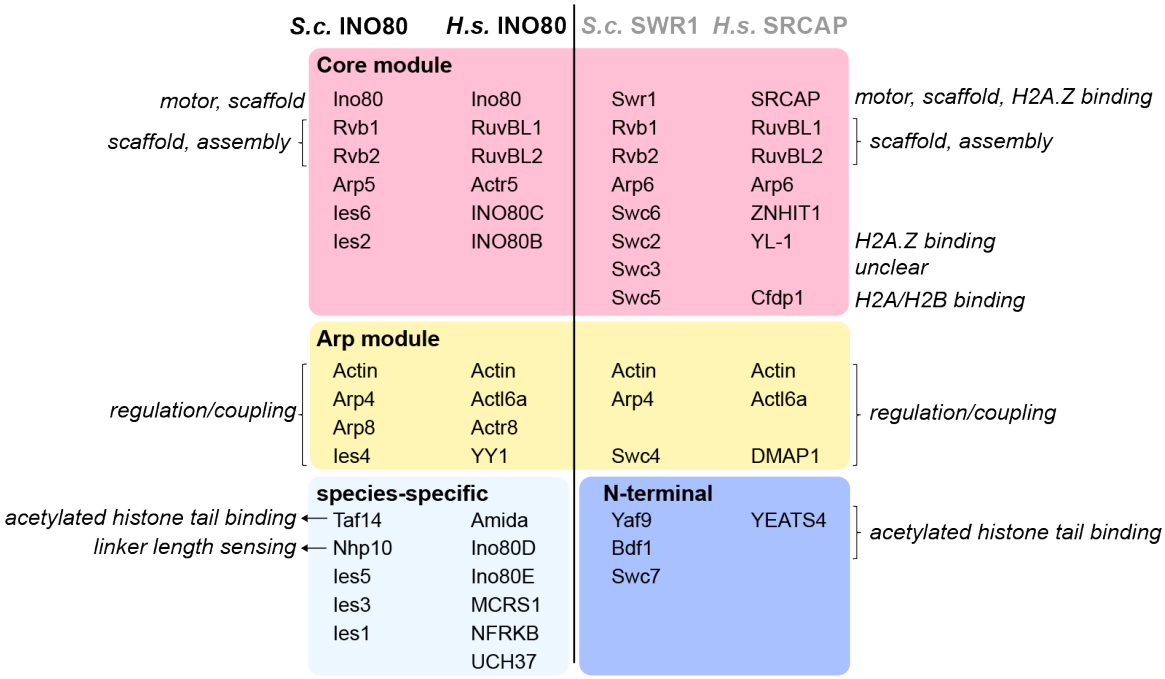


Figure 1.4: **Subunit composition of yeast (*S. c.*) and human (*H. s.*) INO80 and SWR1/SRCAP complexes**. Subunits belong to the Core module (red), Arp module (yellow) or species-specific/N-terminal modules (blue) of the complexes. The Core module and Arp module subunits of INO80 and SWR1/SRCAP share a common architecture and are hence listed next to each other. Functions of the respective subunits are stated on the left for INO80 and on the right for SWR1. The role of the species-specific subunits Taf14 and Nhp10 are only known for yeast.

1.3.2 Common features of the INO80 family

The INO80 and SWR1 complexes are evolutinary conserved 'non-identical twins in chromatin remodeling' [147]. Here, I am mostly focusing on the yeast complexes. On a molecular level, INO80 and SWR1 share a modular architecture with a few shared and some remodeler-specific subunits. Among these shared subunits are Rvb1, Rvb2, actin and the actin-related protein 4 (Arp4) – the nomenclature differs in metazoans (see Figure 1.4). The Snf2-type motor proteins are mostly named after the complex itself but differ in writing between yeast and metazoans [19]. In yeast, the Snf2-type motor proteins are written in small letters: Ino80 for the INO80 complex, Swr1 for the SWR1 complex and Eaf1 for NuA4. In contrast, the motor subunits remain in capital letters in metazoans e.g. INO80 subunit within the INO80 complex. Overall, the Ino80 and Swr1 subunits function as a scaffold for the remaining subunits, giving rise to the modular architecture of the complexes consisting of the Core, Arp and species-specific modules. I will write the Core module in capital letters to distinguish it for example of core features.

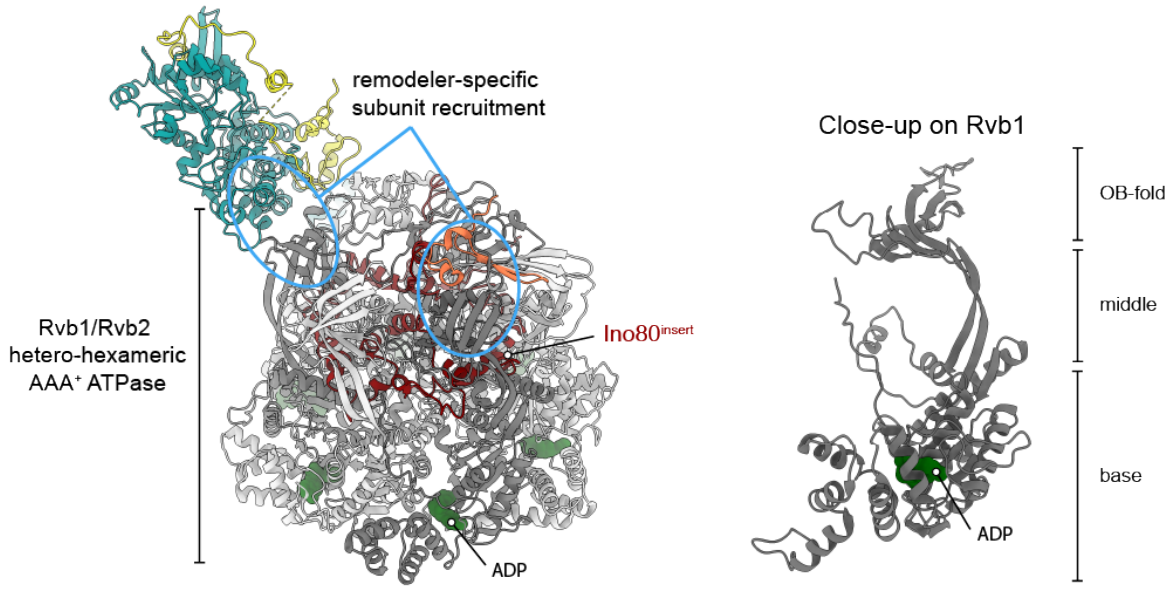


Figure 1.5: **Common features of the INO80 family**. Chromatin remodelers of the INO80 family commonly comprise an insertion domain within their Ino80 or Swr1 motor subunits onto which the Rvb1/Rvb2 hetero-hexameric AAA$^+$ ATPase assembles. Rvb1 and Rvb2 comprise a nucleotide binding pocket at their base that is occupied by an ADP molecule. Further Rvb1/Rvb2 are composed of a middle domain and OB-fold (illustrated on the right). The OB-fold functions as a recruitment platform for remodeler-specific subunits (e.g. in turquoise, yellow and orange). Based on PDB 8AV6.

The Ino80, Swr1 (SRCAP and p400 in humans) motor subunits share common features based on sequence alignments [73]. The two RecA-like lobes form the motor domain. Uniquely, the Snf2-type proteins of the INO80 family contain a long insertion domain of about 300 residues that intersects the C-lobe [73]. The ATPase lobes and insert of Ino80 and Swr1 are part of the Core module, together with the AAA$^+$ ATPase Rvb1/Rvb2 hetero-hexamer and some remodeler-specific subunits (refer to Figures 1.6a, b and 1.8). The Core module subunits facilitate binding to the nucleosomal substrate. The hetero-hexamer has the shape of a barrel that is formed by a ring of alternating Rvb1 and Rvb2 copies comprising a three-fold symmetry [148]. The Rvbs are structured into three domains: the AAA$^+$ at its base, the middle domain and the OB fold on top. Even though each of the

Rvb1 and Rvb2 subunits comprises an ATPase domain, their activity is not essential for nucleosome remodeling of neither INO80 nor SWR1 [141][142]. However, the insert of INO80 was shown to increase the ATPase activity of the Rvbs by 16-fold indicating that the Rvbs serve as a chaperone for complex assembly [149]. Post-assembly, the Rvbs harbor an ADP as shown by cryoEM structures [77][141][150] highlighting that the Rvb1/Rvb2 hetero-hexamer functions as a scaffold once the chromatin remodeler is assembled. The three-fold symmetry of the Rvb1/Rvb2 heterohexamer is broken by the Snf2-type insert domain which shares little of its sequence when comparing different remodeler types or the same remodeler across different species [147]. The insert determines which remodeler-specific subunits are recruited to the Core module. Accordingly, INO80-specific subunits can be exchanged by SRCAP-specific ones *in vivo* by interchanging the corresponding insert domain [151]. Either the insert directly recruits remodeler-specific subunits or the insert induces structural changes in the Rvbs OB domains resulting in subunit recruitment [77] (Figure 1.5).
The N-terminus of Ino80 and Swr1 harbors a helicase-SANT-associated (HSA) helix which assembles actin and Arp4 that share an actin fold, forming together the Arp module. Notably, also the yeast SWI/SNF-family complexes contain the HSA helix but recruit Arp7 and Arp9 [152][153], while SWI/SNF-family complexes from higher eukaryotes also recruit Arp4 (human BAF53) and nuclear actin. The post-HSA helix of Ino80 and the Swr1-Z domain of Swr1 are located in between their respective HSA and ATPase motor domains. Of note is, that Eaf1 from NuA4 contains as well an HSA helix. In addition, the HSA helix binds to DNA, which might interact with linker or promoter DNA in both INO80 [154][150] and the RSC complex [155]. Arps were shown two decades ago to be essential for chromatin remodeling [156]. Furthermore, SWR1 and NuA4 share the Swc4 and Yaf9 subunits. The very N-terminal region of Ino80 and Swr1 assembles further remodeler-specific and in case of INO80 species-specific subunits [13][157].

1.4 INO80

Nucleosome remodeling by INO80 has implications in fundamental processes like DNA transcription, replication and repair [158]. In regard to this, depletion of INO80 is lethal in mice in early embryonic stages [159] as cells fail to repair double stranded DNA breaks (DSBs). The linkage of INO80 to active gene transcription is hijacked by tumor cells through over-expression of INO80 [160] which is generally linked to tumor progression. In turn, down-regulation of INO80 was shown to decrease cell growth and tumorigenesis [160].

1.4.1 Functions of the INO80 complex

Originally, INO80 was identified in a yeast screening in which a set of proteins including INO80 was mutated and kept in media without inositol, introducing stress for the cells. Upon INO80 mutation, a subset of genes failed to be activated [161] giving raise to the complex name: Inositol auxotrophy 80. INO80 also plays a role in other metabolic pathways including glycolysis and the respiratory chain reaction. In addition, INO80 controls osmotic-regulated genes [162] and the expression of its own subunits [163]. At the core of the INO80 functions lays its activity for sliding nucleosomes. Thereby, INO80 regulates access to the genome as DNA sequence features become accessible or inaccessible upon nucleosome re-positioning. This impacts binding of nuclear factors like transcription factors or large chromatin complexes explaining the roles of INO80 in DNA transcription, replication and repair. On a molecular level, INO80 centers mono-nucleosome *in vitro* and spaces nucleosomes evenly [156][164] with an equal length of extra-nucleosomal DNA of approximately 30 bp of DNA [164]. Binding of INO80 requires at least 20 bp of DNA while INO80 remodeling occurs at preferentially

longer stretches of DNA linkers [164]. More recently, INO80 was found to also slide hexasomes [67]. Hexasomes were remodeled 60-times faster compared to nucleosomes which made the authors speculate that hexasomes are even the preferred substrates and could be an intermediate of the sliding reaction.

Role of INO80 around gene promoters and in transcription

It is presumably the long stretch of accessible DNA at NFRs that tethers INO80 to transcription start sites. Additionally, INO80 reads DNA shape features of promoter DNA [165][166][167] and can be recruited by Pol II at sites of stress genes [168]. INO80 positions the −1 and +1 nucleosomes around NFRs as shown by *in vitro* and in *in vivo* studies [34][35][24][36][37] in cooperation with ISWIa [24]. Correct positioning of nucleosomes around transcription start sites is pivotal for transcription activation [32][33]. In line, 90 % of promoter sites are occupied by INO80 in yeast [37]. Downstream nucleosomes are phased by INO80 into regular gene arrays based on the +1 nucleosome as well as barrier factors like Reb1 and DNA ends (as for example present at DNA breaks) [36][169]. The absence of INO80 leads to activation or inactivation of genes at almost the same level [136][163]. A similar result is obtained upon depletion of the INO80 subunits Arp5-Ies6 [170] underlying the role of INO80 in regulating access to the genome especially at transcription start sites. In addition, INO80 averts bidirectional transcription at promoter sites [171] and has implications in the control of long non-coding RNA transcription [172].

Other functions of INO80

Besides the pivotal role of INO80 in transcription and shaping the chromatin architecture around promoter DNA, INO80 is involved in DNA replication. INO80 was found at origins of replication and seemingly has a role in replication elongation as well as restarting stalled replication forks [173]. In addition, INO80 is involved in DNA repair as its deletion in yeast made the cells more sensitive to environmental stresses like UV radiation [174]. In line with this finding, INO80 is recruited to double-strand breaks (DSB) of DNA that occur upon UV radiation [163]. Its role at DSB sites is to create access for the DNA repair machinery to the site of DNA damage. In this light, INO80 removes or even degrades nucleosomes at DSB sites [175][176][177][178] and also assists in the eviction of RNA Polymerase II [179][180]. Of note is that INO80 performs similarly at replisomes by removing nucleosomes [158] or Pol II molecules that collided with the replisome [181]. More generally, INO80 increases chromatin mobility during DSB [182][183] which is required for tethering of the DSB to the nuclear pore complex or Mps3 [184] aiding in DNA repair.
Besides nucleosome sliding, INO80 was found to have a histone exchange activity. INO80 removes the H2A.Z/H2B variant and displaces it with canonical H2A-H2B dimers thereby catalysing the counter-reaction of the SWR1 complex [76]. However, this function is debated in the field and could not be reproduced [133][134]. In regard to this, the depletion of INO80 does not change the occupancy or distribution of H2A.Z containing nucleosomes [185]. Functionally, INO80 might remove H2A.Z from +1 nucleosomes to enable Pol II to transcribe through it [186], however Pol II was also found to remove H2A.Z/H2B dimers on its own. Further investigations are therefore required also as it is unclear how the same remodeler can enable so different reactions as nucleosome sliding and histone exchange.

1.4.2 Complex composition

INO80 is a 15-subunit complex in yeast [174][187] and 16-subunit complex in human [188]. Its modular architecture is composed of three modules: The Core module, the Arp module and the species-

specific module (Figure 1.6a, b). However, the Core module is also referred to as the C-module as its subunits assemble on the C-terminus of the Ino80 subunit. Likewise, the species-specific module is termed N-module. The subunits assemble directly or indirectly onto the Ino80 subunit. Therefore, the remodeler architecture was studied by detecting which subunits are lost upon deletion of a series of Ino80 subunit regions [188]. Ultimately, structures of the INO80 core bound to mono-nucleosomal substrates were determined by cryo electron microscope (human: [141], *C. t.*: [77][150]) in addition to the Arp module (X-ray crystallography: [154], cryoEM: [150]).

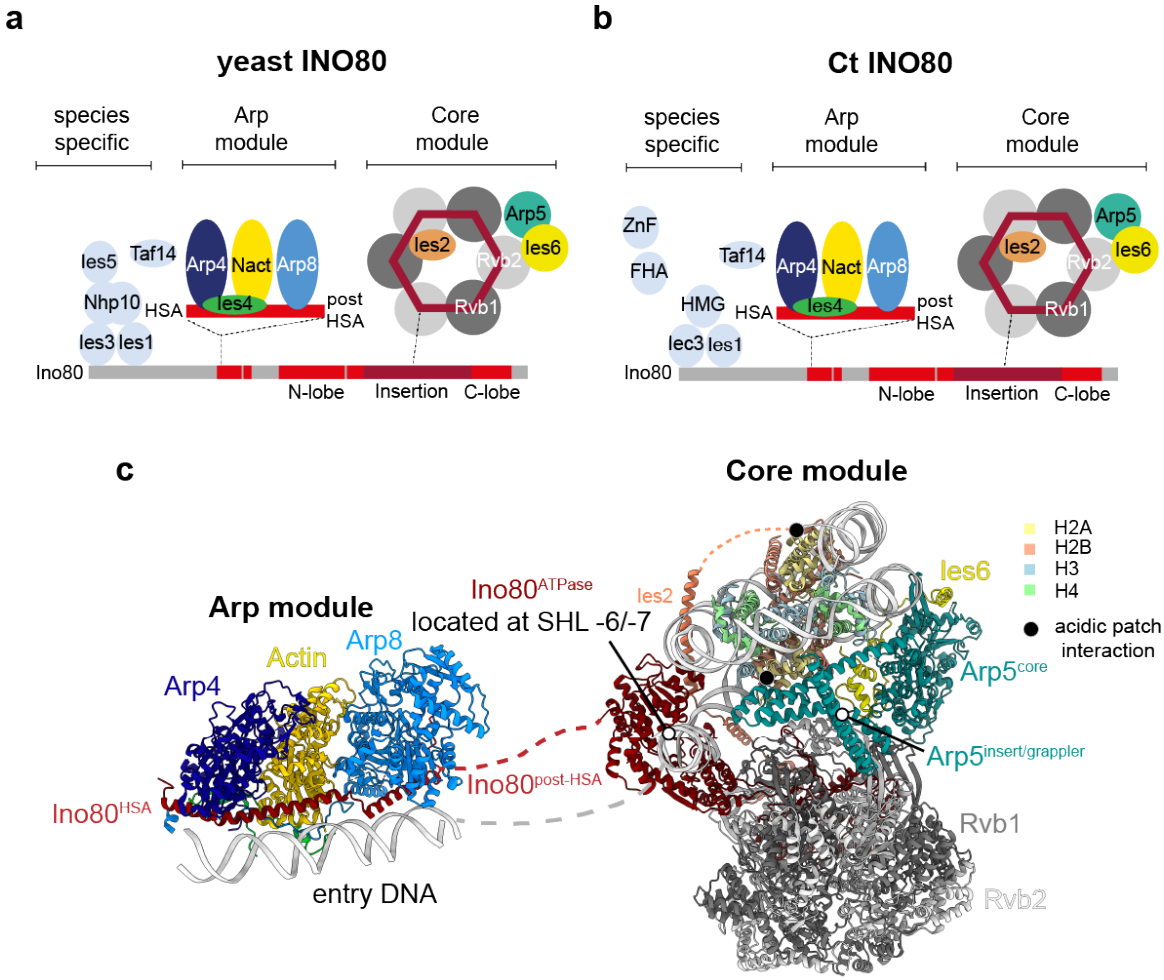


Figure 1.6: **Complex composition of INO80. a**. Subunit composition of yeast INO80 shows a modular architecture comprising a Core module, Arp module and species-specific module. **b** The same holds true for *C. t.* INO80. The Core and Arp modules are conserved among different species, while the subunit composition of the species-specific module differs. **c** Structural knowledge on the Core (PDB 8AV6) and Arp module of INO80 (PDB 8A5P). The Ino80^{ATPase} sits at SHL postion −6/−7 close to the entry DNA. The Arp module sits upstream on the entry DNA connected by the post-HSA helix that sits in between the Ino80 motor and Ino80 HSA helix. The core module contacts the acidic patch (black dot) with the Ies2 and Arp5 subunits.

The Core module

The core of the INO80 complex is composed of the Rvb1/Rvb2 hetero-hexameric AAA$^+$ ATPase, the Ino80 motor lobes and its insert domain that folds into the Rvb1/Rvb2 barrel as well as the sub-

units Ies2, Ies6 and Arp5 (cryoEM structure of the core in Figure 1.6c). Together, all core subuntis are essential for nucleosome binding and DNA translocation of INO80. Unlike the small chromatin remodelers CHD and ISWI, INO80 contacts the nucleosome at multiple sites. The main contact is enabled by the two RecA-like lobes that bind the nucleosome at SHL −6/−7 [76][77][141][150] – a location unique in the field of chromatin remodelers as the other known remodelers recognize instead SHL +/−2. Thereby, the motor lobes sit close to the entry DNA pumping DNA into the nucleosome. Downstream sits the Arp5^{core}/Ies6 dimer at SHL −3 which is also referred to the grip [77] of INO80. Together, Arp5 and Ies6 tether INO80 to the nucleosome and were shown to be essential for INO80 recruitment and coupling the ATPase activity of Ino80 to nucleosome sliding [189][190][191]. The core of Arp5 contains the actin-fold which is intersected by an insertion that is termed grappler [77][150], forming a cross close to the dyad of the nucleosome. While this represents the closed conformation of the grappler, it was also captured in an open, more parallel state [77]. The grappler and more precisely its foot helix contacts the H2A/H2B acidic patch of the proximal nucleosome face with arginine residues commonly used for acidic patch recognition [14]. This interaction is essential for nucleosome sliding but not ATPase activity of INO80 [77]. A second contact of the acidic patch is facilitated by Ies2 – however, at the distal site of the nucleosome (Figure 1.6c).

The Arp module

The Arp module binds upstream of the motor lobes to extra-nucleosomal DNA via the HSA helix of the Ino80 subunit (Figure 1.6c). The HSA helix is connected through the so called post-HSA helix to the Ino80 ATPase at the Core module. Onto the HSA helix assembles the Arp module subunits that comprise an actin-fold, namely nuclear actin, Arp4 and Arp8 and in addition Ies4 (yeast) or YY1 (human) [150]. While the Arp module is not required for nucleosome binding, it is essential to couple the ATPase activity of the INO80 complex to DNA translocation [187]. The coupling is likely enabled through the post-HSA helix that translates binding of extra-nucleosomal DNA via the Arp module to the Ino80 motor domain [71][150]. Thereby, the Arp module functions in reading DNA-length [192][164] as well as DNA shape features [36] which is relevant for creating evenly spaced nucleosome arrays as well as positioning of +1 nucleosomes. Of note is that the human transcription factor YY1 takes the place of Ies4 in humans which raises the possibility that human INO80 recognizes specific DNA sequences through YY1 [150][193].

The species-specific module

The subunits of the species-specific module assemble onto the N-terminus of the Ino80 subunit. As the name suggests, its subunit composition varies among species. This becomes apparent when yeast and *Chaetomium thermophilum* INO80 are compared (Figure 1.6a and b). The latter was used in this thesis for structural studies. In more detail, yeast INO80 contains the yeast specific subunit Nhp10 which is also eponymous for the whole module in yeast – the Nhp10 module. It was described to contribute to determining the inter-nucleosome distancing [194]. Taf14, that is present in both the yeast and *C. t.* complexes, is the only subunit of INO80 described to have a classical histone tail binding motif. Other ATP-dependent chromatin remodelers including SWR1 contain motifs like chromo or bromo domains [158]. Taf14 harbors a YEATS domain that recognizes acetylated or crotonylated histone H3K9 [195] which is correlated with active gene transcription. Overall, the function of the species-specific module however remains largely elusive and might also differ between species.

1.4.3 Mechanism of nucleosome translocation

The Ino80 motor lobes bind close to the entry DNA, placing the Arp module upstream onto the extra-nucleosomal DNA. The INO80 motor translocates likewise to other chromatin remodelers DNA in 1 bp steps. In case of INO80 the DNA is pumped towards the Arp5 grip that is hold in place by the Rvb1/Rvb2 stator element. Pumping of DNA by INO80 does not directly lead to a net DNA translocation around the nucleosome as observed by footprinting studies [76]. Instead, 7-10 bp are accumulated. This results into DNA underwinding and builds up a DNA strain in-between the motor lobes and the Arp5 grip [77]. In this time binding of the Arp module presumably prevents back slippage of the DNA [13]. Once the DNA strain is too large, a conformational rearrangement of the grip domain leads to net DNA translocation. Experimental evidence by single molecule or structural studies are however lacking as one would need to follow the INO80 reaction over multiple ATP hydrolysis steps. The Arp module and/or Nhp10 module in yeast sense DNA-length and DNA shape features [192][164][36] on the entry DNA which resembles for example promoter DNA when INO80 remodels the +1 nucleosome. Once the upstream nucleosome or a favored DNA shape is reached [194], a conformational change in the post-HSA decouples ATP hydrolysis from DNA translocation. This stops the reaction and INO80 releases itself from the nucleosome by one more round of ATP hydrolysis [196]. Of note is that human INO80 functions as a dimer [197] but it remains unclear how a nucleosome can accommodate two copies of such a mega-Dalton sized chromatin remodeler. Less is known on the mechanism of histone exchange by INO80 – especially as no histone chaperons are described for the complex. It is also unclear how INO80 can perform such different reactions and how it differentiates between the two reaction types. However, the described DNA strain is built up close to the proximal H2A/H2B copy resulting in a destabilization of the H2A/H2B-DNA interface [76] which could also explain the eviction of H2A.Z/H2B dimers.

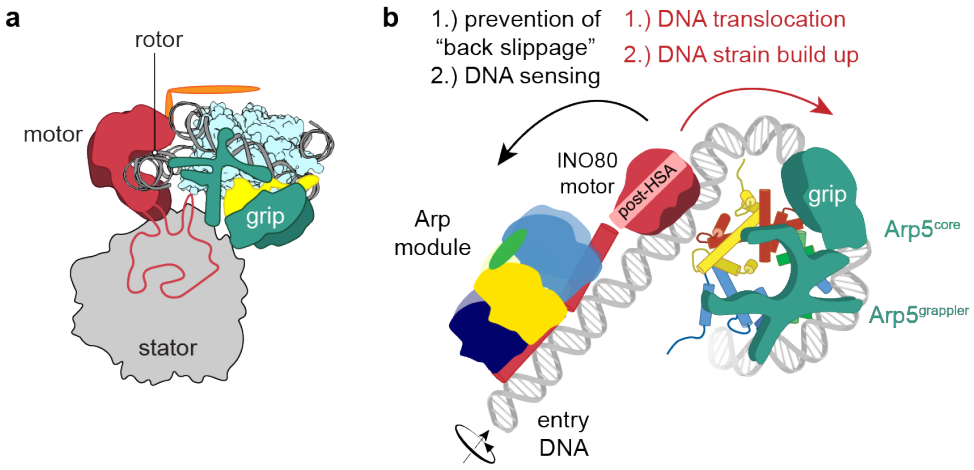


Figure 1.7: **Mechanism of nucleosome sliding by INO80. a** Motor (Ino80), rotor (DNA), stator (Rvb1/Rvb2 hetero-hexamer), grip (Arp5) architecture of INO80 (based on [13]). **b** Ino80 motor uses each ATP hydrolysis cycle to pump DNA in one base pair steps towards the grip. The latter prevents direct DNA translocation. Instead multiple DNA translocation steps are accumulated which builds up DNA strain in between the motor and grip domains. The upstream sitting Arp module prevents back slippage and additionally senses the entry DNA. Once the DNA strain is sufficient, a large conformational change occurs releasing the DNA strain and resulting in a net DNA translocation around the nucleosome (nucleosome sliding).

1.5 SWR1

1.5.1 Functions of SWR1 and H2A.Z

H2A.Z is essential for mammal and drosophila development [49][50]. Yet, SWR1 does not incorporate copies of H2A.Z/H2B randomly into the genome. Instead, SWR1 was shown to have a specificity for +1 nucleosomes at gene start states [53][11], at replication of origins [198] and sites of DNA double strand breaks [199][200]. Thereby, SWR1 plays a pivotal role in DNA transcription, replication and DNA repair and is consequently miss-regulated in cancer including uterine leiomyoma [201]. SWR1 preferably binds to nucleosomes with a long extra-nucleosomal stretch of DNA, tethering it likely to promoter DNA [202]. At its nucleosomal substrates, SWR1 incorporates the histone variant H2A.Z in a stepwise manner by removing the canonical H2A/H2B dimer and replacing it with H2A.Z/H2B.

Role in RNA transcription

SWR1 incorporates H2A.Z specifically at promoter sites – more specifically into the +1 nucleosomes which is located directly downstream of the NFR [202][203][204]. More precisely, H2A.Z/H2B dimers are found in +1 nucleosomes of active and poised genes [52][53] and is used as a switch to facilitate active gene transcription [51]. H2A.Z containing nucleosomes are characterised by weaker dimer-tetramer interactions within the histone core making the nucleosome overall less stable [48]. As a consequence RNA polymerase II (Pol II) can transcribe through H2A.Z containing +1 nucleosomes [185][205]. Even more so, H2A.Z plays a role in Pol II pausing [206][186], Pol II elongation [207] and co-transcriptional splicing in yeast [208]. Overall, H2A.Z occupancy in +1 nucleosomes enables Pol II initiation and its transition from initiation to transcription elongation.
What targets SWR1 to its nucleosomal substrate on a molecular level is incompletely understood. In part, the NFR and acetylation marks target SWR1 to +1 nucleosomal explaining its role in DNA transcription. In more detail, yeast SWR1 harbors the Yaf9 acetylation reader subunit that recognizes H4K9ac and H3K27ac [209][210]. In addition, Bdf1 reads acetylated histones [211]. In humans the Yaf9 homologue associates to H3K14ac and H2K27ac on active genes [212]. The NFR can be scanned by Swc2 as shown by diffusion experiments using optical tweezer technology [213]. On the other hand SWR1 evidently reads out DNA sequence information [214] targeting SWR1 also to intragenic target sites [215]. If sequence information on promoter DNA has a role in SWR1 recruitment to +1 nucleosomes, is however not clear. Lastly, the correct positioning of the +1 nucleosome might play a role in substrate recognition by SWR1. Other chromatin remodelers like INO80 have been shown to function in +1 nucleosome and phasing of downstream nucleosomes and might cooperate with SWR1 to facilitate gene transcription [167][169][216].

SWR1 in DNA repair and other roles

SWR1 incorporates H2A.Z/H2B dimers at the nucleosomes around DNA double strand breaks. H2A.Z functions here as a marker for the recruitment of the DNA repair machinery [199][200]. In yeast, the presence of H2A.Z leads to an increased activity of exonuclease I during DNA repair [217]. Therefore, the lack of H2A.Z and SWR1 results in a higher mutation rate. In addition, SWR1 promotes loading of the non-homologous end-joining (NHEJ)-Ku70/80 complex in yeast [218] and mammals [199].

H2A.Z is enriched in nucleosomes located at replication origins thereby regulating replication initiation and timing of DNA replication [198]. On top of that, H2A.Z was shown to have a pivotal role in chromosome segregation [219].

1.5.2 Complex composition

While H2A.Z was already discovered in 1980 in mouse, it took yet another 20 years to discover the protein complex that catalyses the underlying histone exchange reaction [135][136][137]. The discovery of SWR1 in yeast was complemented by the discovery of its human homologue termed SRCAP complex [220]. In the following two decades SWR1 and SRCAP as well as the H2A.Z/H2B exchange reaction were studied using a mixture of biochemical, structural and single molecule tools. Here, I focus on the yeast SWR1 complex that contains 14 subunits. The complex assembles onto the Swr1 protein creating a similar modular architecture like INO80 [13]. In that way, SWR1 is composed of the Core, Arp and N-terminal module (Figure 1.8). The complex architecture was studied by different biochemical approaches including Swr1 truncation experiments [221] and crosslinking mass spectrometry [222]. More recently, a cryoEM study on SWR1 bound to a mono-nucleosome shed light on the molecular architecture that mostly resembles the INO80 nucleosome complex [142] (Figure 1.8). Yet, the remodeler specific subunit composition of INO80 and SWR1 drives different reaction outcomes on their nucleosomal substrates through a thus far unknown mechanism. The published cryoEM structure comprises only subunits from the Core module that interacts through multiple subunits with the nucleosome.

The Core module

Likewise to the INO80 complex, the Core module of SWR1 assembles onto the Rvb1/Rvb2 hetero-hexamer. The Swr1 subunit is anchored into the barrel shaped Rvb1/Rvb2 sub-complex through its insertion domain that intersects the two RecA-like motor lobes. Three components that sit on top of the Rvb1/Rvb2 hetero-hexamer contact the nucleosome: The Swc6/Arp6 dimer, Swc2 as well as the RecA-like motor lobes. The latter encloses nucleosomal DNA at SHL +/-2 likewise to the chromatin remodelers from the ISWI, CHD and SWI/SNF families [223]. In contrast, INO80 that belongs to the same remodeler family, binds to SHL −6. SWR1 exchanges the H2A/H2B copy that lays on the proximal face of the nucleosome in relation to the binding site of the Swr1 motor lobes. Deletion of Arp6 leads to the loss of histone exchange activity [221][224]. Furthermore, Swc6 was captured in the cryoEM structure close to the motor proximal H2A/H2B, thereby undermining the role of the Arp6/Swc6 in histone exchange. Further, Swc2 folds from its Rvb1/Rvb2 anchor towards the motor lobes and onto the acidic patch of the distal H2A/H2B copy which potentially functions as an anchor during nucleosome remodeling. In addition, Swc3 is part of the Core module as its assembly to SWR1 depends on the presence of Swc2 but structural data is missing.

The Arp module

The Arp module assembles onto the HSA helix of the Swr1 subunit which is located C-terminal of the motor lobes. It harbors the subunits Arp4, actin and Swc4. A dimer of actin and Arp4 binds to the HSA helix with their respective barbed ends as visualized by X-ray crystallography [225]. Lin et al. 2017 [221] showed the presence of a second copy of actin that takes the place of Arp8 in INO80 by Swr1 truncation experiments. In addition, Yaf9 is possibly part of the Arp module as the deletion of Arp4 in yeast strains, abolishes the presence of Yaf9 in the SWR1 complex. Yaf9 harbors a YEATS domain that can bind to acetylated histones. By reading H4K9ac and H3K27ac [209][210] SWR1 can be targeted to +1 nucleosomes as mentioned in the section before. Besides that, the Arp module

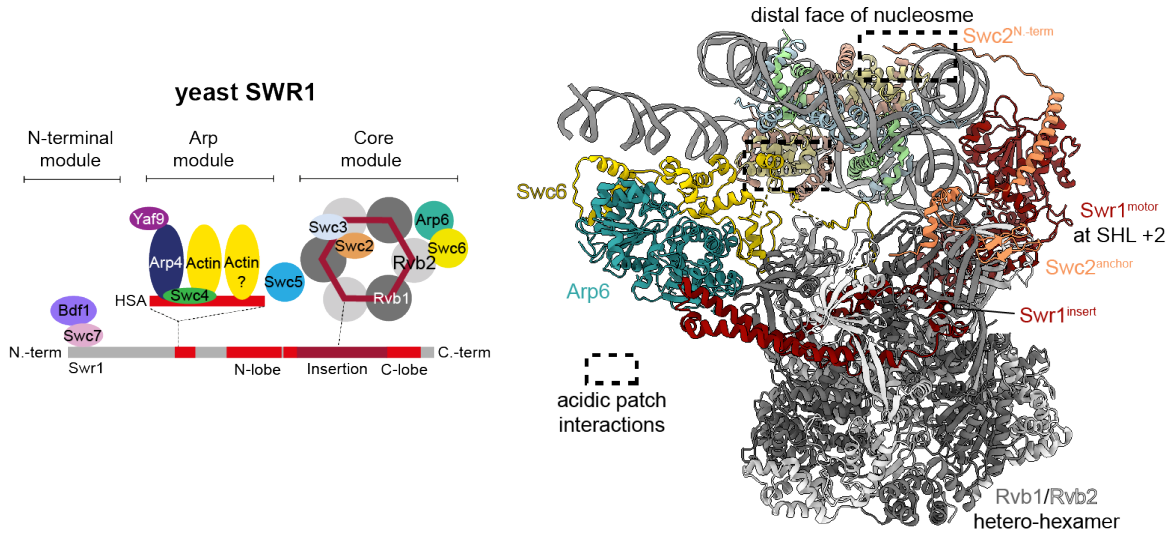


Figure 1.8: **Complex composition of yeast SWR1** . *Left*: Subunit composition based on crosslinking mass spectrometry data [222]. It comprises a Core, Arp and N-terminal module. *Right*: Structural knowledge of the Core module of yeast SWR1 bound to a mono-nucleosome (PDB 6GEJ) [142]. The Swr1 motor sits at SHL+2. Architecture resembles the INO80 structure. The two acidic patch interactions (dotted square) are facilitated by the N-terminus of Swc2 and the Swc6 subunit.

functions in histone exchange as the deletion of Arp4 and Swc4 abolishes histone exchange activity of SWR1 [221]. However, it remains elusive by which mechanism and through which interactions the Arp module contributes to histone exchange. Finally, the Swc5 subunit can be placed in between the Arp and Core modules as it crosslinks close to the Swr1 HSA helix and C-lobe as well as Yaf9 and actin [222]. Swc5 functions as a histone chaperone by binding H2A/H2B as visualized by X-ray crystallography [226]. A recent low resolution cryoEM structure places Swc5 furthermore onto the acidic patch of the nucleosome. This interactions is essential for successful histone exchange as shown with mutants and activity assays [227].

N-terminal module

N-terminal of the Arp module lays the N-terminal module. It comprises the N-terminus of Swr1, Swc7 as well as Bdf1. Bdf1 function in addition to Yaf9 as reader of acetylated histone tails. In addition, its presence is dependent on Swc7 [228] indicating that it rather interacts with Swc7 and not with Swr1. The N-terminal module of INO80 was also termed species-specific module [13] as its subunit composition can vary. It remains unclear if this is also the case for the SWR1 complex.

1.5.3 Mechanism of histone exchange

Binding of SWR1 onto its nucleosomal substrate at SHL+/-2 in its ATP bound state leads to DNA unwrapping prior to ATP hydrolysis [142] resulting in the destabilization of the H3/H4-H2A/H2B interface probing the nucleosome for histone exchange. The entry DNA close to the Arp6/Swc6 grip is lifted off the histone core [142][147]. Unlike other chromatin remodelers no net DNA translocation is required for exchanging canonical H2A/H2B with H2A.Z/H2B [142][143][223] . Instead ATP hydrolysis is used for unwrapping and re-wrapping of the DNA from and around the histone core [142][229][230]. The step of histone exchange is supported by a variety of histone chaperons as for

example the H2A.Z/H2B specific histone chaperone Chz1. In fact Chz1 bound H2A.Z/H2B represents the major substrate of yeast SWR1 [231][232]. The incoming copy of H2A.Z/H2B is then taken over by the Swc2-Z and the Swr1-Z domains respectively [233][234]. Crystal structures resolved the binding of the short helical features of the Swc2-Z and Swr1-Z domains when bound to H2A.Z/H2B dimers respectively. While Swc2-Z is located at the very N-terminus of the Swc2 subunit, Swr1-Z lays between the N-lobe and the HSA helix of Swr1. Binding of H2A.Z/H2B through the Swr1-Z domain stimulates the ATPase activity and thereby histone exchange [234][235]. A recent structural study on human SRCAP [143] showed that the Swc6 homologue binds to the acidic patch of the proximal H2A/H2B copy in relation to the motor domain. Upon ATP hydrolysis the Swc6/Arp6 arm undergoes a conformational rearrangement, potentially facilitating the removal of the H2A/H2B dimer. The outgoing H2A/H2B dimer is taken over by the Swc5 subunit before it is removed from the complex by the histone chaperone Nap1 [236][229]. Overall, the two canonical copies of H2A/H2B are exchanged in a stepwise manner by a monomer of SWR1 with H2A.Z/H2B [221][235][229]. This creates a mixed AZ intermediate nucleosome that contains one copy of H2A.Z/H2B and H2A/H2B respectively. If it is the same copy of SWR1 or another SWR1 molecule that exchanges the second histone dimer is unclear and might be subjected to stochastic events. Regardless, the final product is a ZZ nucleosome that contains two copies of H2A.Z/H2B [214]. On a mechanistic level, SWR1 exchanges the proximal H2A/H2B copy in relation to its motor subunit [142][214]. Which H2A/H2B copy is exchanged first can be DNA sequence dependent [214]. As a consequence, on +1 nucleosomes the distal copy gets exchanged first in relation to promoter DNA [214]. Further, acetylated H4 stimulates the activity of SWR1 through an increased affinity of SWR1 [237] facilitated by the reader proteins Yaf9 and Bdf1. The NuA4 complex of the INO80/SWR1 family acetylates H4 at the initial phase of DNA repair which targets SWR1 to DNA repair sites [238].

1.6 NuA4/TIP60

The yeast NuA4 complex and its human counterpart Tip60 acetylate the histones H4, H2A and H2A.Z [239] as well as more than 250 non-histone targets [240][241][242]. Thereby, the NuA4/Tip60 complexes play a key role in chromatin organisation during DNA repair as well as transcription regulation [238][243][244][245]. Acetylated H2A and H4 residues are subsequently recognized by SWR1, stimulating its histone exchange activity [237]. NuA4 is composed of three conserved modules [246]: the HAT module that harbors the catalytic subunit Esa1 [247], the core module that forms around the largest subunit Tra1 (transcription activator-binding) as well as the TINTIN module which recognizes phosphorylated RNA Pol II and functions subsequently in transcription elongation [248]. The modular architecture was further resolved by multiple cryoEM studies [144][145][146][249] which illustrated that the main contact points to the nucleosome are facilitated by the HAT module. The Core modules of yeast NuA4 and human Tip60 share four subunits with the SWR1 and SRCAP complexes respectively: Swc4/DMAP1, Arp4/BAF53a, actin and Yaf9/YEATS4. Furthermore, the core subunits Eaf1 (yeast) and p400/Domino (human) function as a scaffold for the remaining NuA4 subunits and contain the conserved HSA helix [250]. While Eaf1/p400 lack other Snf2-type features like the RecA motor lobes, it shares the HSA helix with other chromatin remodelers. In addition, NuA4 comprises a Arp module that resembles the one from SWR1 and are formed by Swc4, Arp4, actin and the Eaf1/Swr1 HSA helices.

1.7 CryoEM as a tool for structural determination

Ever since the resolution revolution in 2014 [251], cryo electron microscopy (cryoEM) became the method of choice for the structure determination of large protein complexes. The advancements in resolution was mainly driven by the implementation of direct electron detectors that are more sensitive and faster as previously installed CCD detectors. Compared to X-ray crystallography, cryoEM enables structural studies of not only large proteins and protein complexes but also to probe for multiple conformational states from the same sample. Advances in software tools allow to push the size and resolution limits to proteins as small as 40 kDa [252] and atomic resolutions around 1.25 Å [253][254]. In addition, subvolumes of cells are imaged by cryo electron tomography (cryoET) visualizing proteins in their cellular and hence native environment (reviewed in [255][256]). The foundations for cryoEM laid the insight that water freezes in a vitreous manner if cooled down at a rate of 10^6 K/s [257][258]. Therefore, proteins in aqueous solutions can be imaged by electrons, also as electrons scatter water differently from biomolecules. The development of the vitrification technique together with hardware and software advances was appreciated with the Nobel prize in 2017 for some of the pioneers in the field: Jacques Dubochet, Joachim Frank and Richard Henderson.
Protein samples are prepared for the cryoEM workflow, by applying them to a metal support called grid (reviewed in [259]). The round grids are commonly 3 mm wide and made of copper with an option for a gold surface. They comprise regularly positioned holes that support ice formation as well as automated acquisition of images on an electron microscope. Samples are vitrified with the aid of automated vitrification devices under controlled humidity and temperature conditions. In this process, the user applies 3-4 µl of protein sample onto a grid which is placed in the humidity chamber. Excess liquid is blotted away by filter paper with a user-defined force and duration. This leaves a thin film of about 100 nm, that is sufficiently thin for electrons to pass through. Next, the grid is rapidly plunge frozen in liquid ethane or a ethane/propane mixture which yields the required cooling rate for vitreous ice formation. Sample preparation is usually an iterative process where one screens the grids on an electron microscope, judges the grids visually or based on small datasets and improves the sample preparation. Protein complexes can fall apart in the thin ice film and optimal sample conditions have to be identified in an iterative approach. Screening involves testing of different grid types as well as grid supports including 2 nm of carbon film, graphene or graphene oxide [260]. However, while they might stabilize protein complexes, these supports can lead to preferred orientation of the particles in relation to the support resulting in missing views and hence insufficient quality of the data [261]. Preferred orientation can also originate from the air-water interface which is hydrophobic and can therefore align particles homogeneously. This bias can be overcome by the use of detergents which cover the interface instead of the protein of interest. In addition, the air-water interface leads to protein degradation which is to be avoided [262].
Once an optimal sample is identified, a dataset of a few thousand images (called micrographs) can be obtained at a high end microscope that comprises an electron source of 300 eV. The electron dose is carefully chosen to at the one hand enhance particle contrast and on the other hand avoid sample damage by electrons. The signal to noise ratio of proteins is low and lower the smaller the particles are. In order to increase contrast a defocus of 0.5-3 µM is applied during data acquisition which can be later accounted for by making use of the contrast transfer function (CTF).
Subsequently, the 2D micrographs are subjected to one of many available cryoEM data processing tools including RELION [263], CryoSPARC [264] or CysTEM [265]. No matter which tool is chosen, the following workflow (reviewed in [266][267]) is applied to reconstruct a 3D protein model that can be used to built atomic protein structures:

1. Pre-processing of the micrographs
2. Particle picking
3. Iterative rounds of 2D and 3D classifications
4. 3D refinements
5. Particle polishing and post-processing (optional)

The overall principle of the data processing is that the bad signal-to-noise ratio of proteins can be overcome by averaging thousands of individual copies present on the micrographs. Under the assumption that the particles are randomly distributed on the micrographs, a 3D volume can be calculated from the obtained 2D views. This is achieved by first correcting the micrographs for motion that occurs during the multiple frame recording and to correct for the applied defocus. Subsequently, particles are picked by automated software tools, some of which are based on neural networks (these days) [268]. The identified particle coordinates are now used to extract the protein molecules with a user defined box size that depends on the pixel size used for acquisition as well as the size of the protein of interest. The obtained 2D images can be used for 2D classification and to calculate an initial model of the best 2D classes. This model can be used as a low-pass filtered template in the subsequent steps. However, 2D classification should only be used to filter out obvious false-positive picking hits as one runs otherwise in the risk of excluding rare views. Next, the 2D views are used for 3D reconstruction and subsequent 3D classification. This step is very powerful, as it allows to identify different conformational states or sub-complexes from the same dataset. For this purpose and in order to sort out the best particles possible, multiple rounds of 3D classifications are performed. Finally, the angles of the best volume are refined which can be also focused on mobile domains to enhance the underlying resolution. In addition, particle re-centering, subtraction or MultiBody refinement [269] can be performed to enhance the resolution of mobile domains. Also, tools can be used to analyse the flexibility of the sample (3DFlex in CryoSPARC [270] or CryoDRGN [271]) in the first place. In a final step per particle ctf refinements and motion correction can be performed in RELION which typically enhances the resolution. This is complemented by a post-processing job in which the obtained map is sharpened with a B-factor (done in RELION).
The obtained EM map is used to build or fit an atomic protein model. Until recently, *de novo* model building or creation of homologous models based on published structures, were the only options to obtain a protein structure. Nowadays, structure prediction tools like AlphaFold [272] and AlphaFold Multimer [273] are used as a starting point for model building or can be even automatically performed (ModelAngelo ([274]). The models are then further improved in dedicated model building softwares like COOT [275] and ChimeraX Isolde [276] and refined (phenix [277], refmac servelcat [278]), followed by quality control (MOLPROBITY [279]).

1.8 Aims of this PhD thesis

The mega-Dalton sized chromatin remodelers of the INO80-family share a modular architecture, yet shape the chromatin landscape by distinct remodeling reactions. Even in light of recent advancements through cryoEM, the molecular mechanisms of these enigmatic enzyme complexes are still poorly understood.
The INO80 chromatin remodeler structures of the Core and Arp modules and their interplay, are well-characterized [154][77][141][150]. These studies laid the basis to understand how INO80 acts on its nucleosomal substrates. However, emerging evidence suggests, that non-canonical nucleosomes, including hexasomes, play a key role in chromatin regulation. INO80 was recently found to slide hexasomes, yet it remains an open question of how INO80 acts on hexasomes on a molecular level [67]. I hypothesized that INO80 must adopt a completely different mode of remodeling on hexasomes as they lack one copy of H2A/H2B which was previously suggested to be essential for nucleosome remodeling. To resolve this conundrum, I sought out to:

- Implement the recombinant protein production, enzymatic activity assays, and cryoEM structural studies of the evolutionarily conserved *C. t.* INO80.
- Determine a high-resolution structure of INO80 in complex with a hexasome.
- Biochemically probe the reaction mechanism of INO80 on hexasomes by using site-directed mutagenesis and enzymatic activity assays.

The SWR1 histone exchanging complex is less well studied in respect to its molecular structure and mechanism. The published SWR1 mono-nucleosome structure lacks many subunits including the Arp module and only resolved a limited number of SWR1-histone core interactions [142]. Furthermore, the role of the Arp module remains elusive as it was not studied in context of the full-length SWR1 complex and only minimal structural knowledge exists. In summary, it is necessary to solve cryoEM structures of the full-length complex bound to its nucleosomal substrate. This will not only enable an understanding for the substrate specificity and mechanism of histone exchange by SWR1, but also shed light on how INO80 and SWR1 facilitate such different reaction outcomes on their nucleosomal substrates. I decided to use the conserved SWR1 complex from the thermostable *Chaetomium thermophilum* (*C. t.*) fungi, as proteins from this organism proved successful in cryoEM for INO80. The SWR1 project can be divided into the following parts:

- Identify and recombinantly express the so far uncharacterized SWR1 complex from the thermophilic fungus *C. t.*
- Structurally and biochemically characterize *C. t.* SWR1 by using an integrative approach of cryoEM, AI-powered structure predictions, as well as crosslinking mass spectrometry. This part of the study will initially focus on the nuclear actin-containing Arp module as the basis for further characterizing the structure and mechanism of the entire complex.
- Establish recombinant expression of *S. c.* SWR1 to measure its histone exchange activity in the context of native gene promoter sequences to understand the functional impact of its mechanism.

2 Results

2.1 Implementation of biochemical and structural experiments for INO80

In this section, I will present how I implemented the recombinant protein production of *C. t.* INO80 as well as the cryoEM workflow for INO80 bound to a mono-nucleosome. This lay the basis for the following chapter that focuses on how INO80 acts on hexasomes on a molecular level.

2.1.1 Purification strategy

Chaetomium thermophilum INO80 was characterized in the Hopfner lab in Munich by my supervisor Sebastian Eustermann. He had cloned a minimal INO80 complex that misses some of the species-specific subunits (Ies1, Iec3, HMG, ZnF and FHA) and does not include the N-terminus of the Ino80 subunits (residues 1-721). This construct is sufficient to resolve the structure of the INO80 Core and Arp modules as published previously [150]. I used this construct for recombinant protein production from insect cells and the establishment of biochemical and structural tools for INO80 in our newly founded lab.
In more detail, the gene sequences of INO80 are partitioned among two plasmids. One containing the core subunits that contact the nucleosome: Ino80, Rvb1, Rvb2, Ies2, Arp5 and Ies6, while the other plasmid harbors the Arp module and the remaining of the species-specific subunits namely actin, Arp4, Arp8, Ies4 and Taf14 (Figure 2.1a). For the generation of baculoviruses, I prepared Bacmid DNA for each of these two plasmids by blue-white selection and viruses were generated using Xtreme reagent. After one round of virus amplification, the generated baculoviruses were used to co-infect Hi5 insect cells. Protein expression was enabled for about 65 hours. I purified *C. t.* INO80 within one day by an anti-FLAG purification step as the Ino80 subunit was double FLAG tagged, followed by an anion exchange step (see workflow in Figure 2.1b). This twofold purification strategy yielded INO80 complexes in high purity (gel of purification in Figure 2.1c). Finally, the recombinantly produced INO80 was used for biochemical and structural studies.

2.1.2 Nucleosome design and nucleosome preparations

Recombinant nucleosomes were required in this thesis for both biochemical and structural experiments. For this purpose human histones were used from The Histone Source. Histone octamer reconstitution, DNA preparation and nucleosome reconsitution were executed by the lab manager Olga Kolesnikova.
As DNA, the asymmetric Widom 601 sequence [280] of 147 bp with a long linker DNA of 80 bp was used creating 0N80 nucleosomes. The nomenclature describes that there are 0 bp of linker DNA on one side of the nucleosome core, while there are 80 bp on the other side of the 601 sequence. In more detail, the 601 sequence harbours a TA-rich and TA-poor side separated by the dyad that creates a pseudo symmetry of the nucleosome (illustrated in Figure 2.2a). The 80 bp linker is placed next to the TA-poor side. The TA-poor side contains less T and A nucleotides resulting in lower affinity of

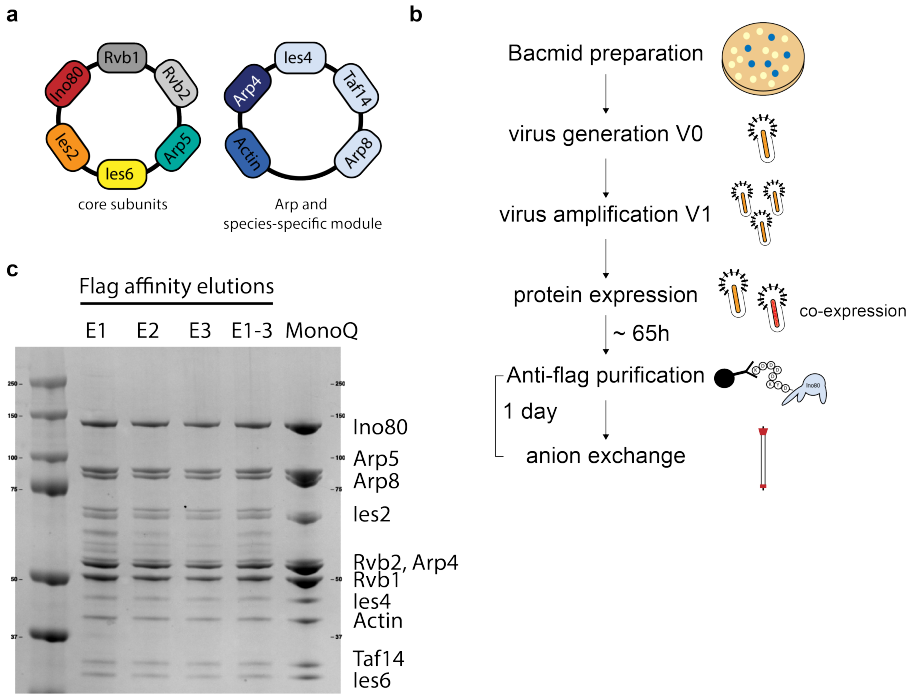


Figure 2.1: **Purification strategy of *C. t.* INO80. a** Construct overview of the two INO80 plasmids. **b** Purification strategy. **c** SDS-PAGE of the two-step INO80 purification. Elutions (E) from FLAG affinity step are shown alongside the main peak fraction of the MonoQ anion exchange chromatography.

the DNA to the histone core. As a consequence, hexasomes can be reconstituted when the presence of H2A/H2B dimers is limited. The H2A/H2B close to the TA-poor side will be missing, creating a longer linker DNA.
The 601 sequence is very potent to wrap around the histone octamer core and is therefore used for an uniform positioning of the DNA [280]. The long linker of 80 bp enables sufficient INO80 binding which requires at least 60 bp of extra-nucleosomal DNA for successful binding [194].
PCR-amplified DNA is then mixed with purified histone octamers at high salt concentrations. Dialysis of the two components into lower salt results in nucleosome reconstitution (example gel in Figure 2.2b).
I used the purified nucleosomes to test the biological activity of recombinant *C. t.* INO80. For this purpose, I chose an INO80 sliding assay which probes not only the ATPase activity of the Ino80 motor domain but also outputs the biological relevant activity of INO80 which is the sliding of nucleosomes. I utilized the 0N80 nucleosomes in which the nucleosome core particle is end-positioned. Upon addition of ATP, INO80 uses ATP hydrolysis to slide the nucleosome into a centered position on the DNA. The slid product has about 40 bp of DNA on each side of the nucleosomal core (see scheme in Figure 2.2c). The end-positioned and central-positioned nucleosomes run differently on a Native PAGE. Therefore, one can follow the reaction on a Native PAGE over time (Figure 2.2d). Already after 30 s one begins to see the sliding activity in form of a a band shift on the Native PAGE. Half of the nucleosomes are slid after 10 min. The assay clearly shows that the recombinantly expressed *C. t.* INO80 is biologically active and can be used for subsequent structural studies.

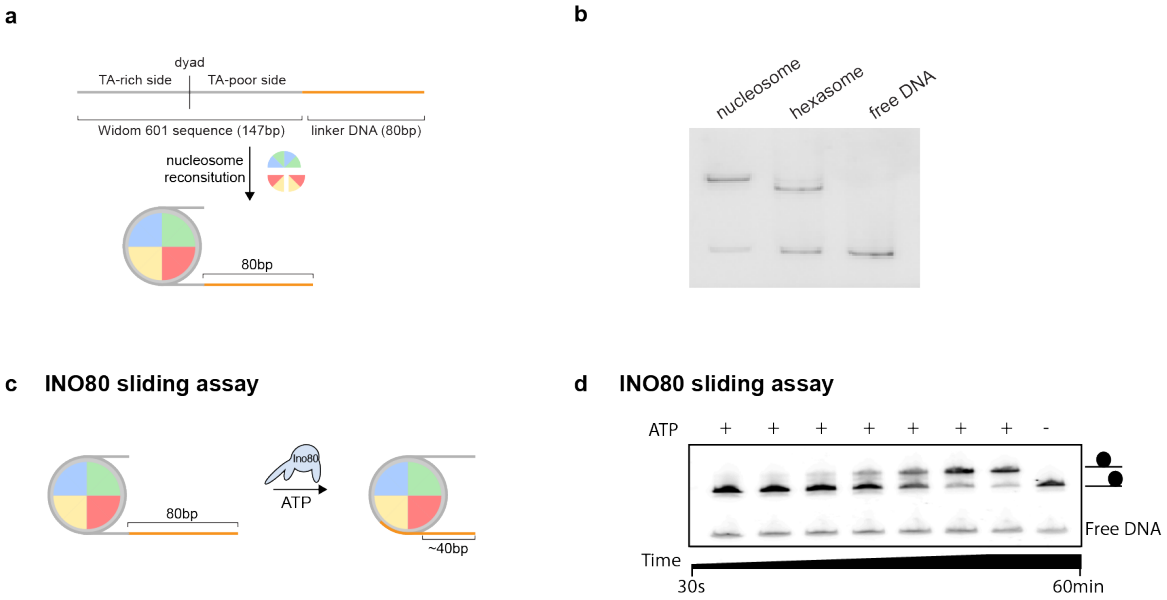


Figure 2.2: **Nucleosome preparation and INO80 sliding assay. a** Nucleosome design of 0N80 nucleosomes. The 601 sequence has a TA-rich and TA-poor side. I used nucleosomes with 80 bp of extra-nucleosomal DNA. **b** Agarose gel of an exemplary nucleosome reconstitution. Typically, three peaks were observed in the anion exchange chromatography step resembling fractions of nucleosomes, hexasomes and free DNA as shown on the gel. Free DNA is part of the nucleosome and hexasome samples to some degree. Gel kindly provided by Olga Kolesnikova. **c** Schematic illustration on how the INO80 sliding assay works. End-positioned 0N80 nucleosomes are slid into a central position upon addition of ATP. **d** Native PAGE showing the result of the sliding assay. Lowest band represents free DNA that is always part of nucleosome preparations. Middle bands are end-positioned nucleosomes (educt) while the upper most bands are INO80-slid nucleosomes (product). Also indicated by little cartoons on the right.

2.1.3 CryoEM on the INO80-nucleosome complex

I used 0N80 nucleosomes that were produced by Olga Kolesnikova and recombinant *C. t.* INO80 to prepare cryoEM grids and thereby establish the data collection and data processing pipeline of our lab. Therefore, I collected 11000 micrographs on a Titan Krios microscope equipped with a K2 direct electron detector.

RELION 3.0 [263] was used for data processing. First, automated picking was performed on the motion corrected and ctf corrected micrographs (Figure 2.3a). Picked particles were subjected to 2D classification which is illustrated in Figure 2.3b. Clearly visible is the INO80 Core module bound to a nucleosome. The latter are mostly captured from the side as one can see the two gyres of DNA. Besides the INO80 core, one can see a smeared out domain which indicates its flexibility in relation to the core module. This represents the post-HSA that connects to the Arp module of INO80. A selection of good 2D classes was used to generate an initial model of INO80-nucleosome particles. The initial model was used for multiple rounds of 3D classifications on the full dataset to mainly separate apo INO80 from nucleosome bounds ones. 3D refinement followed by post-processing and Bayesian polishing resulted into a map with a final resolution of 3.3 Å (Figure 2.3c). The final EM density map (Figure 2.3d) was used to fit the published INO80-nucleosome structure into the density (Figure 2.3e; PDB code: 6FML). Overall, the Core complex of INO80 has a Pac-Man-like shape. Its two arms – the Ino80 motor domain and the Arp5/Ies6 grip – enclose the nucleosome.

The grips are hold in place by the hexameric Rvb1/Rvb2 AAA$^+$ ATPase that functions as a scaffold. The Ino80 lobes are tethered to the Rvb1/Rvb2 barrel through a long insertion in-between the N and C-terminal lobe. The two main contacts to the nucleosome are facilitated by the Ino80 motor lobes and the Arp5-core/Ies6 grip domain. The motor lobes bind the nucleosome at SHL -6 (superhelical location), while the Arp5 subunit has a secondary contact point enabled through its cross-shaped grappler domain which folds both against the DNA and the histone core. The latter is enabled by the Arp5 foot helix that contacts the acidic patch of H2A/H2B. I additionally see some density for Ies2 close to the nucleosome and on the Rvb1/Rvb2 interface.

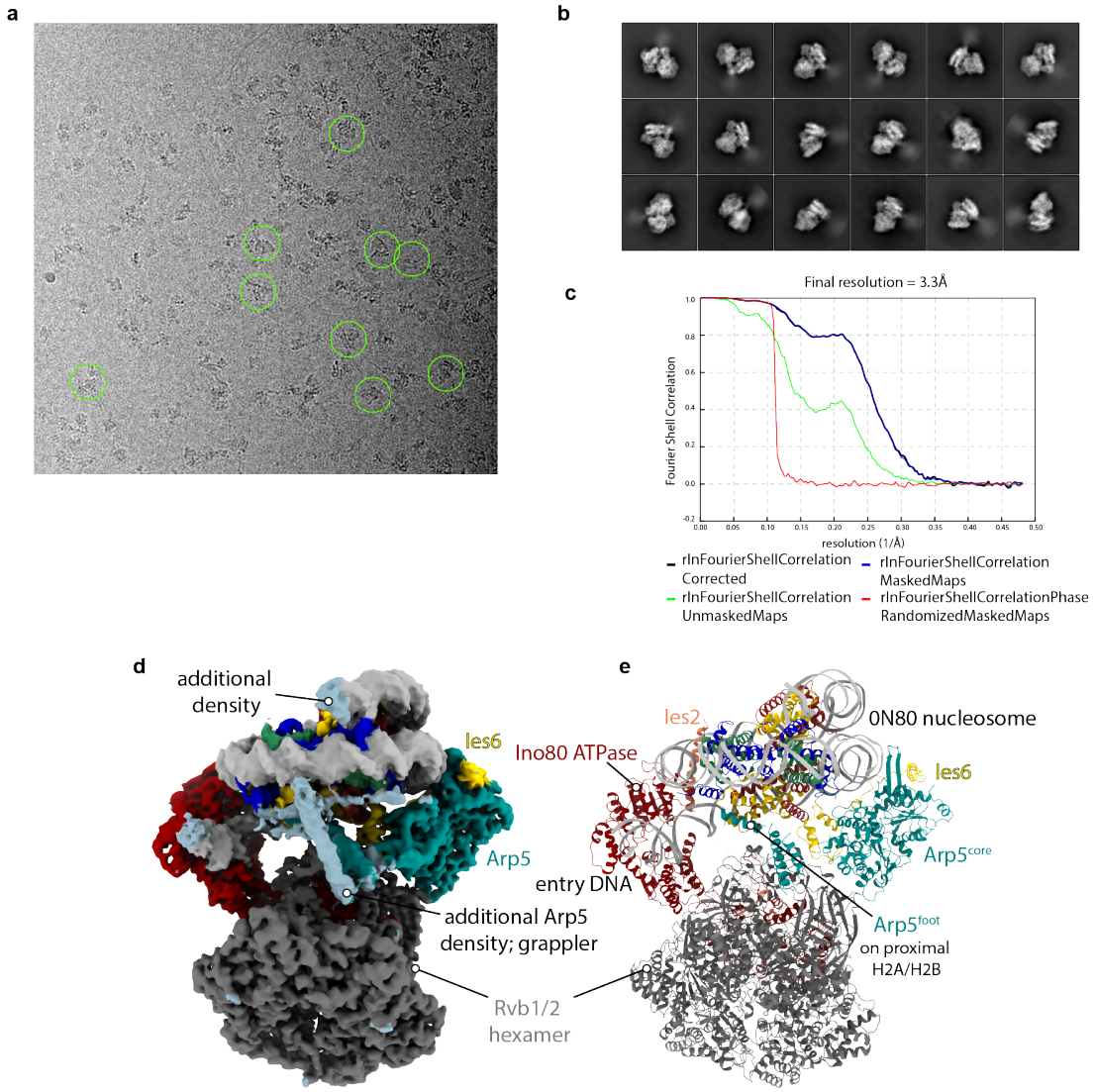


Figure 2.3: **Structural studies on *C. t.* INO80 in complex with a nucleosome. a** Example micrograph showing picked particles in green circle. **b** 2D classes of INO80-nucleosome complex with a box size of 360 Å. **c** FSC curve illustrating the final resolution of 3.3 Å. **d** EM density map of INO80-nucleosome complex. Rvb1/2 hetero-hexamer is colored in dark gray, Arp5 in cyan, Ies6 in yellow, Ino80 motor in dark red, Ies2 in orange, DNA in light gray, H2A in yellow, H2B in red, H3 in blue and H4 in green **e** Rigid body fitted structure of the INO80-nucleosome based on PDB 6FML. Colors likewise to **d**.

In comparison to the published INO80-nucleosome structure in 2018 [77], I resolved some additional density (colored in light blue) which corresponds to some additional DNA bases and the Arp5 grappler. As an INO80-nucleosome structure including the Arp5 grappler was published in 2022 by the Hopfner lab [150], I am not discussing the result further in the results part here – more details can be however found in the introduction to this PhD thesis and in Section 2.2.3. Of note is however, that the Rvb1 and Rvb2 subunits are occupied by an ADP molecule respectively, despite the absence of nucleotides from the purification strategy as well as the presence of $CaCl_2$. Seemingly, the ADP molecules were pulled out endogenously from the expression system. In summary, I showed that the EM pipeline and our transmission electron microscope setup are sufficient to resolve the INO80-nucleosome complex at near atomic resolution. Finally, the establishing of the biochemical and structural tools for INO80 at EMBL Heidelberg, laid the foundation for the next chapter: The study of INO80 bound to a hexasome substrate.

2.2 Structural studies of INO80 bound to a hexasome

A recent publication of the Narlikar lab [67] showed that *Saccharomyces cerevisiae* INO80 can not only slide hexasomal substrates but also does so in a preferred and faster manner compared to nucleosomes. Min Zhang – a postdoc from the lab – and myself were inspired by this finding and asked ourself how INO80 acts on hexasomes on a molecular level. The mechanism remained especially elusive, as hexasomes comprise only one acidic patch on the remainng H2A/H2B dimer while INO80 contacts both acidic patches on nucleosomes [77]. In addition, nucleosome sliding is dependent on the acidic patch. In order to understand the dual substrate specificity of INO80, we teamed up to perform cryoEM on INO80 bound to a hexasome. We divided the work as follows: Min Zhang prepared the hexasomal and nucleosomal substrates required for this study. Protein production was performed by myself, Min Zhang and Franziska Kunert – a PhD student from the Hopfner lab in Munich (Franziska Kunert: mutants and *S. c.* INO80). For the cryoEM part, Min Zhang prepared the cryoEM grids, collected the cryoEM data and conducted the data processing of the obtained micrographs. Finally, I built the protein structures into the obtained cryoEM maps. For validation of the structure, I performed INO80 sliding assays. In the revision process of our paper, Min Zhang did some extra sliding assays that were required from the reviewers. Our work was published in Science 2023 [281]. While I phrased the results in my own words for the presented thesis, some figures in Section 2.2 resemble in parts the figures presented in our paper [281].

2.2.1 The INO80 and hexasome samples

For the structural studies on the INO80-hexasome complex, we used the N-terminal truncated INO80 construct from *Chaetomium thermophilum* as described in Section 2.1.1. The very same construct was used for structural studies before [77] and I showed by activity assays (Figure 2.7d) that the truncated construct is active on hexasomes and is similarly active in comparison to full-length *Saccharomyces cerevisiae* INO80 on nucleosomes. The truncated INO80 construct is hereafter referred to as INO80 wild type (wt) to distinguish it from later used mutant INO80. Its subunit composition is illustrated in Figure 2.4d. For the biochemical assays, Franziska Kunert prepared three different Arp5 mutants of INO80 (gels of purification in Figure 2.4a): The Arp5-heel mutant (R501E/K502E), Arp5-heel-foot mutant (R501E/K502E/Q507A/R509A/M510S/K511A/I513S) and the Arp5-DBD (DNA binding domain) (K88A/R90A/R92A/K93A/K96A/R112A/R116A). In addition to *C. t.* INO80, I used INO80 from *S. c.* (purified sample in Figure 2.4b) to compare the sliding activity of *S. c.* and *C. t.* INO80 on hexasomes and nucleosomes. This was especially important as endogenous *S. c.* was used in the

study that showed that hexasomes are the preferred substrate of INO80 [67]. In comparison to our *C. t.* INO80 construct *S. c.* INO80 contains the Nhp10 module with its species-specific subunits Ies1, Iec3, Ies5, Nhp10 and Taf14 (Figure 2.4d).
For the preparation of 0H80 hexasomes, we made use of the asymmetry of the 601 Widom DNA sequence as described in Section 2.1.2. Using limiting conditions of H2A/H2B, the TA-poor site does not reconstitute a copy of H2A/H2B as it has an intrinsically lower affinity for the dimer compared to the TA-rich side. As the linker of 80 bp is close to the TA-poor side of the 601 sequence, a total linker length of about 125bp is created (see Figure 2.4c).

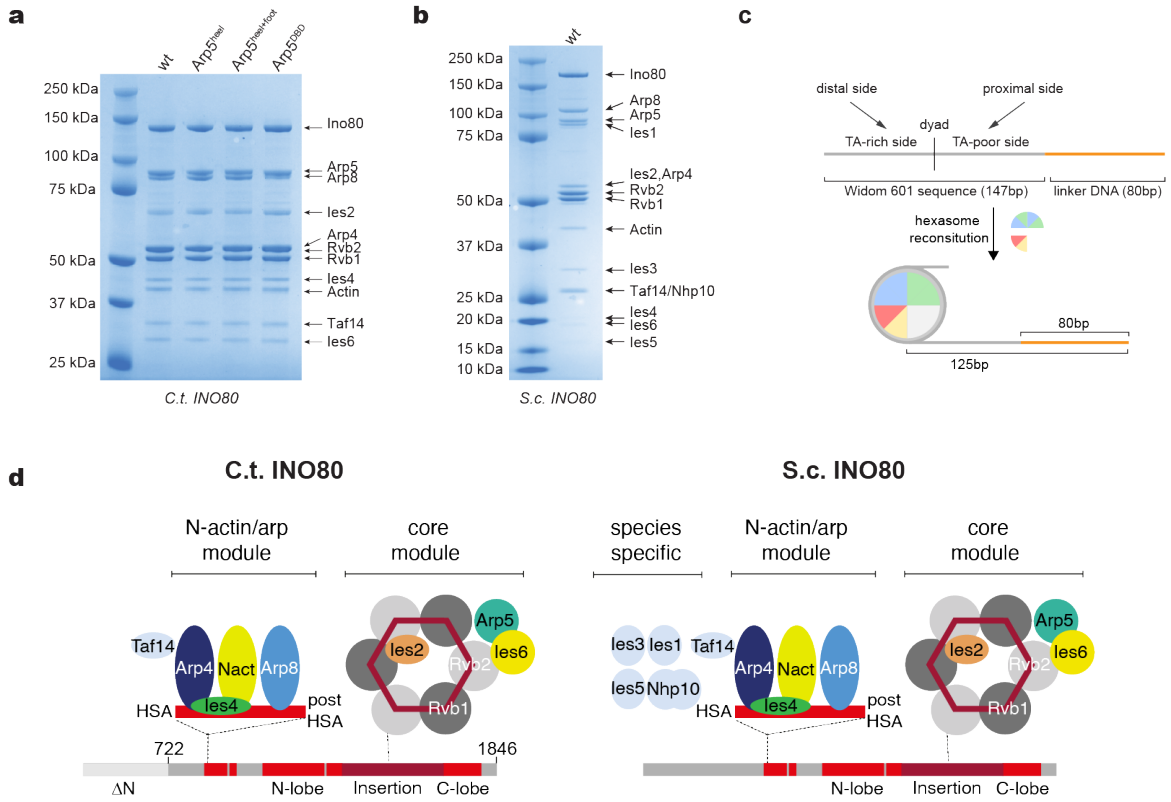


Figure 2.4: **INO80 and hexasome samples. a** SDS-PAGE of wt *C. t.* INO80 and Arp5 mutants. Identities of subunits are labeled respectively. **b** SDS-PAGE of purified *S. c.* INO80. **c** Design of hexasomal DNA and architecture of reconstituted hexasomes. Linker DNA is composed of 80 bp of DNA plus 45 bp that are unwrapped from the histone core due to the missing H2A/H2B copy. In combination, the extra-nucleosomal DNA has a length of about 125 bp. **d** Complex architecture of *C. t.* INO80 (*left*) and *S. c.* INO80 (*right*). The *C. t.* INO80 construct lacks the N-terminal region of the Ino80 subunit as well as some subunits of the species-specific module.

2.2.2 CryoEM and model building of the INO80-hexasome complex

As mentioned before, Min Zhang prepared grids of the INO80-hexasome sample and collected 15384 micrographs at a magnification of 105000 on a Titan Krios transmission electron microscope equipped with a K3 camera (example micrograph in Figure 2.5a). In 2D classifications, both classes for the INO80 core bound to the hexasome (Figure 2.5b) as well as the Arp module bound to DNA were visible (Figure 2.5c). The Arp module of INO80 is flexibly connected to the Core module by the post-HSA (data not shown).

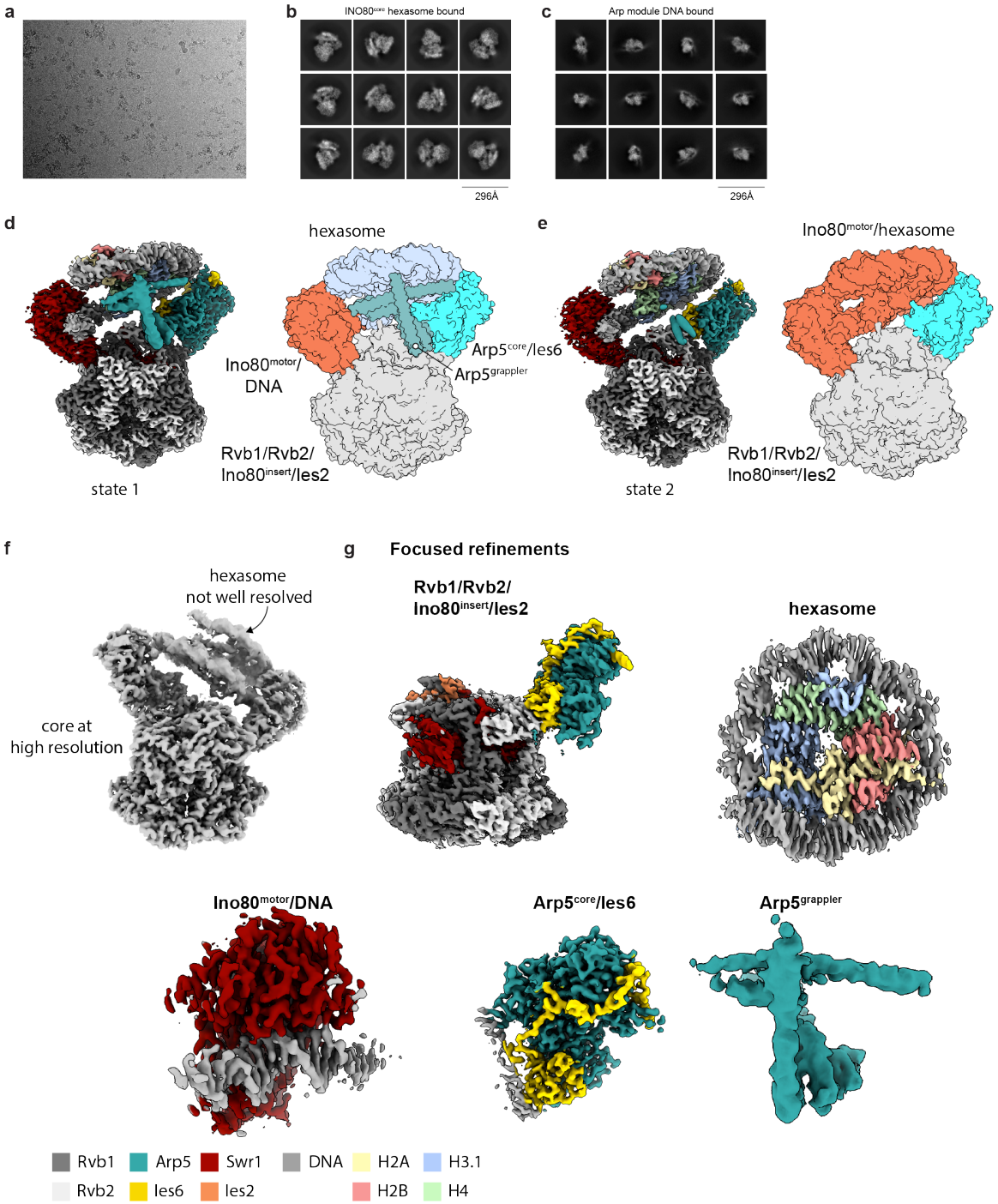


Figure 2.5: **Data processing of INO80-hexasome and focused refinements of state 1. a** Example micrograph **b** Selection of 2D classes of the INO80 core bound to hexasome. **c** 2D classes of the INO80 Arp module bound to extra-nucleosomal DNA. **d** cryoEM map of state 1 (*left*) and overview of focused refinements (*right*). **e** cryoEM map of state 2 (*left*) and overview of focused refinements (*right*). **f** Example map (EMDB 17019) of the INO80-hexasome complex highlighting the flexibility of the hexasome in relation to the Ino80 core underneath. Consequently, the hexasome density shows less features and is of lower resolution. **g** Focused refined maps of state 1. Resolved parts are labeled by a title to each EM density. Color code of the presented subunits on the bottom.

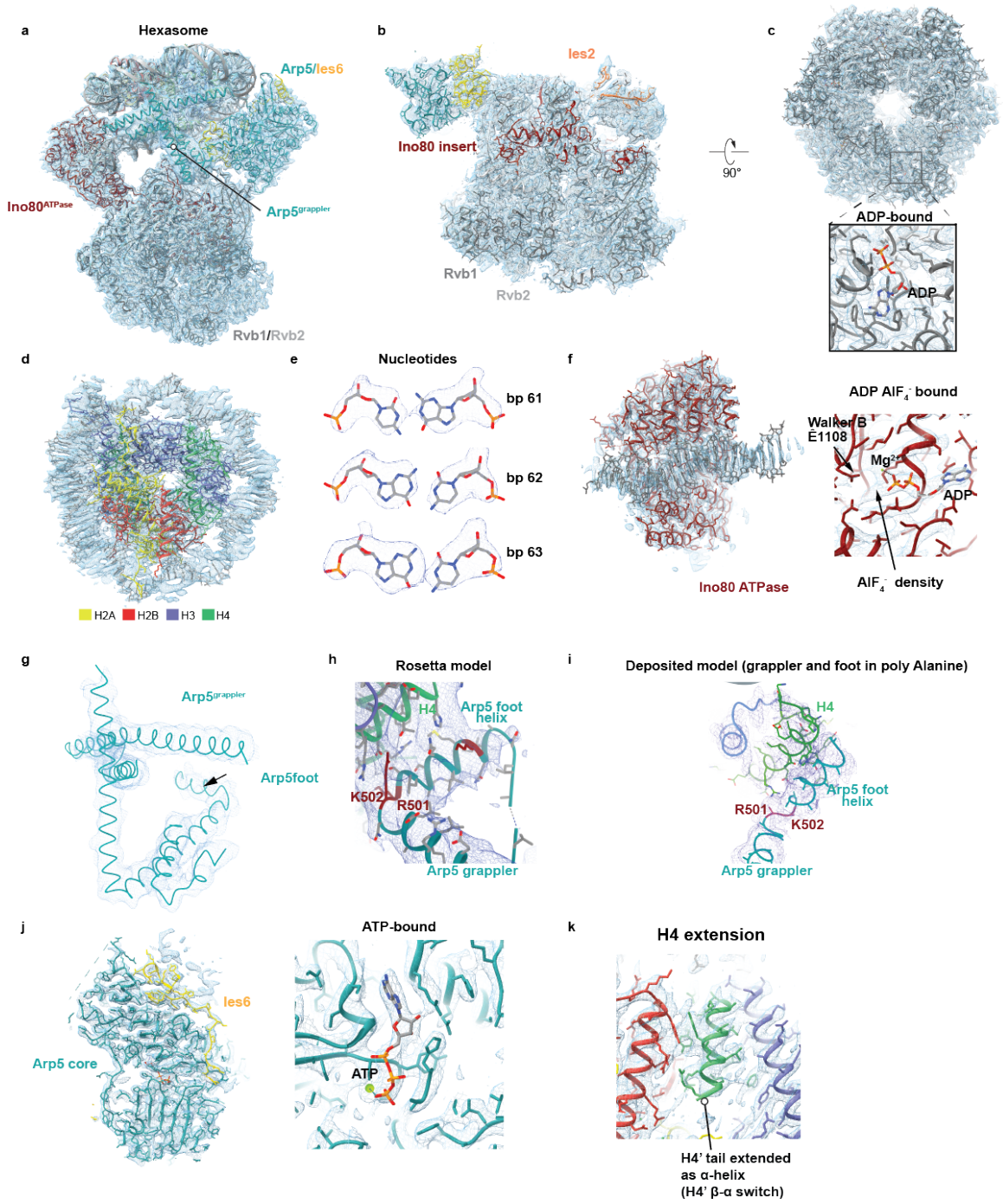


Figure 2.6: **Models of state 1 fitted into cryoEM maps.** **a** Overall model of state 1. Respective subunits are labeled. **b** Rvb1/Rvb2 core with Ino80 insert, and Ies2. **c** Bottom view of Rvb1/Rvb2 hetero-hexamer. In close up, one of the in total 6 bound ADP molecules **d** Top view of the hexasome. **e** View of nucleotides 61 – 63 in their respective EM density showing that I could assign the sequence register of the hexasomal DNA **f** Ino80 ATPase bound to hexasomal DNA. On the right, active site of Ino80 subunit with a bound ADP molecule and a magnesium ion. Extra density for AlF_4^- as well as Walker B residue is labeled. **g** Structure of Arp5 grappler has been built only with alanines due to the limited resolution. Arp5 foot is labeled separately. **h** Rosetta modelled Arp5 helix. Extra labeled are subunits K501 and R501 as well as K511 as they were used later as mutants. **i** Close up of Arp5 heel (R501 and K502, in red) and foot as they fold alongside the H4 $\alpha 2$ helix (in green). **j** Built models for Arp5 core and Ies6 with close up view of Arp5-bound ATP molecule. **k** Close up of extended H4' helix.

Therefore, particles for the Core and Arp modules were picked and processed separately. Yet the high degree of flexibility only allowed to reconstitute the Arp module at 3.7 Å resolution. For that reason the structure was only rigid body fitted based on [150]. Continuous rounds of 3D classifications and refinements of the INO80 Core module resulted into two maps termed state 1 and state 2 (Figure 2.5d and e respectively). They mainly differ in their binding mode of the Arp5 subunit as shown in more detail in Section 2.2.5. The maps of state 1 and state 2 were resolved at a resolution of 3.15 Å and 3.08 Å respectively. However, the local resolution differs among the various complex domains. Especially the interface of INO80 towards the hexasome and the hexasome itself are of worse resolution as the hexasome undergoes a briefing motion. This prevented to obtain high resolution reconstruction of the INO80 and hexasome domains at the same time (illustrated in Figure 2.5f). Due to the motion of the hexasome, the Ino80 motor lobes, the Arp5/Ies6 dimer as well as the Arp5 grappler show heterogeneity resulting in maps that are not sufficient to build atomic models. To overcome this hurdle, Min Zhang performed focused refinements by creating a mask around different regions: 1) Rvb1-Rvb2-Ino80 insert-Ies2. Arp5 and Arp6 were also present but of worse resolution and were therefore built into another focused refined cryoEM density. 2) Hexasome with histone core and hexasomal DNA 3) Ino80 motor bound to hexasomal DNA 4) Arp5 core-Ies6 and 5) the Arp5 grappler domain (illustrated in Figure 2.5g). For the refinement of the hexasome it was necessary to perform in addition particle subtraction. In this step, only the density for the hexasome was kept to improve the angular assignments independently of the remaining INO80 core. With this method, Min Zhang obtained an hexasome structure at 3.18 Å resolution. Overall, the density of state 2 is of worse quality compared to state 1 and the density of the Arp5 grappler is hardly visible. Therefore, I first performed model building of state 1 and used them as a starting point for state 2.

The flexibility of the INO80-hexasome complex posed further challenges to the subsequent model building procedure. First, I built the atomic models in the focused maps based on previously published structures (7OHC for nucleosome, 6FML for Arp5 core and Ies6, 3I62 for ADP•AlF_4^- and 8AVG for remaining parts) followed by model building in COOT [275]. Next, models were improved using iterative rounds of the molecular dynamics programme ISOLDE [276] (done by Thomas Hoffmann) and real-space refinements in Phenix [277] (done by myself). Lastly, composite models were generated by rigid body fitting the individual models into their respective composite map (state 1 or 2), connecting the chains were necessary and adapting the side chains at the map interfaces. The interface of the hexasome and INO80 Core complex was carefully inspected and residues were only interpreted if they were clearly visible in the overall maps. Furthermore, as the interface of the hexasome and the Arp5-foot was of limited resolution Frank DiMaio modeled the Arp5-foot using Rosetta. With his model, I was able to build the residues of the Arp5-foot into the respective EM density, while we deposited the PDB structure with only an alanine backbone. For validation, tables with statistics on the obtained maps and built protein structures can be found in the methods section in Tables 4.1 and 4.2.

The quality of the focused refined maps and the built models can be seen in Figure 2.6. Illustrated are only the maps and model for state 1 as it is the one mainly visualized throughout this thesis due to its better quality around the hexasome. As the resolution of the hexasome is at 3.18 Å, I was able to assign the sequence register of the nucleotides (see Figure 2.6e). In addition, the quality of the maps was sufficient to see various nucleotides bound to INO80 subunits. The active site of Ino80 harbors a ADP•AlF_4^- molecule that was added prior to sample preparation (Figure 2.6f). It resembles a transition state of ATP which leads to a displacement of the N and C lobes leading to a more enclosed binding of the motor to the hexasome. Therefore, the complex is more stable and a higher amount of hexasome-bound complexes was observed compared to preparations without ADP•AlF_4^-. Even though INO80 was purified in presence of $CaCl_2$ which stabilizes ATP, some subunits contain ADP molecules. In more detail, Arp5 harbors an ATP molecule (Figure 2.6j) while each of the Rvb1

and Rvb2 subunits contain ADP (Figure 2.6c). The resolution of the obtained cryoEM maps was sufficient to built the ADP and ATP molecules into the density.

2.2.3 *C. t.* INO80 recognizes hexasome by multiple binding sites

Overall, the structures of INO80 bound to a hexasome (Figure 2.7a) and a nucleosome (presented earlier in Figure 2.3d, e) resemble each other. The two main contact points remain to be the Ino80 motor and Arp5/Ies6 domain. However, I did not observe any density for the Ies2 throttle helix which contacts on nucleosomes both DNA and through its extension the acidic patch of the distal H2A-H2B copy as shown in [77]. In contrast, only the Ies2-anchor sitting on top of Rvb1 was captured. This indicates a higher flexibility of the underlying domain and possibly the hexasomal particle itself.

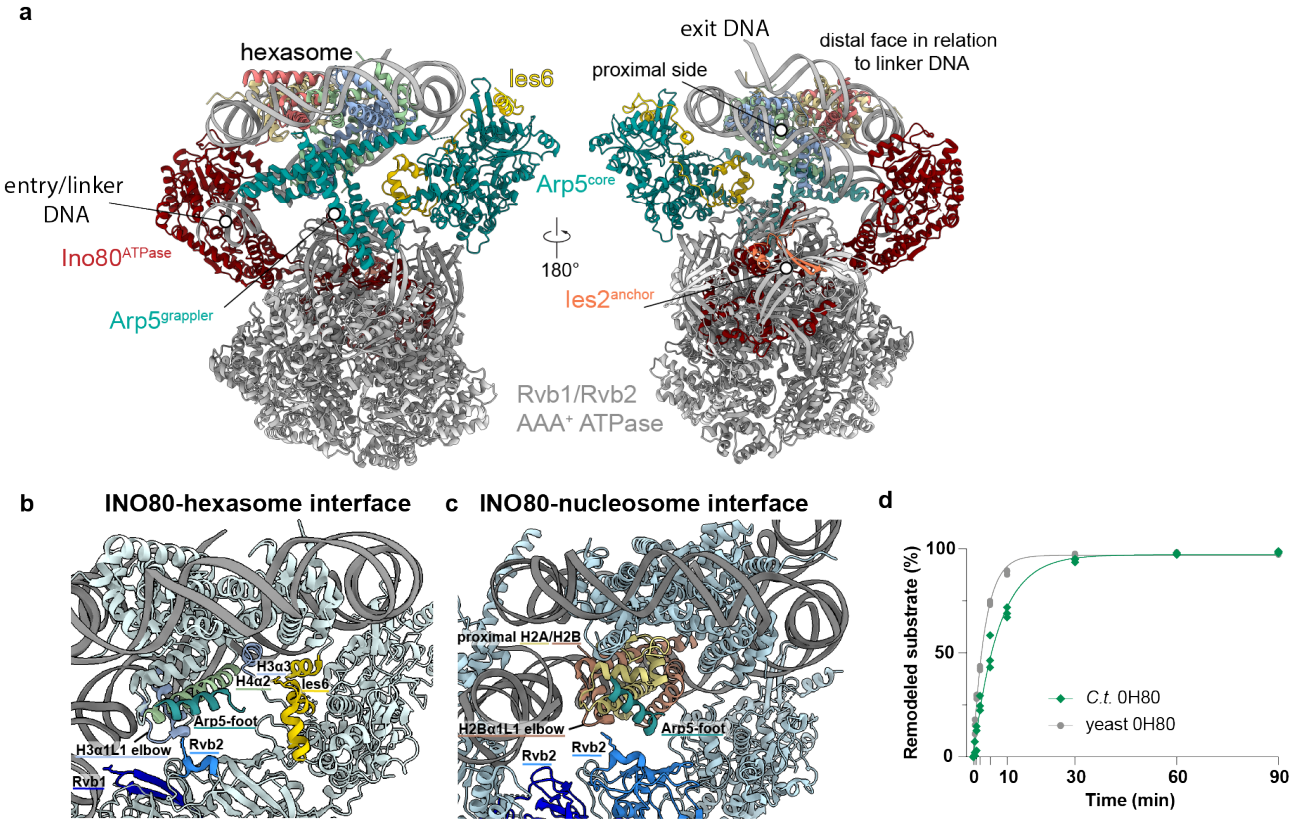


Figure 2.7: **CryoEM Structure of INO80 bound to its hexasomal substrate.** **a** Structure of *C. t.* INO80 bound to a hexasome (PDB 8OO7, from this study) in two views. The major DNA contacts are made by the Ino80 ATPase and the Arp5/Ies6 grip. Secondary contacts are illustrated in **b** (Same view as in panel a, *left*). Rvb1 residue R166 binds to the entry DNA, Rvb2 residue R212 contacts the H3α1L1 elbow and Ies6 contacts both the DNA (via a helix from T43 to Q52) as well as the H3α3 helix that is bound by Ies6 residue Q202. In addition, the Arp5-foot helix binds alongside the H4α2 helix. **c** Similar view as in panel **b** but of the INO80-nucleosome complex. The same Rvb1/2 interface and the Arp5 foot contact here the acidic patch of the proximal H2A/H2B dimer as well as the H2Bα1L1 elbow. **d** Sliding assay of *C. t.* INO80 and yeast INO80 on 0H80 hexasomes. *C. t.* INO80 is on the same substrate almost as active as yeast INO80.

Also some other secondary binding sites of INO80 of the hexasome are different compared to the INO80-nucleosome complex. The discrepancy arises from the missing H2A/H2B copy in the hexasome which consequently exposes the underneath H3/H4 interface. In more detail, the proximal

H2A/H2B copy is absent in relation to DNA linker. The proximal H2A/H2B copy lays in close proximity to the Rvb1/Rvb2 hetero-hexamer in context of nucleosome binding while the distal copy is solvent exposed. The uncovered H3/H4 tetramer offers novel binding sites for the INO80 complex which stabilizes the exposed H3/H4 interface in return (Figure 2.7b). In this context the INO80 subunits Rvb2, Arp5 and Ies6 bind to the nearby copies of H3 and H4. In depth, Rvb2 residue R212 contacts the H3α1L1 elbow. The same Rvb2 residue binds to the H2Aα1L1 elbow in context nucleosomes (as illustrated in Figure 2.7c). Notably, these histone elbows are known hotspots for chromatin binding factors [14] and are also used by INO80 to recognize both nucleosomes and as shown here – hexasomes. In addition, the Ies6 residue Q202 points in the hexasome structure towards the H3α3 helix and the Arp5-foot helix folds against the H4α2 helix. Of note is, that the same Arp5-foot helix is in contact with the acidic patch of the proximal H2A/H2B, when bound to a nucleosome. However, this H2A/H2B copy is not available for binding on the hexasome – instead the Arp5-foot binds to H4. Besides the histone contacts, I observed some additional DNA binding sites of INO80 to the hexasome: the arginine residue 166 of Rvb1 points towards the backbone of the entry DNA and the Ies6 helix from T43 to Q52 folds alongside DNA.

The remaining INO80 subunits assembly onto the N-terminus of the Ino80 subunit. The protrusion helix and the post-HSA helix connect the Ino80 Core to the Arp module which contains the subunits nuclear actin, Arp4, Arp8 and Ies4. The post-HSA helix is however not visible in the structure. Hence, the connection between the INO80 Core and Arp module is more flexible in the INO80-hexasome complex compared to the nucleosome-bound state. The Arp module binds the extra-nucleosomal DNA 20 bp upstream of the DNA entry site of the hexasome [281]. The N-terminal region of INO80 – also denoted as species-specific module – is absent from this structure due to the high degree of flexibility in-between the Arp module and the N-terminus. The species-specific module contains only the Taf14 subunit in the used construct. Accordingly, the Nhp10 module subunits are lacking in comparison to yeast INO80. Nonetheless, I could show that the *C. t.* INO80 construct has a similarly high sliding activity compared to yeast INO80 which does contain the Nhp10 module (Figure 2.7d). In conclusion, the Nhp10 module is not crucial for successful hexasome or nucleosome sliding by INO80.
I showed that the overall INO80 architecture remains the same when binding to a hexasomal substrate instead of a nucleosome. However, unique hexasome features alter the INO80-histone core interface. Therefore, I will present the adopted binding mode of INO80 on hexasomes in Figure 2.11 and Section 2.2.8.

2.2.4 Motor domain comparison between different nucleotide states

The INO80-hexasome structure is in an active state: The nucleotide binding pocket in-between the N and C-lobes harbors an ADP•AlF_4^- molecule which was added shortly before sample vitrification. The magnesium and AlF_4^- ions mimic the transition state of ATP. This can be also seen by the fact that the magnesium ion is coordinated by Ino80 residue T1003 while the AlF_4^- is bound by E1108 (Figure 2.8c). The latter residue is described as Walker B residue [282] – when mutated to a non-charged residue, Ino80 fails to catalyse ATP hydrolysis thereby loosing its sliding activity on nucleosomal substrates.
Together with the INO80-hexasome structure, there are now Ino80 structures in three different nucleotide states available: nucleotide-free (PDB 8AV6), bound to ADP•AlF_4^- (this study) and in the ADP•BeF_3^- (PDB 8ATF) state. While ADP•BeF_3^- mimics ATP, the ADP•AlF_4^- state resembles the ATP hydrolysis transition state.

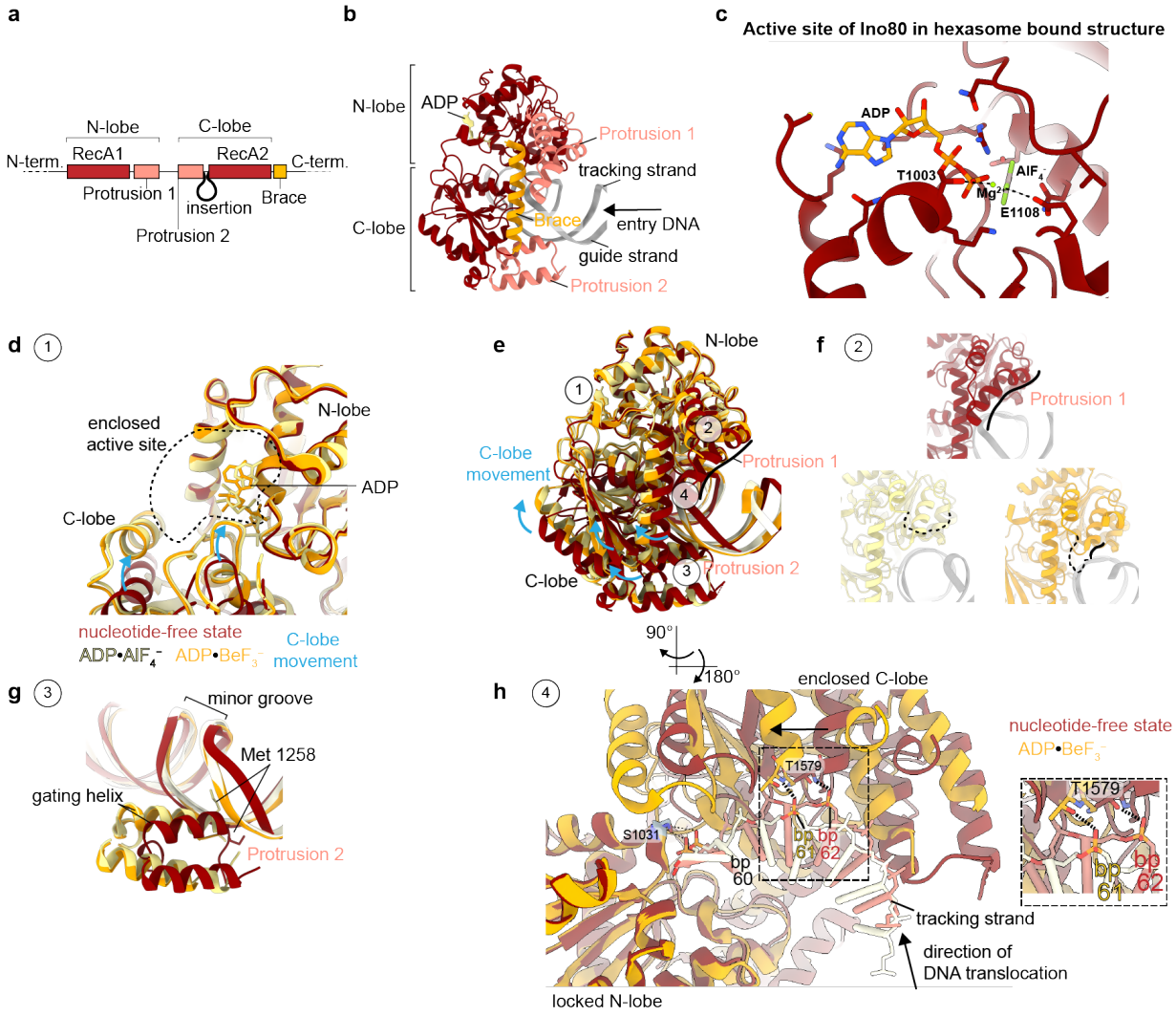


Figure 2.8: **Comparison of Ino80 motor domain in relation to nucleotide state. a** Architecture of the Ino80 motor domain. **b** Structure of the Ino80 motor domain – adapted from the nucleotide-free structure of PDB 8AV6. For visualization an ADP molecule is placed in the active site of Ino80. **c** Active site of the Ino80 motor protein with a bound ADP•AlF_4^- molecule. Residue T1003 contacts the magnesium ion while the as Walker B known residue E1108 binds the AlF_4^-. **d** Active site in presence of different nucleotide analogues. I compared the active sites of nucleotide-free Ino80 (PDB 8AV6, in dark red) with ADP•AlF_4^- (PDB 8OO7, in light yellow) and ADP•BeF_3^- (PDB 8ATF, in orange) bound complexes. Upon binding of ADP•AlF_4^- or ADP•BeF_3^- the active site is closed at the C-lobe interface (indicated with blue arrows). **e** This causes a large movement of the C-lobe. This overview highlights different sites of alterations among the different states: 1) The active site 2) Destabilisation of Protrusion 1 within the N-lobe 3) Insertion of the gating helix of Protrusion 2 into the minor grove 4) Translocation of 1 bp of DNA. Close up views of events 1)-4) visualised elsewhere in the Figure as indicated. **f** Comparison of Protrusion 1 among the different nucleotide states. While the large Protrusion 1 helix is visible in the nucleotide-free state, it is not or only partially present in the ADP•AlF_4^- and ADP•BeF_3^- state respectively. **g** M1258 of the gating helix inserts in the ADP•AlF_4^- state into the minor groove. The same residue is part of the connecting loop in the nucleotide-free state and is not resolved in the ADP•BeF_3^- state. **h** Close up view of the C-lobe interactions with the tracking strand. Shown are only the nucleotide-free and ADP•BeF_3^- state. Residue T1579 interacts with base pair 62 while the C-lobe movement in the ADP•BeF_3^- state causes a shift by one base pair causing an interaction of the residue T1579 with base pair 61. Interaction visualized with dotted black line. Close up of the are on the right in dotted square.

Therefore, comparing the three different Ino80 structures allows to follow the early reaction step of DNA translocation. As introduced earlier, the Ino80 motor domain is composed of the N and C-lobes which bind the entry DNA at SHL−6. (Figures 2.8a and b). Both lobes contain conserved RecA1 (N-lobe) and RecA2 (C-lobe) ATPases which form the core of the two motor lobes. In addition, both lobes contain Protrusion 1 and 2 helices which are involved in DNA binding. The C-terminal region of the C-lobe harbors the Brace helix which folds on top of both lobes – thereby connecting the two lobes. The active site of the Ino80 motor lays in the back part of the N-lobe in relation to the DNA binding interface. ATP binding causes the closure of the active site as detectable in the ADP•AlF_4^- and ADP•BeF_3^- bound structures. In more detail, the closure is facilitated by a movement of the C-lobe towards the N-lobe as depicted in Figure 2.8d. The closure of the active site causes a large rearrangement of the motor lobes (Figure 2.8d, e) as the complete C-lobe moves towards the N-lobe, thereby leading to an overall closure of the motor domain onto the DNA. This movement results in the insertion of the gating helix (Protrusion 2, C-lobe) into the minor groove of the DNA. Along the trajectory of nucleotide free to ATP bound to ATP hydrolysis states M1258 of the gating helix gets inserted into the minor groove (Figure 2.8g). Besides the rearrangements of the gating helix, also Protrusion helix 1 is repositioned upon ATP binding and hydrolysis. While the long helix of the Protrusion 1 is well-ordered in the nucleotide-free state, it is not visible and only partly visible in the ADP•AlF_4^- and ADP•BeF_3^- states respectively (Figures 2.8f). The concrete function is elusive, however, the Protrusion 1 helix is placed close to the guiding DNA strand. Therefore, a role in DNA translocation is likely. Possibly, the DNA gets first translocated along the tracking strand by one base pair. Next, the Protrusion 1 facilitates DNA translocation of the same base pair on the guiding strand. DNA translocation already occurred in the ADP•AlF_4^- and ADP•BeF_3^- states. For clarity the ADP•BeF_3^- and nucleotide-free state are compared in the following as Ino80 is bound to a hexasome in the ADP•AlF_4^- state. The described movement of the C-lobe leads to a large rearrangement of the C-lobe helices close to the tracking strand (Figures 2.8h). As a consequence, the helix tip around residue T1579 changes from an interaction with base pair number 62 to base pair 61. However, the DNA translocation is not yet propagated around the nucleosome core. This is visible as base pair 60 is still bound in both states to serine 1031. I hypothesize, that the release of the phosphate in the ATP cycle repositions the C-lobe back thereby leading to the propagation of the DNA translocation.

2.2.5 Alternative binding modes of the DNA binding domain of the Arp5 core

As described in the model building section (Section 2.2.2), two distinct states of the INO80-hexasome complex were classified. The two states differ in their globular architecture. This becomes apparent when the two states are aligned on the hexasome. While the motor domain location is fixed, the Arp5/Ies6 grip as well as the Rvb1/Rvb2 hetero-hexamer change their conformation between the two states. In more detail, especially the Rvb1/Rvb2 stator element is placed closer to the hexasome in state 1 compared to state 2 (Figure 2.9a). On a molecular level, the apparent movement of the Arp5 grip, translates into a different position of the DNA binding domain (DBD) of the Arp5 core on the hexasomal DNA. The DNA binding domain consists of a small hairpin between residues 90-96. While the DBD in state 1 binds to the major groove between SHL0 (dyad) and SHL+1, it is inserted in state 2 into the minor groove at SHL+1. Therefore, not only the position changes but also the binding mode alters from major groove to minor groove binding (Figure 2.9b, c). The interaction is in both cases mainly facilitated by residue R92, which contacts the DNA sugar backbone (Figure 2.9d and e). Overall, the DBD consists of many positively charged residues – mainly arginines and lysines. A selection of these residues was mutated to alanines to assess the role of the DBD in hexasome and nucleosome remodeling (Figure 2.9f, g) (K88A/R90A/R92A/K93A/K96A/R112A/R116A). As a result of the introduced non-charged residues in the DBD-mutant, a similarly reduced INO80 activity is

observed on both hexasomes and nucleosomes. Likely, the reduction arises from the decreased binding capability indicating a role of the Arp5/Ies6 grip in progressive DNA translocation.

2.2.6 H3/H4 tetramer is exposed in the context of hexasomes

The captured hexasome has a footprint of only 101 bp of DNA (Figure 2.10b). This is significantly less compared to a nucleosome where approximately 147 bp of DNA wrap around the histone core. The decreased DNA footprint arises from the loss of one H2A/H2B copy resulting in unwrapping of 46 bp of DNA off the histone core. Looking at the hexasome from the side, one can acknowledge that the hexasomal particle consists only of one single DNA gyre compared to two within a nucleosome. In the presented structure, the proximal H2A/H2B copy is missing. This is remarkable, as the distal H2A/H2B copy is the one being solvent exposed.
Thereby, the underlying H3/H4 tetramer becomes solvent exposed instead. Both the smaller DNA footprint and the exposed H3/H4 tetramer constitute an altered interaction surface for INO80 as outlined in Section 2.2.3.

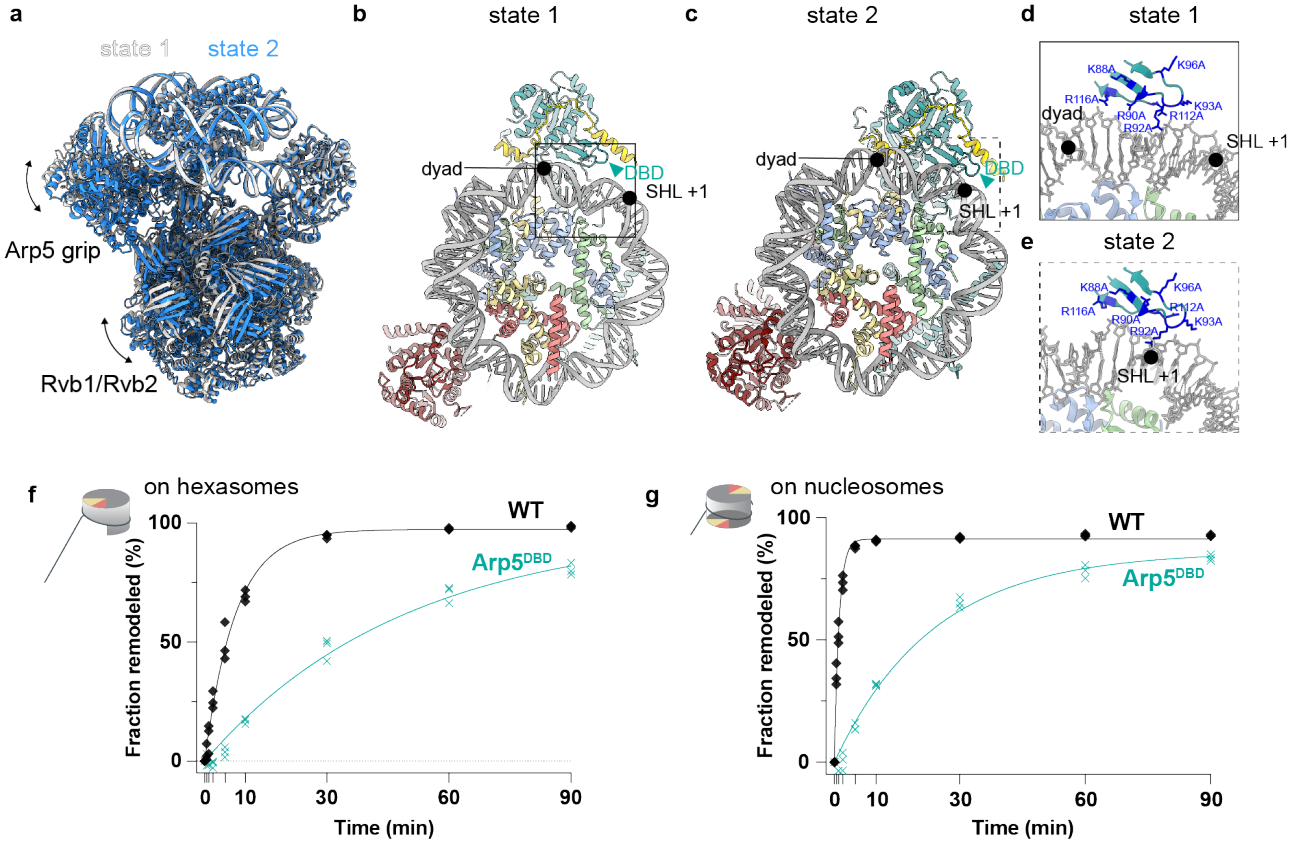


Figure 2.9: **Dual binding modes of the Arp5 DBD domain. a** View of state 1 (light gray) and state 2 (blue). The structures were aligned on the hexasome. Arp5 grip and Rvb1/Rvb2 subunits differ in their location in between the two states. **b, c** Bottom view of the hexasome with Ino80 and Arp5 core. Arp5 DBD domain binds to major groove between the dyad and SHL+1 in state 1 (**b**), while the same domain inserts into the minor groove close to SHL+1 (**c**). Dyad and SHL+1 are highlighted with black dots. **d,e** show close-up views of the DBD as indicated by dotted boxes in b and c. Residues of the DBD-mutant are shown in blue and labeled respectively. **f, g** Sliding assays of wt *C. t.* INO80 (black graph) and DBD-mutated INO80 (cyan graph) on 0H80 hexasomes **f** and 0N80 nucleosomes **g**.

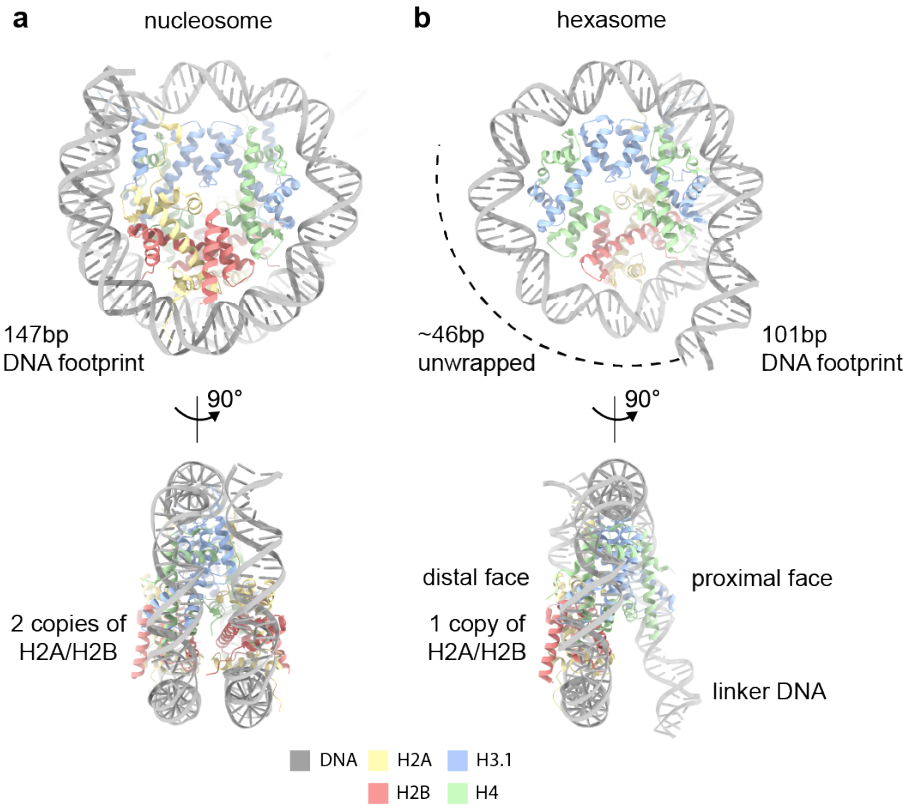


Figure 2.10: **Comparison of nucleosome and hexasome structure. a** Top and side view of a nucleosome (PDB 8AV6). The nucleosome has a DNA footprint of 147 bp. The side view reveals the presence of two gyres of DNA that wrap around two copies of H2A/H2B and one H3/H4 tetramer. **b** Same view for hexasome (this study). The absence of one H2A/H2B leads to the unwrapping of 46 bp of DNA (depicted as dotted line) leaving a DNA footprint of 101 bp. In addition, the side view at the bottom illustrates that the proximal H2A/H2B copy is missing which leads to the presence of only one gyre of DNA.

2.2.7 INO80 spin-rotates upon hexasome binding

INO80 binds hexasomes differently compared to a nucleosome substrate. To illustrate this, we first need to look at the binding mode of INO80 on nucleosomes. Therefore, I focus on the two main hexasome interacting subunits – the Ino80 motor domain as well as the Arp5-core and Arp5-grappler elements. As shown by Eustermann et al. 2018 [77], the Ino80 motor domain binds to SHL-6 which lays close to the entry DNA of the nucleosome (Figure 2.11a). The Arp5 grip binds to SHL-3 which is about three helical turns towards the C-terminus of the motor domain binding site. As a secondary contact site of Arp5, its cross shaped grappler domain is placed close to the dyad. As described previously, the proximal H2A/H2B copy is missing in the hexasome when bound to INO80. The proximal H2A/H2B is located in between the Ino80 motor domain and the Arp5 grip. As the loss of the H2A/H2B copy leads to the unwrapping of 46 bp of DNA, SHL-6 lays now in the unwrapped and therefore disordered part of the DNA. As a consequence, the entry site of the hexasome is shifted in comparison to the nucleosome. Therefore, the Ino80 motor lobes bind instead to superhelical location −3 (Figure 2.11b). Besides the DNA unwrapping caused by the missing H2A/H2B dimer, an additional 8 bp are lifted off the histone core by the Ino80 motor. The novel binding mode is remarkable, as it was suggested earlier [75][2] that ATP-dependent chromatin remodelers bind either

to SHL+/-2 or SHL-6. Therefore, SHL-3 represents a novel binding site for chromatin remodelers. In addition, it was not seen before, that a chromatin remodeler has a dual binding site on a nucleosome-like particle. The hexasome with its smaller DNA footprint represents a non-canonical nucleosome [55][56]. Binding to the hexasome is facilitated by INO80 through an altered binding mode to adapt to the new interfaces present in the hexasome.
The Ino80 motor lobes are not the only domain changing its contact site in context of hexasome binding. The reason for this lays in the fixed motor/rotor/stator/grip architecture of INO80 as introduced in Section 1.2.

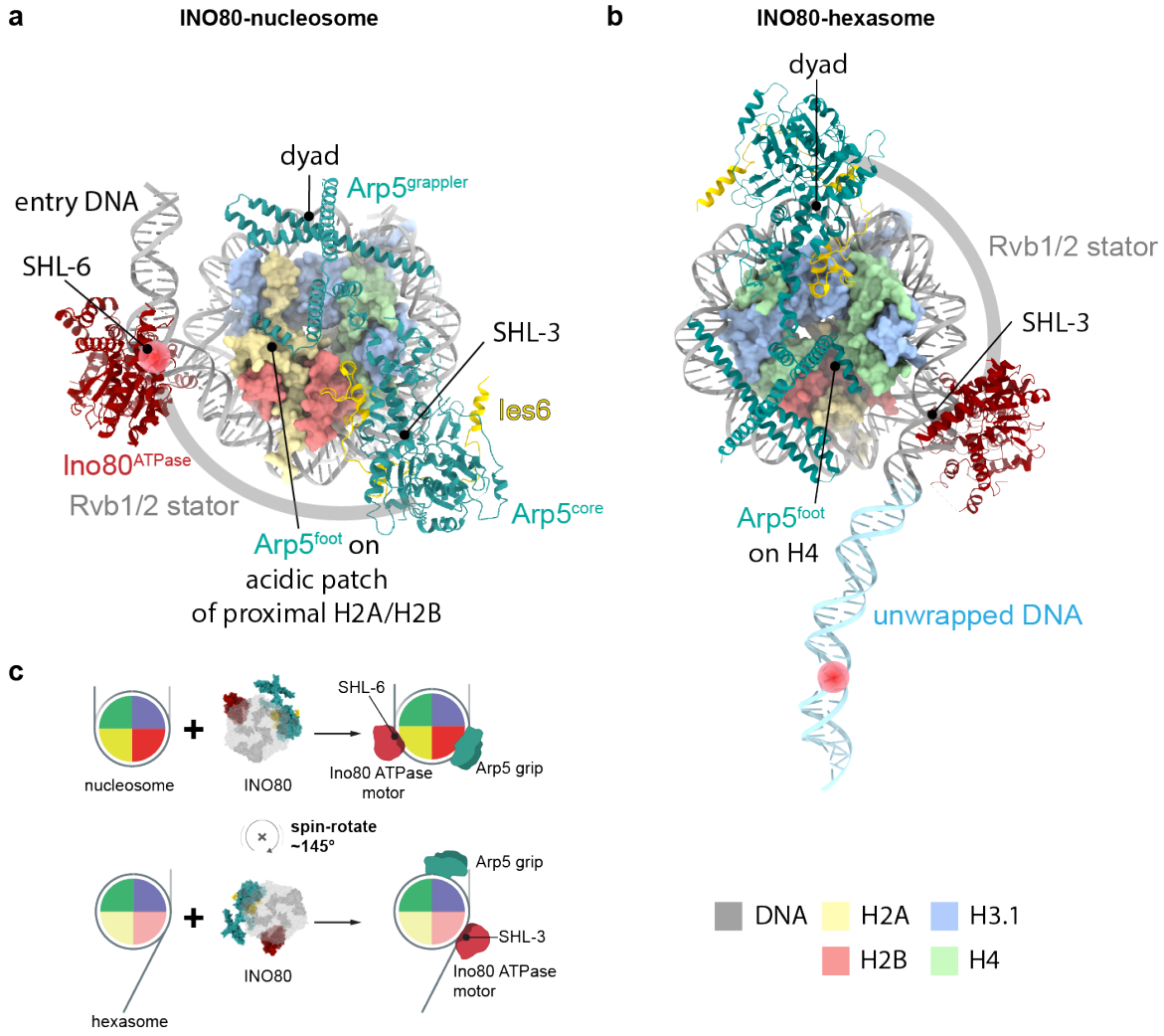


Figure 2.11: **Dual binding mode of INO80 on nucleosomes and hexasomes. a** INO80 bound on nucleoesome substrate. For clarity only the Ino80, Arp5 and Ies6 subunits are illustrated. The Ino80 motor domain binds at SHL position -6, while the Arp5/Ies6 grip is placed at SHL-3. Location of the Rvb1/2 stator is indicated with a gray bar. **b** Same representation of INO80 on a hexasome. The hexasome is shown in the same orientation compared to the nucleosome on the left (see dyad position). As a consequence of the DNA unwrapping off the hexasome core particle, SHL-6 lays in the unwrapped DNA part (red dot). Therefore, the Ino80 motor binds instead to SHL-3 and the Arp5/Ies6 to the dyad. The unwrapped DNA (colored in light blue) was modeled as B-DNA for illustration. **c** Schematic representation of the different INO80 binding modes. INO80 undergoes a 145° spin rotation to bind the hexasome in relation to a nucleosome.

The Rvb1/Rvb2 stator element places the motor and grip domains of INO80 about three helical turns apart onto the canonical and non-canonical nucleosome substrates. A change in the binding site of the Ino80 motor therefore results into an adapted mode of binding of the Arp5 elements. In more detail, the Arp5 grip binds the hexasomal DNA at SHL0 (the dyad) – again three helical turns away from the motor domain. This places also the Arp5-grappler to a new location. While bound to the dyad on nucleosomes, it switches to SHL+3 which is nearby the newly exposed H3/H4 histones. Therefore, the grappler is placed where the proximal H2A/H2B dimer is located on a nucleosome. On a hexasome, the Arp5 grappler element binds to the newly exposed DNA which was covered by the second gyre of DNA before. The latter does not exist on a hexasome due to the unwrapping of DNA. On a globular level, the change of binding site means that the INO80 complex binds to the hexasome in a 145° spin-rotated mode (Figure 2.11c). This does not only change the interaction sites of INO80 to the DNA but denotes that the INO80-histone core interface is different. Consequently, the Arp5-grappler element recognizes the H3/H4 interface as described in Section 2.2.6. In more detail, the Arp5-foot folds alongside the H4α2 helix. As a consequence of the hexasome binding mode, DNA is translocated now alongside the H3/H4 interface towards the Arp5 grip. In contrast, the DNA is translocated on a nucleosome at the interface close to the proximal H2A/H2B dimer – the one missing in the hexasome structure.
In summary, I showed that INO80 adopts a dual binding mode on hexasomes and nucleosomes. On top of that, INO80 binds the hexasome at SHL-3 – a novel binding site in the field of chromatin remodelers. Other chromatin remodelers - except for INO80 – bind on SHL+/-2 [72][106][122][121]. Hexasome binding of INO80 is facilitated by a 145° spin-rotation which positions the complex in a novel histone context. Therefore, the Arp5 foot is placed onto the exposed H3/H4 interface.

2.2.8 Sliding of hexasomes by INO80 is independent of the acidic patch

The spin-rotated binding mode of INO80 places the Arp5-grappler element onto the exposed H3/H4 interface and the nearby DNA (Figure 2.12a, b). In the context of hexasomes, a stretch of DNA from SHL+2 to SHL+4 gets exposed due to the unwrapping of DNA. This stretch of DNA is bound by the N-terminal helix of the cross-shaped grappler element. At the same time, the tip of the N-terminal helix points towards the new entry DNA site at SHL-3. Overall, the Arp5 grappler is stabilizing the exposed H3/H4 interface. For the INO80-nucleosome complex, two distinct conformations of the Arp5 grappler were observed [150]: one inhibitory and one permissive. Here, I only observed the permissive conformation indicating that the INO80-hexasome complex was trapped in its sliding prone state.
In our structure, the Arp5-foot helix that extends from the Arp5 grappler, is placed onto the H3/H4 interface. The same helix is located at the acidic patch of the proximal H2A/H2B dimer in context of nucleosome binding. In more detail, the Arp5-foot folds against the H2A-α2 helix on a nucleosome (Figure 2.12c). In addition, the Arp5-heel which connects the rest of the Arp5 grappler with the foot, is in close contact with the H2A-α2 helix. Due to the lack of the proximal H2A/H2B dimer and the spin-rotated binding mode of INO80, the same Arp5 helix folds on the hexasome against the H4α2 helix (Figure 2.12d). This interface lays close to the H3α1L1 elbow which is specifically recognized by Rvb2 (see Section 2.2.3). The H3 elbow is replaced in the nucleosome structure by the H2Bα1L1 elbow (compare Figure 2.12c, d). In conclusion, altered interactions of INO80 make up for the missing H2A/H2B dimer which is mainly facilitated by the Arp5 subunit. The Arp5 subunit has a dual substrate specificity for both the acidic patch of H2A/H2B and the H3/H4 interface. Thereby, INO80 recognizes the hexasome specifically by using similar architectural features to its nucleosome binding mode.
In order to probe the specificity and relevance of the Arp5-foot interaction on the hexasome, I

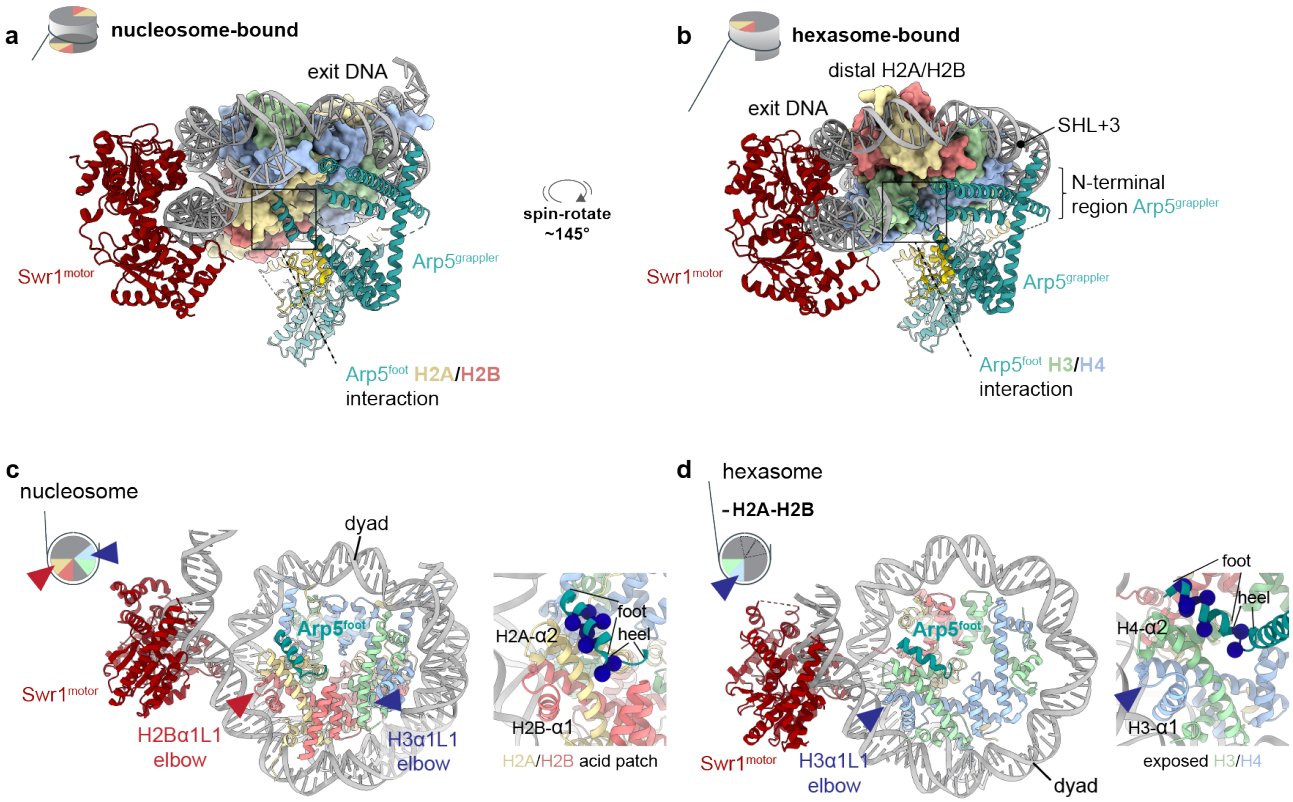


Figure 2.12: **Arp5-foot helix interactions on nucleosomes and hexasomes. a** Arp5-foot interaction on hexasome. The foot helix binds to the acidic patch of the proximal H2A/H2B as marked with a square. **b** Spin rotation of INO80 places the foot helix on the exposed H3/H4 interface. For better representation, the hexasome was rotated compared to the nucleosome on the left. **c** Bottom view of the nucleosome. The foot helix colored in cyan interacts with the H2Aα2 helix of the proximal acidic patch. H2Bα1L1 and H3α1L1 elbows are indicated by arrows on the structure and the pictogram at the top. Close-up on the right indicates what residues have been mutated for the sliding assays. **d** Same representation for the hexasome. Arp5 foot helix folds against the H4α2 helix on the exposed H3/H4 interface. The spin rotated binding mode of INO80 places the foot close to the H3α1L1 elbow.

performed activity assays of INO80. Therefore, I used the introduced sliding assay (see Section 2.1.2) which allows to probe different hexasome, nucleosome and INO80 substrates. First, I tested an Arp5-heel mutant where only mutations to R501 and K502 were introduced (R501E and K502E). Secondly, I performed the same experiment with INO80 where not only the heel residues but also some foot residues were mutated (R501/K502E/Q507A/R509A/M510S/K511A/I513S) (Figure 2.13a). All experiments were run in triplicates on 0H80 and 0N80 substrates. In addition to the INO80 mutants, wild type INO80 was used for comparison.

When either of the Arp5 mutants is used to probe the sliding activity on nucleosomes, a reduced activity is observed. Furthermore, The Arp5-heel mutant retains some activity, while the Arp5-heel+foot mutant abolished its sliding activity almost completely (Figure 2.13c). On contrary, the same INO80 Arp5 mutants on 0H80 hexasomes not only maintain the activity, but show an even higher sliding activity compared to wild type INO80 (Figure 2.13d). Taken together, the Arp5-heel and foot have a distinct function on nucleosomes and hexasomes. While both elements are substantial for successful nucleosome sliding, they have a regulatory function on hexasomes. Therefore,

the sliding activity is increased upon mutation of the Arp5-foot. A similar phenomena is observed for INO80 on nucleosomes: an H2A.Z mimicking mutant of H2A that is located close to the Arp5 foot leads also to an increase in sliding activity [77].
Next, I wanted to know if hexasome sliding by INO80 is dependent on the remaining distal acidic patch of the H2A/H2B dimer. The novel placement and function of the Arp5-foot suggests that hexasome sliding is independent on the remaining acidic patch. In order to probe the activity, I used the very same sliding assay – this time using nucleosomes and hexasomes comprising acidic patch mutations on H2A (E61A/E64A/D72A/D90A). These mutations alter the H2A surface from a negatively charged surface to a neutral electrostatic potential. INO80 shows no activity on the acidic patch mutated nucleosomes as previously described in the literature [77]. However, the sliding activity of INO80 on the hexasome acidic patch mutant, remains at a similar level compared to wild type hexasomes. Consequently, hexasome sliding is indeed independent of the acidic patch which confirms the INO80-hexasome structure.

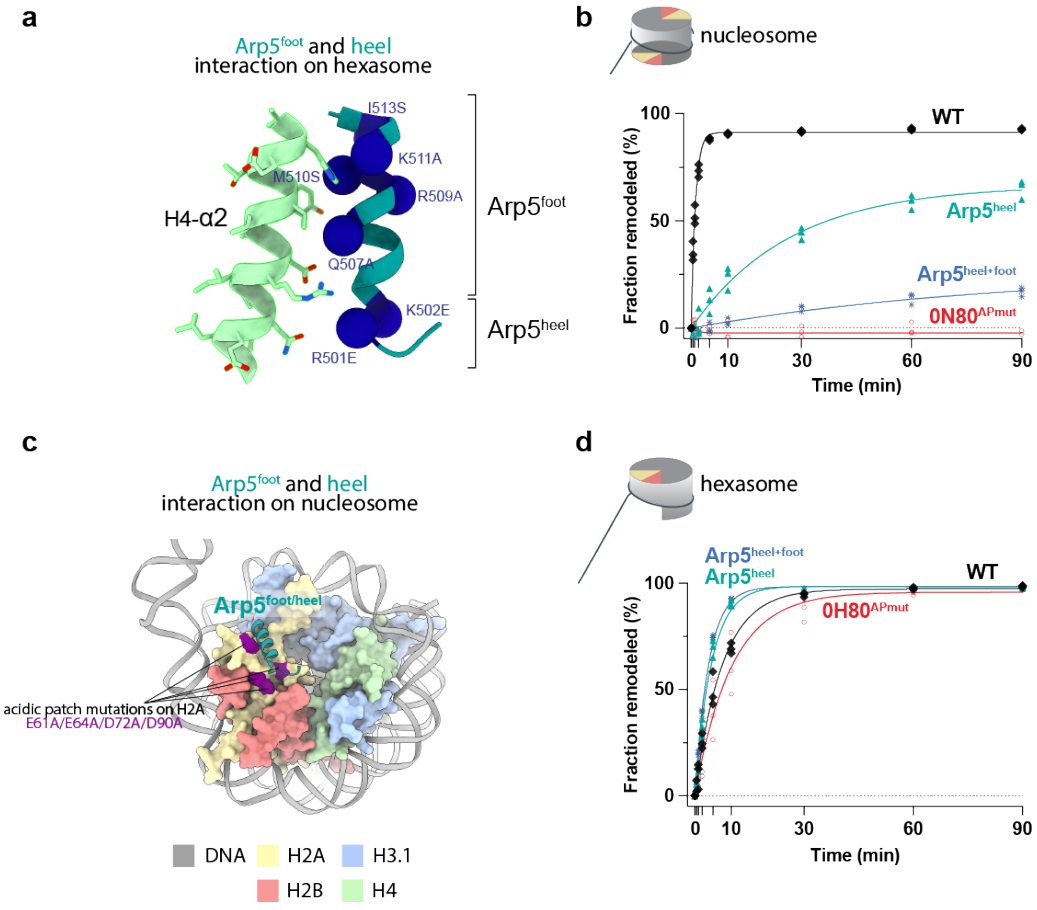


Figure 2.13: **Sliding assays of Arp5-foot and acidic patch mutants. a** Close-up of the Arp5-foot (cyan) and the H4-α2 helix (green). Mutated residues on the foot are labeled respectively and are represented as blue spheres. **b** Sliding assays of wild type INO80 or Arp5 mutated INO80 on 0N80 nucleosomes and 0N80-acidic patch mutated nucleosomes. Activity is measured as fraction of remodeled nucleosomes in percent. **c** Bottom view of a nucleosome showing the proximal H2A/H2B and the Arp5-foot helix as it interacts with the acidic patch of the nucleosome. Acidic patch mutated residues are colored in purple. **d** Likewise sliding assay on 0H80 and 0H80-acidic patch mutated hexasomes.

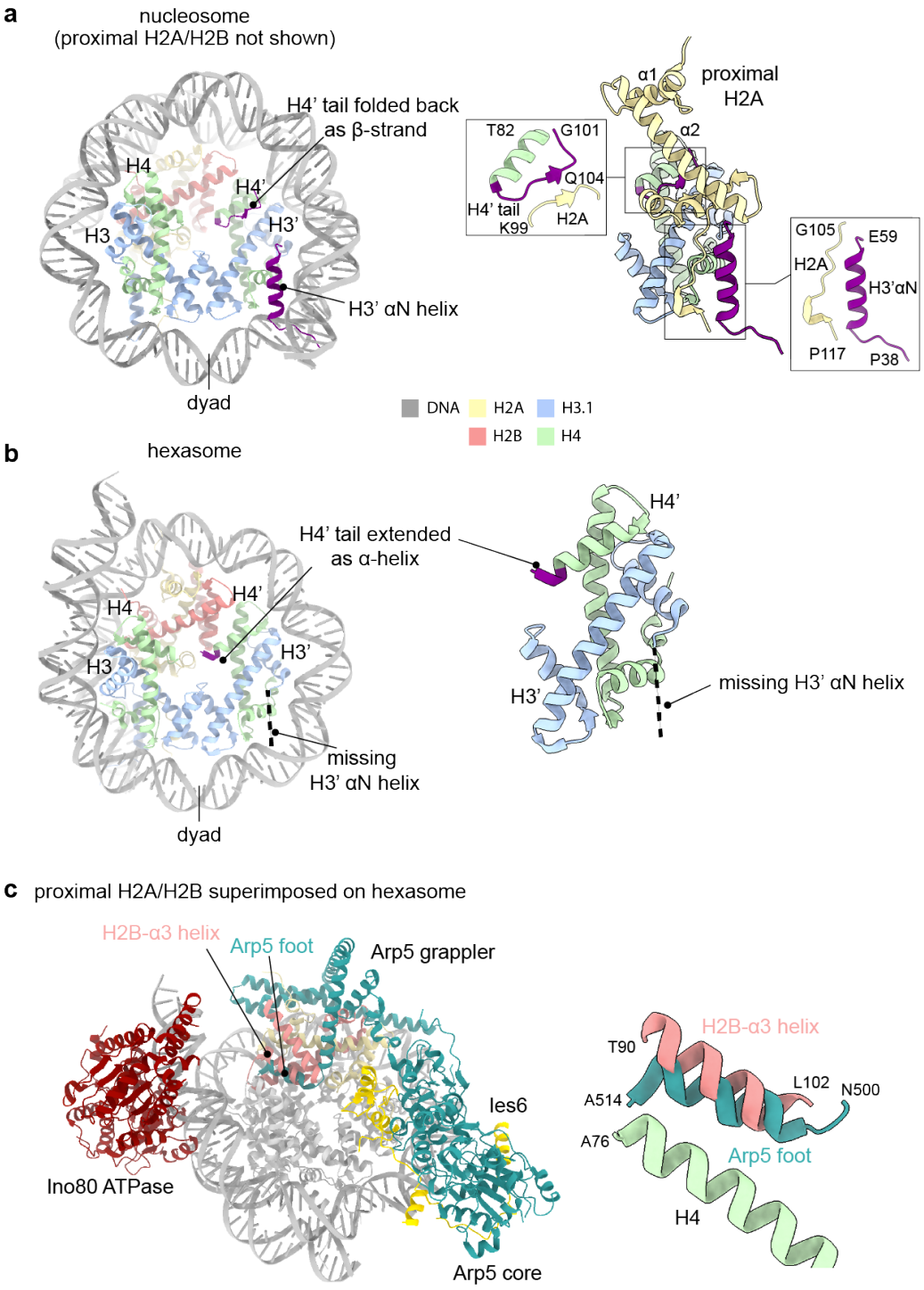


Figure 2.14: **Altered features of the hexasomal histone core. a** On a nucleosome the H4' tail is folded back as β-strand (in purple) interacting with the β-strand from the proximal H2A. The H3' αN helix (also in purple) is ordered as it folds against H2A. *Left:* Overview of the location of H4' and H3' features on the nucleosome (proximal H2A is not shown for clarity). *Right:* Close-up views of H4' and H3' features interacting with the proximal H2A copy. H3 and H4 on the left side of the nucleosome are in contact with Arp5, while H3' and H4' are the copies on the right. **b** Same view of a hexasome (*left*). Due to the missing proximal H2A, the H4' tail undergoes a β-α switch and extends as a helix into the cavity of the hexasome. H3' αN helix is unstructured and is not visible in the EM map. **c** Superimposition of the proximal H2A/H2B onto the hexasome structure. The proximal H2B-α3 helix of a nucleosome gets replaced by the Arp5-foot helix in the context of hexasome binding of INO80. Close-up shown on the right. H4-α2 helix interacts with the Arp5-foot helix on hexasomes (in green).

2.2.9 H3 and H4 constitute altered features due to the absence of one H2A/H2B copy

The absence of one H2A/H2B copy has also consequences on the hexasome architecture itself. These changes occur on the H3 and H4 copies that are in contact with the proximal H2A/H2B copy in the context of a nucleosome. I denote this H3 and H4 copy as H3' and H4'. The copies of H3 and H4 close to the Arp5-foot, that are interacting mainly with the distal H2A/H2B copy, remain denoted as H3 and H4. On a nucleosome the C-terminus of the H4' tail folds back to the H4' helix as a β-strand. Its shape is stabilized by an interaction of the H4' β-strand with the β-strand of the proximal H2A docking domain (Figure 2.14a). In proximity, the same H2A docking domain stabilizes the N-terminal helix of H3' denoted as H3' αN helix. As the proximal H2A is missing on a hexasome, the described H3' and H4' histones get restructured showing a so far unanticipated plasticity of the histone core structure. The missing H2A docking domain leads to an extension of the H4' tail helix. In more detail, the C-terminus of the H4' tail undergoes a β-α switch. The H4' helix therefore reaches into the central cavity of the hexasome, filling the characteristic hole of the histone core (Figure 2.14b). In addition, the H3'αN helix gets unstructured and is no longer visible in the EM density map. It was stabilized by the H2A docking domain before, which is missing in the hexasome structure. The missing proximal H2A/H2B copy exposes the whole H3/H4 interface and thereby the complete H3/H4 tetramer. The new interface is stabilized by multiple INO80 features as described earlier in the thesis. The loss of the proximal H2A/H2B is thereby compensated by a closer interaction of the Rvb1/Rvb2 stator and novel interaction sites of the Arp5 subunit. These are possible due to the spin-rotated binding mode of INO80. The latter results into a replacement of the proximal H2B-α3 helix with the Arp5-foot (Figure 2.14c). In that way, the Arp5-foot mimics the H2B-α2 helix in the context of hexasome binding. As a consequence, the H3/H4 interface gets stabilized by the Arp5-foot. In the context of a nucleosome the H4α2 helix was folded against the H2B-α3 helix. Now, on the hexasome, the same H4 helix is stabilized by the Arp5-foot. This emphases once more the specificity of the INO80-hexasome binding as well as the dual binding mode of the Arp5-foot helix.

2.2.10 Competition assays - Limited amount of extra-nucleosomal DNA makes hexasomes the preferred INO80 substrate

I showed previously that INO80 is active on a hexasomal substrate (Section 2.2.3). However, the question remains if INO80 has a preference for either nucleosome or hexasome substrates. In order to probe the activity of INO80 on either of the two substrates, I performed sliding assays as described earlier. In more detail, I probed 0N80 (601 with 80 bp of extra-nucleosomal DNA) and 0H80 which have been used for the structure determination. However, the 0H80 hexasomes contain 40 bp more extra-nucleosomal DNA in relation to the 0N80 nucleosomes due to the unwrapping of DNA from the hexsomal histone core.
Therefore, I probed the activity also on 0H40 hexasomes which contain beside the 40 bp of extra-nucleosomal DNA, 40 bp additional base pairs in the linker DNA that unwrapped from the histone core. Consequently, the 0H40 are comparable with 0N80 nucleosomes in terms of their extra-nucleosomal DNA length (Figure 2.15) as they comprise as well 80 bp as linker DNA. I observed for *C. t.* INO80 a faster kinetic for 0N80 compared to the two different hexasome substrates indicating that nucleosomes are the preferred substrate of *C. t.* INO80 (illustrated in Methods Section 4.1). For yeast INO80 a similar sliding activity was observed for 0N80 and 0H40. 0H80 with its longer linker DNA showed here a faster kinetic compared to the other two substrates.
This type of experiment is however limited, as the sliding assays are performed separately on hexasomes and nucleosomes. As another aspect, I observed more free DNA in the hexasome samples

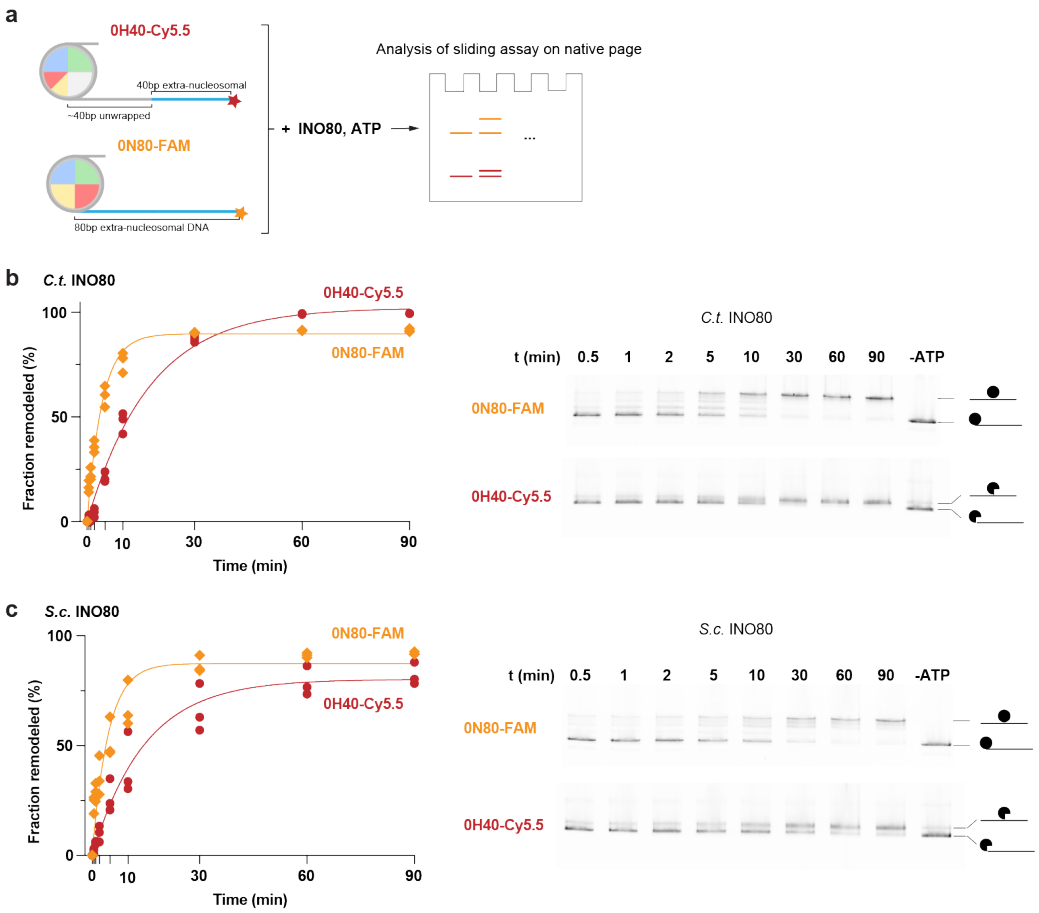


Figure 2.15: **Competition assay of nucleosome and hexasome substrates. a** Pictogram of 0H40 and 0N80 substrates. They contain the same length of extra-nucleosomal DNA due to the 40 bp of DNA unwrapping from the hexasomal histone core. Activity assays of 0H40 and 0N80 are performed in the same tube. Band shifts of remodeled vs non-remodeled substrates are detected for the hexasomes via Cy5.5 fluorescence and in case of 0N80 with fluorescein (6-FAM). **b** Competition assay of *C. t.* INO80 on 0N80 and 0H40 substrates. Native PAGE scan on the right was quantified by measuring the intensity of unremodeled substrates over time and plotting the fraction of remodeled substrates (graph on the left). **c** Same as in **b** but with *S. c.* INO80.

which competes with the hexasome for INO80 binding. I therefore devised a competition assay in which sliding of Cy5.5 labeled nucleosomes and fluorescein labeled hexasomes are probed in the same reaction tube. While I conducted the experiment, Min Zhang kindly provided the labeled hexasomes and nucleosomes. The competition assay was performed likewise to the sliding assays, but now using a limited amount of INO80 in comparison to the nucleosomal and hexasomal substrates. I subsequently ran the samples on Native PAGE and imaged the gel for Cy5.5 and fluorescein fluorescence in two consecutive steps. I used 0N80 and 0H40 with comparable extra-nucleosomal DNA to rule out a preference for a certain linker length. In this case, one can directly probe the inherent preference for the altered histone interface. Here, for both *C. t.* and yeast INO80 a faster kinetic was observed for the nucleosome substrate (Figure 2.15b, c). The saturation point of the remodeling reaction on nucleosomes is already reached after about 10 min while it takes almost 1 h on the hexasome substrate. Interestingly, *C. t.* INO80 remodels 0H40 hexasomes close to completion while yeast INO80 does not reach the same level. In summary, nucleosomes are the preferred substrate

of INO80 when the amount of extra-nucleosomal DNA is comparable. A reason for the observed kinetics is possibly the second acidic patch of INO80 and the interaction with the Arp5-foot. As described earlier, the Arp5-foot has a regulatory role when binding the H3/H4 interface on hexasomes. Hence its presence makes the remodeling reaction less efficient and overall slower compared to the remodeling of nucleosomes.
Next, I wanted to know if conditions exist, where the hexasome becomes the preferred substrate of INO80. Therefore, I chose substrates with limited amount of extra-nucleosomal DNA. It is known from the literature that INO80 requires at least a length of 60 bp [194] to place the Arp module on the DNA connecting two nucleosomes. In this manner, INO80 senses DNA shape features [36][169] and couples its ATPase activity to productive nucleosomes sliding. I chose to compare 0N60 nucleosomes with 0H20 hexasomes which again comprise the same length of extra-nucleosomal DNA (Figure 2.16a). In this case, the preference of INO80 for nucleosomes over hexasomes becomes even more apparent. Nonetheless, half of the hexasomes are still remodeled within a time span of 90 min. In contrast to the 0H20 hexasomes, INO80 is almost completely inactive on 0N20 nucleosomes. This resembles a scenario within yeast gene bodies were nucleosomes are spaced by about 18 bp of DNA [283]. Here, INO80 is not active but when a hexasome is formed, the unwrapped 40 bp of DNA create a linker length of 60 bp of DNA - enough for INO80 to translocate DNA and slide the hexasome. Therefore, hexasomes become a preferred substrates in packed chromatin regions.

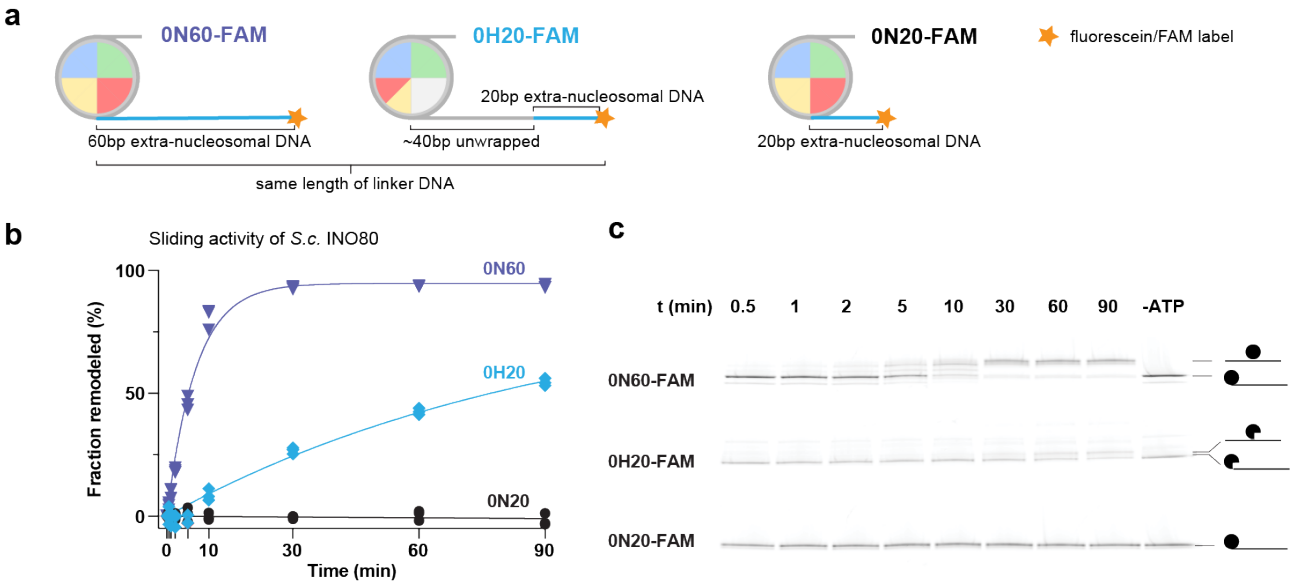


Figure 2.16: **Limited amount of extra-nucleosomal DNA makes hexasomes the preferred substrate. a** Schematic illustration of used substrates: 0N60, 0H20 and 0N20. 0N60 and 0H20 harbor the same amount of extra-nucleosomal DNA. **b** Sliding assays of yeast INO80 on 0N60, 0H20 and 0N20. Sliding assays were run separately in this case. Analysis of remodeled substrates (*left*) of the Native PAGEs (*right*). Substrates were 6-FAM labeled.

2.3 Biochemical characterization of yeast and *Chaetomium thermophilum* SWR1

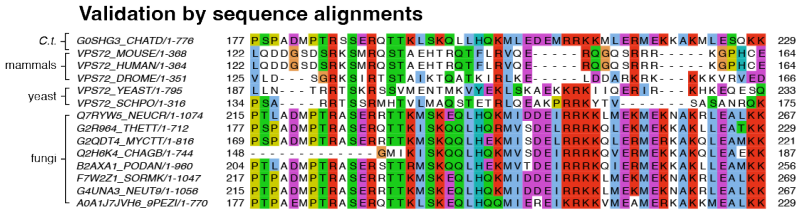
In the following sections, I am focusing on the biochemical and structural characterization of the SWR1 complex aiming to understand the mechanism of its histone exchange reaction (outlined in aims of this thesis Section 1.8). In my structural studies, I started with the Arp module of SWR1 which enabled the interpretation of my full-length apo-SWR1 cryoEM data. Finally, I screened SWR1 samples in complex with a nucleosome for structure determination by cryoEM. I used SWR1 from *Chaetomium thermophilum* (hereafter: *C. t.*) for the structural characterization of SWR1, while I used SWR1 from *Saccharomyces cerevisae* (hereafter: *S. c.*) for genome wide biochemical studies to understand how SWR1 reads out genomic information to enable its substrate specificity of +1 nucleosomes.

a

S.c. SWR1 **Blast** → *N.c.* SWR1 **Blast** → *C.t.* SWR1

	S.c. SWR1	*C.t.* SWR1
Swr1	Q05471	G0RXU5
Swc2	Q03388	G0SHG3
Swc3	P31376	G0S2Q0
Swc4	P53201	G0S9A4
Swc5	P38326	G0SHE0
Swc6	Q03433	G0SGS0
Swc7	Q06707	???
Rvb1	Q03940	G0RYI5
Rvb2	Q12464	G0RWC2
Actin	P60010	G0SE15
Arp4	P80428	G0SBW4
Arp6	Q12509	G0S9Q2
Bdf1	P35817	G0RYX1
Yaf9	P53930	G0S5N1

b



c

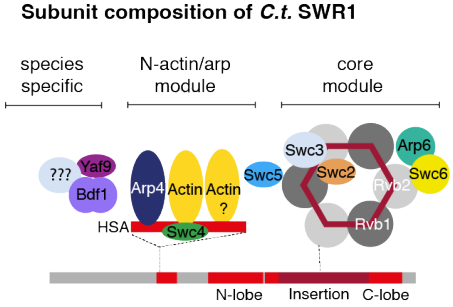


Figure 2.17: **Subunit identification of *C. t.* SWR1. a** Table showing the uniprot identifiers for *S. c.* SWR1 and the corresponding ones for *C. t.* SWR1 that were identified using Blast®® (National Library of Medicine) searches. **b** Exemplary sequence alignment of *C. t.* Swc2 against mammalian, yeast and fungi sequences for validation purposes. Plot was generated in Jalview [284][285]. **c** Subunit composition of *C. t.* SWR1. Complex is divided in Core, Arp and species-specific modules.

2.3.1 Characterization of *Chaetomium thermophilum* SWR1

The composition of the *C. t.* SWR1 is incompletely annotated and is not validated as the *C. t.* proteome was automatically annotated [286]. Therefore, I had to start with identifying the subunit composition of *C. t.* SWR1 and validate the underlying protein sequences. As *S. c.* SWR1 is well annotated and characterized biochemically and structurally, I utilized its amino acid sequences to identify the *C. t.* homologues. I performed sequence similarity searches by using Blast®® as well as the database from the Bork lab [286].

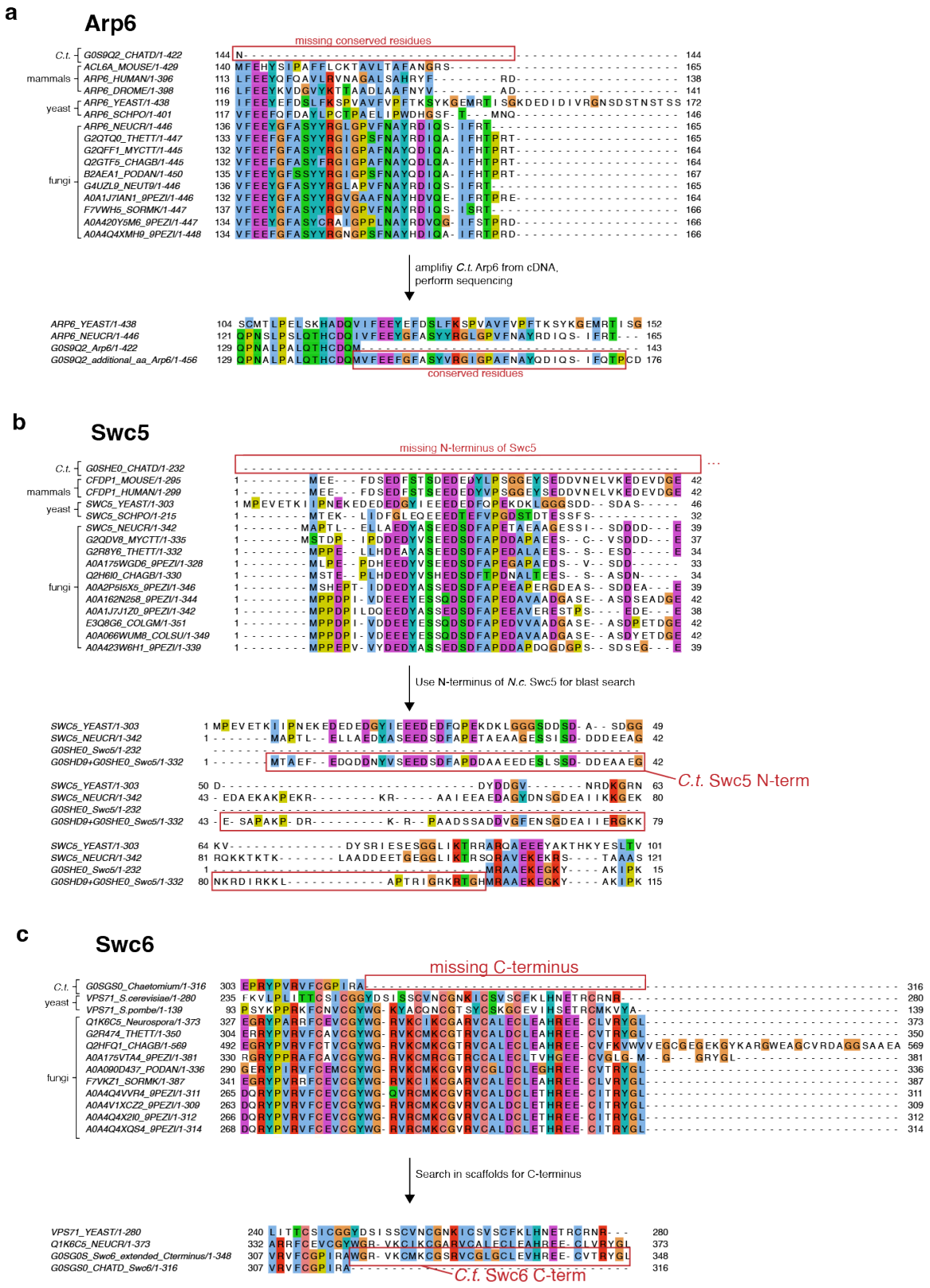


Figure 2.18: **Investigation of *C. t.* Arp6, Swc5 and Swc6 amino acid sequences. a-c** show at the top, sequence alignments against mammals, yeast and fungi sequences of Arp6 **a**, Swc5 **b** and Swc6 **c**. While *C. t.* Arp6 lacks a stretch of conserved amino acids within the protein, *C. t.* Swc5 misses a part at its N-terminus and Swc6 on the C-terminus. Bottom parts of each panel represent alignments of the respective yeast subunit against *Neurospora crassa*, *C. t.* and the newly annotated *C. t.* subunits.

As an intermediate step, I identified the SWR1 subunits from *Neurospora crassa* (*N. c.*) which is a fungi relative of *C. t.* and is well characterized. Based on these sequences, another round of Blast® searches was executed to identify the respective subunits in *C. t.* This method worked for 13 out of the 14 known yeast SWR1 subunits (Table with uniprot identifiers in Figure 2.17a). A Swc7 orthologue could not be identified in any other species but *Saccharomyces* yeast. Conclusively, the amino acid sequence of Swc7 is yeast-specific. Nonetheless, a homologue to Swc7 could exist in *C. t.* that does not share its amino acid sequence but 3D structure. I therefore used pulldown experiments of *C. t.* SWR1 in addition to bioinformatic tools to investigate the presence or absence of a Swc7-like subunit (see Section 2.3.5) in *C. t.* SWR1.
As the *C. t.* proteome was automatically annotated [286], I had to validate the identified SWR1 amino acid sequences. For this purpose I performed sequence alignments (ClustalO in Jalview [287][284][285]) against human, mouse, drosophila, yeast as well as 10 close relatives of *C. t.* (example in Figure 2.17b). The alignments revealed that the three subunits Arp6, Swc5 and Swc6 are incompletely annotated. In more detail, an internal part of about 30 amino acids was clearly missing in *C. t.* Arp6 as the 10 close relatives of *C. t.* showed a high degree of conservation in that region (Figure 2.18a). I therefore amplified Arp6 from a cDNA library that was created by Olga Kolesnikova for the subsequent cloning of the recombinant SWR1 complex. Sequencing of the amplified Arp6 sequence, revealed that *C. t.* shares the high degree of conservation with its fungi relatives. In contrast to Arp6, the sequence alignments of Swc5 with vertebrate and fungi homologues, indicated that a stretch of about 90 amino acids is missing from Swc5's N-terminus (Figure 2.18b). I therefore used the seemingly conserved N-terminal region of *Neurospora crassa* Swc5 and submitted it to another round of Blast® search. This revealed that the N-terminal region of *C. t.* Swc5 was falsely annotated as a separate protein. Swc5 consists of the two uniprot entries G0SHD9 (N-terminus) and G0SHE0 (C-terminus). Lastly, the sequence alignments for Swc6 indicated that a stretch of 30 amino acids was missing in comparison to *Neurospora crassa* and other fungi relatives. This part is particularly well conserved in fungi (Figure 2.17c). Thus, I used the nucleotide sequence of the conserved region to search in the scaffolds of the *Chaetomium thermophilum* genome using the database from the Bork lab [286].
In summary, the usage of Blast® allowed me to identify the *C. t.* SWR1 composition illustrated in Figure 2.17c. It resembles yeast SWR1 except for the potential absence of Swc7. However, the presented composition might be incomplete as there could be additional subunits being part of the species-specific module. Therefore, pulldown experiments from *C. t.* were performed as presented in Section 2.3.5.

2.3.2 Bioinformatic tools identify a candidate for *C. t.* Swc7

As described in the previous section, I could not identify a homologue for yeast Swc7 in *C. t.* using a sequence based search approach. Nonetheless, a Swc7-like protein with a conserved fold could exist in *C. t.* comprising a different amino acid sequence. I therefore used the Dali [288] tool which searches in protein databases for structurally homologue proteins. The Dali search identified a structrual similarity of Swc7 and human Brd8 which is a subunit of the human NuA4/Tip60 complex [246]. This human super-complex shares subunits with the yeast homologues of the NuA4 and SWR1 complexes. The AlphaFold model of human Brd8 (Figure 2.19d) showed that the first 130 amino acids have a similar fold compared to yeast Swc7 (Figure 2.19c). Overall, Brd8 incorporated the Swc7-fold while Swc7 remained a separate protein in yeast (Figure 2.19a). Both Brd8 and Swc7 harbor three smaller helices and a long one. While the smaller helices are intercepted with small helical features in Swc7, Brd8 comprises a more classical helix-turn-helix architecture. Despite the high degree in structural similarity, yeast Swc7 and human Brd8 do not share a common amino acid sequence (Fig-

ure 2.19b) and only show little similarity. I conclude that evolutionary Swc7 evolved in its amino acid content while there was a pressure on keeping the 3D structure. Human Brd8 contains besides the Swc7-fold two Bromo domains which bind acetylated H3/H4 tails [211]. It shares this feature with yeast Bdf1 that is an integral part of yeast SWR1 (Figure 2.19d). Yeast Bdf1 contains a NET domain in comparison to Brd8. The long loop connecting the NET domain helices folds in the presence of Swr1 into a β-sheet alongside the Swr1 residues 243-247 which remain unstructured in absence of Bdf1 (Figure 2.19e, based on AlphaFold Multimer models [273]). The NET domain therefore tethers Bdf1 to the SWR1 complex while the bromo domains are connected by an intrinsically disordered region and are hence ready to bind acetylated H4. As the NET domain is missing in Brd8, I speculate that the Swc7-fold functions in tethering Brd8 to the remaining NuA4/Tip60 complex.

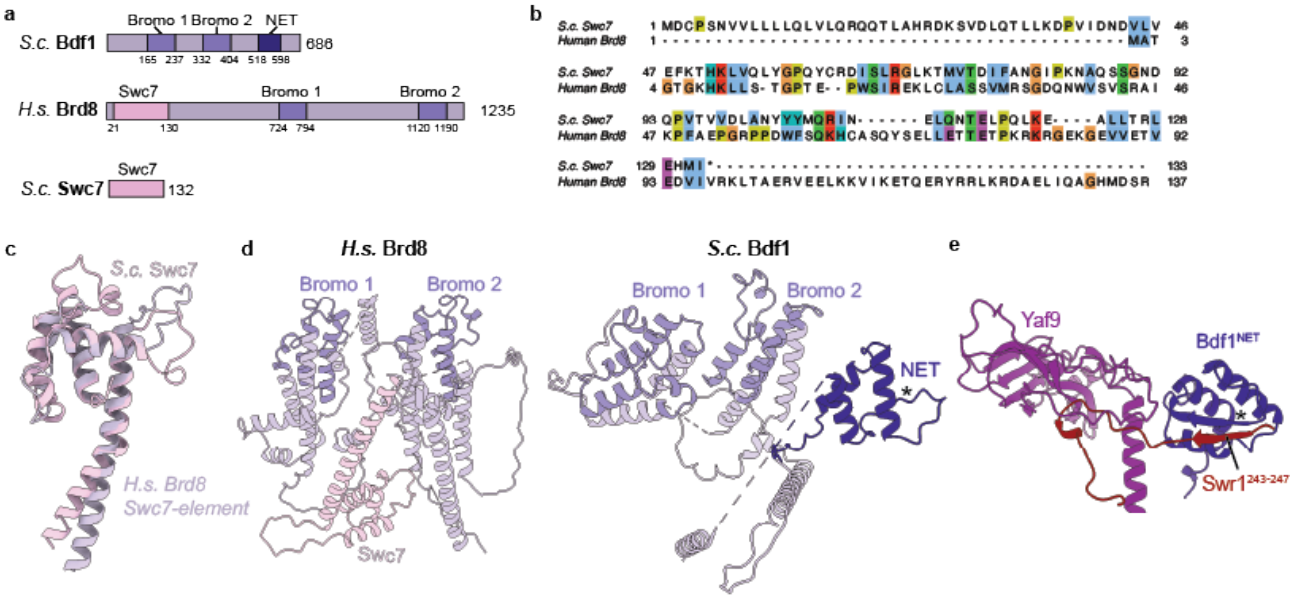


Figure 2.19: **Swc7-element in human Brd8. a** Domain overview of yeast Bdf1, human Brd8 and yeast Swc7. Numbers indicate the range of amino acids. **b** ClustalO alignment [287] created in Jalview of yeast Swc7 and human Brd8. There is some similarity with however little sequence consensus. **c** Comparison of the AlphaFold predictions of yeast Swc7 (pink) and the Swc7-element of Brd8 (light purple) **d** (*left*) AlphaFold prediction of human Brd8 showing its two Bromo domains and the Swc7-element (*right*) Same representation of yeast Bdf1 which does not contain the Swc7-element but a NET domain which **e** is a putative anchor of Bdf1 to bind the N-terminal region of Swr1 (based on AlphaFold multimer prediction)

The finding of the Swc7-element in Brd8, provided me the opportunity to use its structure and amino acid sequence to look for Swc7 in *C. t.* For this purpose I turned to FoldSeek which uses similar to Dali, algorithms to search for structure similarities from databases like the pdb and AlphaFold but in a much faster fashion [289]. I used the Brd8-Swc7 element that corresponds to the first 130 amino acids of human Brd8 to search for a similarly folded protein in *Neurospora crassa* (*N. c.*) I had to look first for Swc7-like proteins in *N. c.* as Foldseek does not allow to search for *C. t.* proteins directly. Indeed Foldseek returned one hit which I used for a sequence Blast® search that identified the uniprot entry G0SFI0 of *C. t.* Next, I compared the predicted structures and amino acid sequences of yeast Swc7, the Brd8-Swc7 element and the hits from *C. t.* and *N. c.* (Figure 2.20a). Strikingly, the hits from *C. t.* and *N. c.* resemble the structure of the Swc7 element. However, the helix-turn-helix motif of the Swc7 element shows altered features in *C. t.* and *N. c.* While there are three short helices and a

subsequent long helix in the human Brd8 and the yeast Swc7, the hits for *C. t.* and *N. c.* are composed of a helix-turn-helix and a third helix that continues as a long helix. The relative position of the long helix thus differs between Brd8/Swc7 and the fungi variants. The described helix-turn-helix motif is denoted as Myb domain in the corresponding uniprot entries for the *C. t.* and *N. c.* hits. The Myb domain is a common fold in some transcription factors found in plants [290].

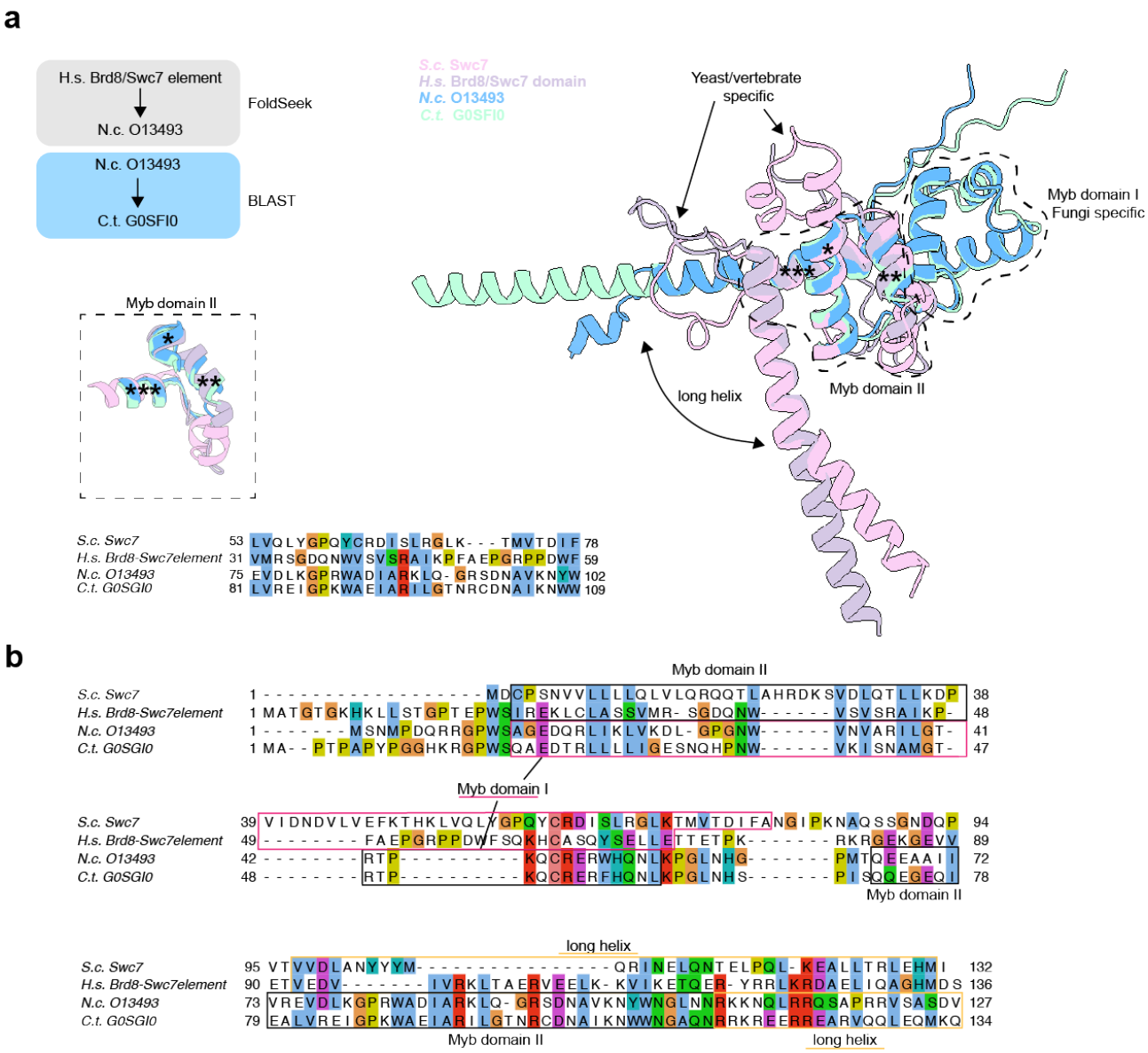


Figure 2.20: **Identification of Swc7 like protein in *Chaetomium thermophilum*. a** (top) Approach. (*right*) Aligned AlphaFold models for yeast Swc7 (pink), human Brd8 (light purple) and the candidates for *C. t.* (green) and *N. c.* (blue) Swc7. Structures aligned on Myb domain II of the *C. t.* and *N. c.* candidates. Myb domain I is specific to the fungal proteins and sits at their N-terminus. The three helixes of the Myb domain II are marked with asterisks. (*left*) Sequence alignment of Myb domain II. The structure shows only the residues that align in terms of amino acid sequence. The Myb domain II of *C. t.* and *N. c.* resembles Brd8 at their N-terminus while they are more similar to yeast Swc7 at the C-terminus. **b** Sequence alignments of yeast Swc7, human Brd8-Swc7 element and the *C. t.* and *N. c.* Swc7 candidates.

At the N-terminus of the Swc7-like *C. t.* and *N. c.* protein sits a second Myb domain in which the third helix remains to be short. I denoted the Myb domain that is specific to fungi as *Myb domain I* while the shared Myb domain is called *Myb domain II*. Alignment of the amino acid sequences of the Myb domain II shows that it consists mainly of non-charged residues in all four proteins. In addition, the *C. t.* and *N. c.* Swc7-like proteins resemble Brd8 at the N-terminus of the motif, while they are more similar to yeast Swc7 at the C-terminus. When I aligned the full-length proteins, some conserved motifs were visible showing that the *C. t.* and *N. c.* Swc7-like proteins share more of their sequence with Brd8 (Figure 2.19b). However, the alignment tool ClustalO [287] aligned the Myb domain I with Myb domain II of Brd8 and Swc7 and parts of the fungal Myb domain II with the long helix of Brd8 and Swc7. For this reason and as this is only a bioinformatical approach, an experimental validation is required to confirm that the *C. t.* Swc7 candidate is an integral part of *C. t.* SWR1. For this purpose I used a pulldown approach presented in Section 2.3.5.

2.3.3 Recombinant protein production of *C. t.* SWR1

The sequence and structure similarity searches were not sufficient to identify a Swc7 homolog in *C. t.* with high confidence. Nonetheless, the identified Core and Arp module subunits are sufficient to study the SWR1 interaction with nucleosomes by biochemical and structural methods. In addition, it was shown by *in vitro* histone exchange assays that only Swr1, Swc2, Swc5, Swc6, Arp4, Arp6 and Yaf9 are essential [228][224] – all of which are present in my recombinant SWR1 construct. For these reasons, I decided to use a recombinant protein production strategy to obtain high protein yields and be able to perform site-directed mutagenesis studies for understanding structure-function relations after obtaining structural insights into SWR1. To clone recombinant *C. t.* SWR1 I applied a similar strategy as described for *C. t.* INO80 [77] were the subunits are assembled on Multibac vectors [291] followed by Cre recombinations for baculovirus-mediated insect cell expression.
I cloned the 13 subunits into two plasmids as the total length of all gene expression cassettes would have been too large for efficient insertion into the baculovirus genome. For this purpose I cloned the gene expression cassettes coding for subunits from the same module on one plasmid. This aimed to facilitate the assembly of stable modules during recombinant expression. In more detail, I cloned each subunit in either so-called 'doner' or 'acceptor' Multibac vector which were previously modified to contain a polyhedrin and SV40 terminator designed for insect cell protein expression [291]. Pairs of two subunits were created using InFusion reactions (Takara Bio US). Finally, the plasmids were derived by combing the acceptor vector with the donor vectors in a stepwise manner (Figure 2.21a). As each individual vector contains a unique antibiotic resistance, clones can be selected that contain all desired gene expression cassettes. For quality control, I sequenced the two-subunit pairs and performed test PCRs of the Cre recombined plasmids to check for mutations and that all subunits were correctly assembled. Subsequently, baculoviruses were generated likewise to INO80. The Bacmid DNA prepared for the virus production was carefully analysed by PCRs (Figure 2.21) using primers that generated products of about 250 bp. This allowed me to qualitatively verify that the Bacmid DNA contained all subunits encoding gene cassettes for stoichiometric expression of the protein complex.
For pilot expressions and purifications of recombinant *C. t.*, I used the INO80 purification protocol described in Section 2.1.1 and adapted it based on [221] and [142]. The protocol mainly differed from the one for INO80 by the adjusted pH value and the usage of zinc chloride as Swc6 is a reported zinc binding protein based on conserved zinc binding residues in its sequence. A C-terminal double FLAG tag on the Swr1 subunit allows to perform an anti-FLAG affinity step followed by anion exchange chromatography to separate DNA bound complexes. FLAG-tag affinity purification from 25 ml insect cell expression culture showed a similar band pattern compared to its yeast SWR1

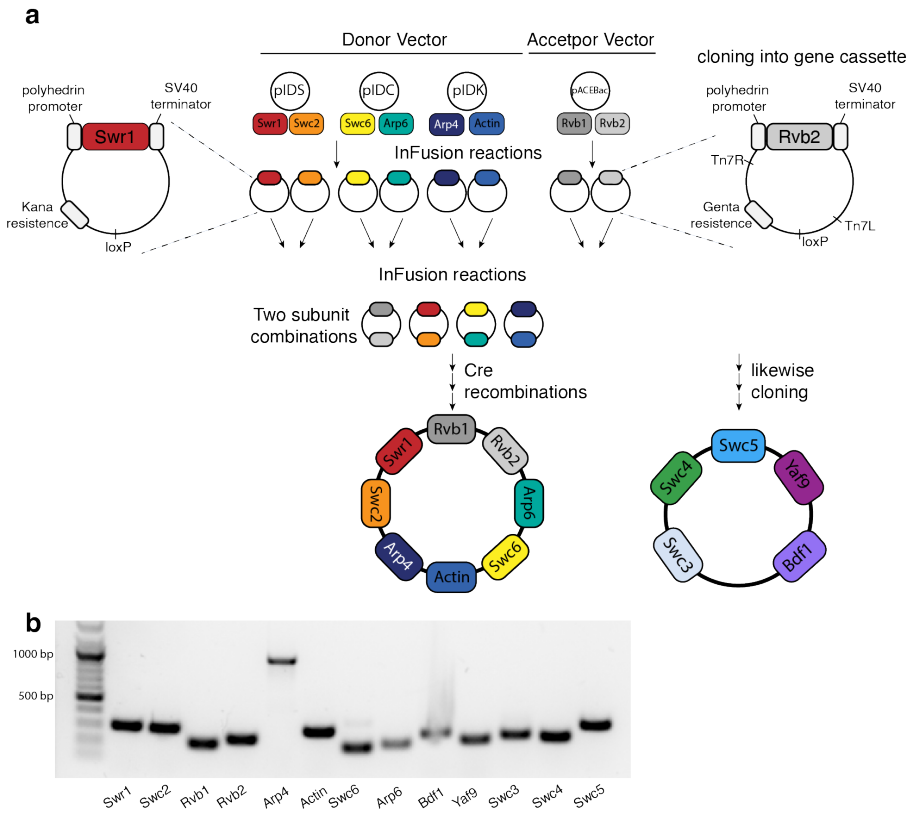


Figure 2.21: **Cloning of *C. t.* SWR1 complex. a** Overview of cloning procedure and division of subuntis across the two plasmids. Close-up view for an example donor and acceptor vector on the left and right respectively. The loxP site enables the Cre recombinations. The acceptor vector contains an additional Tn7R and Tn7L site to allow for the integration of the viral genome upon baculovirus production. **b** Agarose gel showing presence of all 13 subunits in the Bacmid DNA used for baculovirus generation.

counterpart [142] and I therefore scaled the purification up to 2 l which became the regular volume for SWR1 purifications aimed at cryoEM structural studies. An example purification is depicted in Figure 2.22a,b. The anti-FLAG purification proved to be sufficient to remove most of the insect cell proteins. In the subsequent ion exchange purification step, I typically observed one peak at high protein concentrations and a second representing SWR1 bound to DNA and histones as also visible on SDS-PAGE. The DNA-bound peak could not be reduced by adding DNAse in the lysate step nor by varying the sonication conditions. The identity of the bands on the SDS-PAGE was verified using in-gel digestion and subsequent mass spectrometry analysis which had been performed by the EMBL proteomics core facility. In addition, the presence of all 13 subunits and their stoichiometry was also analyzed by in-solution proteolytic digest followed by mass spectrometry. Both the SDS-PAGE and mass spectrometry analysis revealed that the Bdf1 subunit was only present substoichiometrically. However, I observed in some of the purifications that Bdf1 was still present after anti-FLAG affinity purification (Figure 2.22c) while it was not stably bound during the ion exchange step. This indicated that the association of Bdf1 to the SWR1 complex might be sensitive to ionic strength. Usage of potassium chloride instead of sodium chloride had a positive yet unreproducible effect on the presence of Bdf1. The unstable interaction of Bdf1 with the *C. t.* SWR1 complex can be explained by the non-identified Swc7 subunit (see Section 2.3.1 and 2.3.2). It was shown in yeast that the recruitment of Bdf1 to the SWR1 complex is dependent on the presence of Swc7 which I could

not identify for *C. t.* SWR1 [228]. On the other hand, my pulldown experiments (Section 2.3.5) showed that recombinant SWR1 can stably associate with endogenouss Bdf1 possibly with another yet unidentified Swc7 orthologue from the lysate. This indicates that Bdf1 is an integral part of the *C. t.* SWR1 complex. SWR1 was proposed to contain two copies of actin, residing in its Arp module [221]. Unlike INO80, SWR1 does not contain Arp8 and a second actin copy could take its place.

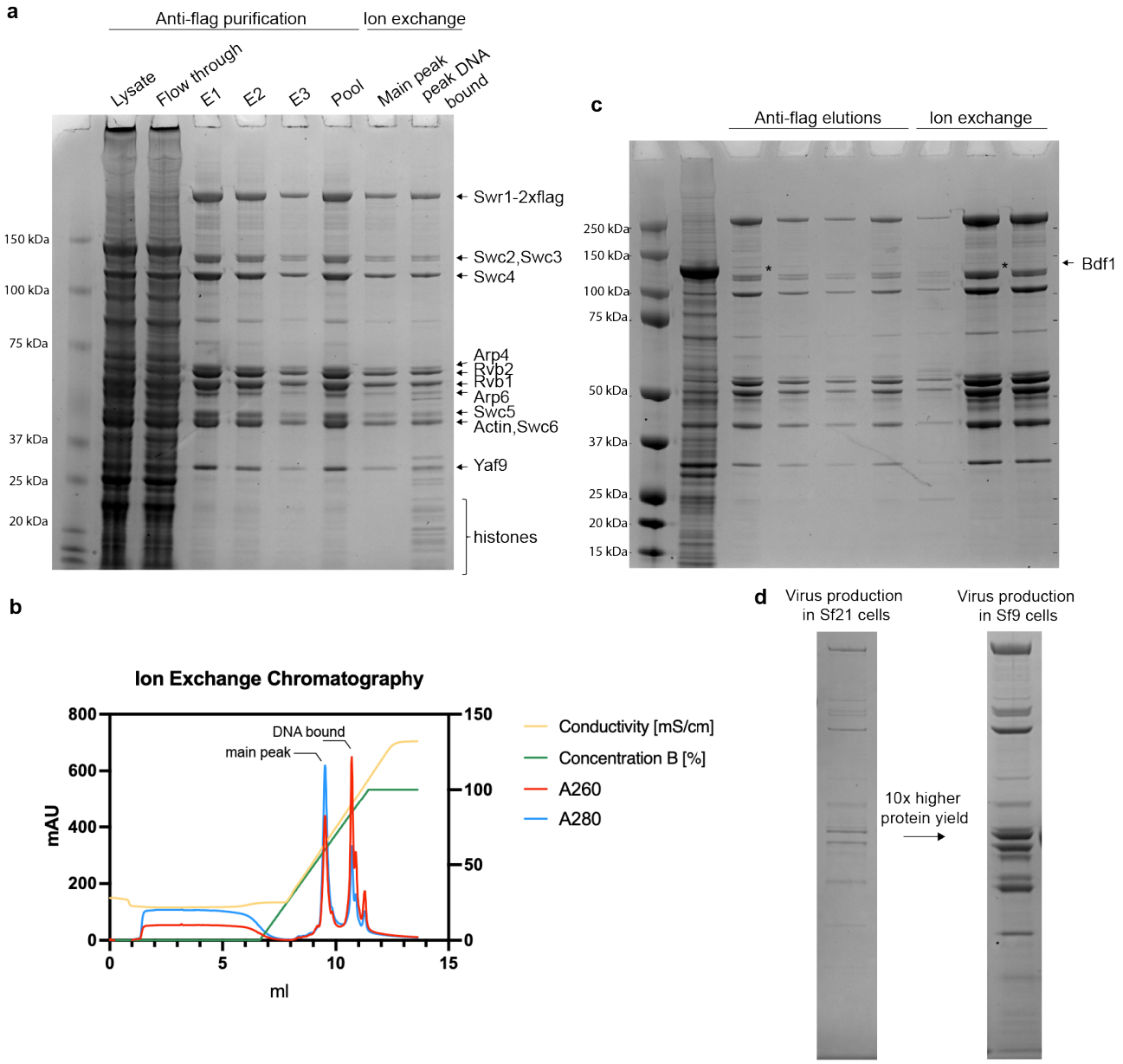


Figure 2.22: **Protein purifications of *C. t.* SWR1. a** SDS PAGE showing lysate, flow through and elutions of the anti-FLAG purification as well as fractions for the ion exchange chromatography step. Subunits are labeled according to mass spectrometry analysis **b** Chromatogram of ion exchange step showing the absorbance at 280 nm for protein (blue line), at 260 nm indicating presence of DNA (red line), the percentage of buffer B (green line) and the measured conductivity (yellow line). **c** Example gel where Bdf1 was present faintly in the FLAG elutions and not any more after the ion exchange chromatography. **d** SDS-PAGEs illustrating the increased amount of SWR1 complex after changing the baculovirus production from Sf21 to Sf9 cells. Comparable amounts were loaded on the gel.

The published crystal structure of the truncated SWR1-HSA helix bound to one actin and Arp4 is not sufficient to address this question and further biochemical and structural studies are thus required. I performed additional biophysical experiments presented in Section 2.3.4 as it remained challenging to estimate the actin stoichiometry of the recombinantly expressed *C. t.* SWR1 complex by Commassie stained SDS-PAGE or by mass spectrometry analysis.
Lastly, I want to comment on the protein yields obtained. Initially I produced the baculoviruses in Sf21 insect cells while I performed the protein expression in Hi5 cells. Using this approach, I obtained only about 100 µl of protein sample at a concentration of 800 nM. This prevented the cryoEM sample preparation of the SWR1-nucleosome complex as I had to assemble the complex at a minimum of 600 nM to reach sufficient coverage of the cryoEM grids. Adapting the baculovirus production from Sf21 to Sf9 cells helped to increase the protein yields ten fold based on the calculated obtained amount of protein. This optimized protocol also worked to reach a higher protein yield for the INO80 complex. Therefore, Sf9 cells are now generally used in the lab for baculovirus production.

2.3.4 Biophysical characterization of *C. t.* SWR1

I used different biophysical methods to further characterize the mega-Dalton sized recombinant *C. t.* SWR1 complex. Mass spectrometry and SDS-PAGE analysis indicated that the subunit Bdf1 was present in sub-stoichiometric amounts as mentioned in the previous section. The other 12 subunits of *C. t.* SWR1 were robustly detected. To analyse the stoichiometry and homogeneity of the sample, I utilized mass photometry – a single molecule technique that allows accurate determination of the molecular weight of protein(/nucleotide) samples [292]. The mass photometry analysis of *C. t.* SWR1 shows at least four different species (see Figure 2.23a).
The main peak at 1067 kDa is consistent with a SWR1 complex that carries two actin molecules and lacks Bdf1, as already shown on SDS-PAGE and by mass spectrometry. The calculated molecular weight of this complex would be 1043 kDa which is close to the observed mass and within the error range of +/- 20 kDa [292]. A second peak at 930 kDa represents a sub-population that is 113 kDa smaller. The closest to this value is Swc3 with 83.4 kDa. Thirdly, a sub-fraction at 362 kDa is detected possibly consisting of the Rvb1/Rvb2 hetero-hexamer which has a calculated mass of 310.5 kDa. In summary, the majority of the sample contains the SWR1 with two actin copies.
Next, I performed crosslinking mass spectrometry together with Mandy Rettel from the EMBL proteomics core facility. Crosslinking of protein complexes allows to further probe interacting subunits within the SWR1 complex. In addition, the obtained crosslinked peptide-pairs can be used for an integrated structural characterization of (low resolution) cryoEM densities [294]. Prior to crosslinking, I added recombinant *C. t.* Bdf1 as well as *C. t.* H2A.Z/H2B dimers (further described in Section 2.6) to probe if Bdf1 interacts with the remaining subunits without Swc7 and to identify the histone-interacting subunits. Crosslinking was performed in the presence of 0.25 mM bis(sulfosuccinimidyl)suberate (BS3) which contains two NHS ester groups and crosslinkes primary amines as for example present within lysine residues. Accordingly, the crosslinking reaction can be quenched by adding an excess of Tris-HCl that contains a primary amine group. I performed the crosslinking reactions (Figure 2.23b), while Mandy Rettel conducted both the trypsin digest as well as the mass spectrometry experiments. Due to the complexity of the sample it was important to apply a 120 min gradient allowing the detection of more peptides. I depicted the crosslinks using the xiView tool [293] from the Rappsilber lab (Figure 2.23c). The crosslinking experiment shows the modular architecture of SWR1. The core subunits Rvb1, Rvb2, the Swr1 motor and insert regions as well as Swc2, Arp6 and Swc6 interact as already reported by prior cryoEM [142]. Additionally, crosslinks between Swc2 and Swc3 are observed indicating their interaction. The reported Arp module subunits actin, Arp4 and Swc4 show crosslinks however none to Swr1 as it would be expected

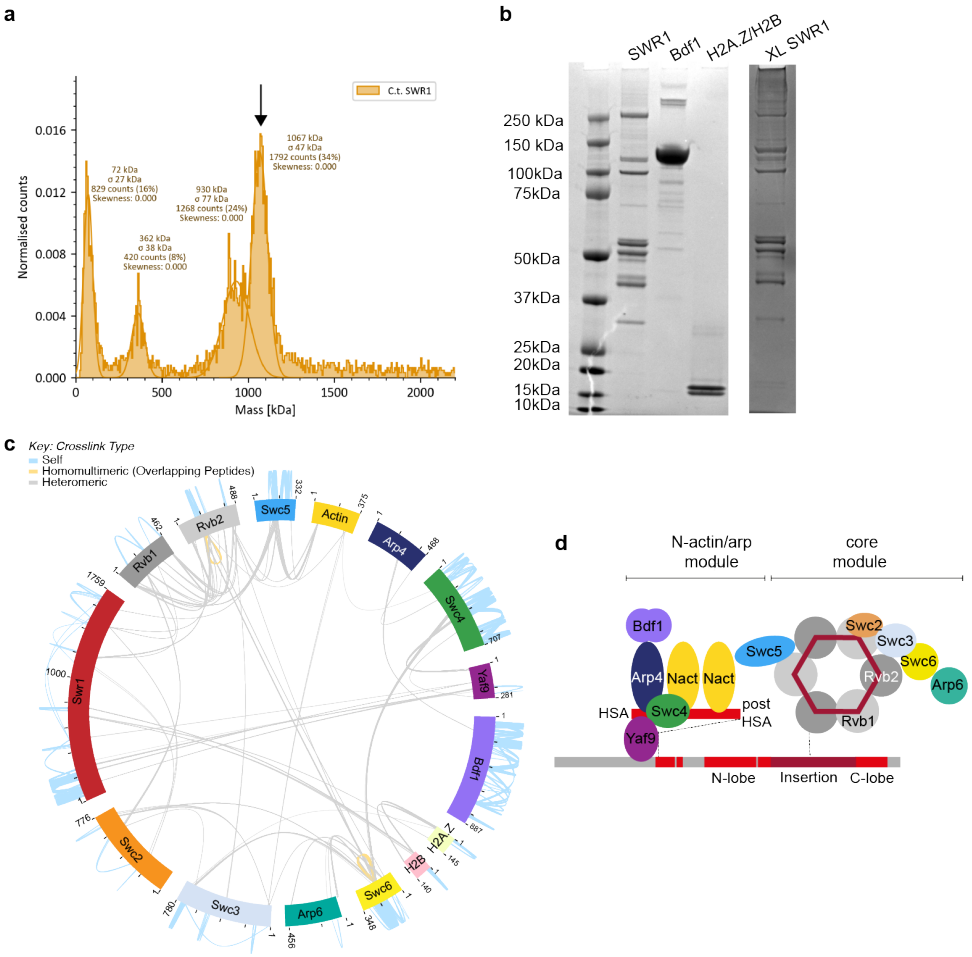


Figure 2.23: **Biophysical characterization of *C. t.* SWR1. a** Mass photometry of *C. t.* SWR1. Shown are the normalised counts. The protein mass was calibrated using IgG and Tg. The first peak originates from the buffer. The fourth peak represents *C. t.* SWR1 in absence of Bdf1 and presence of two actin molecules. **b** SDS-PAGE of input for crosslinking (*left*) and the final crosslinked sample in presence of 0.25 mM BS3 (*right*). **c** Result of crosslinking experiment illustrated with xiView [293]. **d** SWR1 architecture based on the presented crosslinking data.

based on the ternary Arp4/actin/Swr1-HSA crystal structure [225]. Yaf9 crosslinks to both Swc4 and the N-terminal region of Swr1. Bdf1 that was added to the SWR1 complex, shows only an interaction with Arp4. Swc5 is linked to Rvb1, Rvb2 and actin, placing it in between the Core and Arp modules. Besides the internal SWR1 crosslinks, interactions between Swr1 and both H2A.Z and H2B were detected. This concerns the Swr1 residues around 680-760 which lay between the HSA helix and the N-terminal motor lobe. Yeast Swr1 was shown to bind yeast H2A.Z/H2B dimers [234] with an overall unstructured region that corresponds to residues 770-795 in *C. t.* Swr1. This region lays adjacent to the crosslinked region. Overall, the observed binding region of H2A.Z/H2B to *C. t.* SWR1 seems to be in agreement with the literature. Besides Swr1, Swc2 interacts with H2A.Z/H2B based on the literature [233] - a crosslink between Swc2 and the histone dimer is however absent in the presented crosslinking data. This led me to the SWR1 architecture depicted in (Figure 2.23d). Swc2 and Swc3 assemble onto the Rvb1/Rvb2 hetero-hexamer in proximity to the Arp6/Swc6 arm explaining crosslinks between Swc3 and Swc6. Within the Arp module, Yaf9 is an integral part indicated by the number of crosslinks between Swc4 and the N-terminus of Swr1. Swc5 is located

at the intersection of the two modules as also shown in [228]. Lin et al. deleted different ranges of the Swr1 subunit and identified the lost subunits [221]. Interestingly, the Arp module subunit Swc4 crosslinks to Swc6 of the Core module indicating that the Arp module might fold back to the core of SWR1. Bdf1 likely associates with the Arp module of SWR1 but the crosslinking data is too sparse for a confident placement of Bdf1. In summary, the performed crosslinking of *C. t.* SWR1 is a valuable addition to the previous MS-XL experiments of yeast SWR1 [222] that yielded a much more sparse dataset. In their case, Yaf9 was only crosslinked to Swc5 and links between both Swc2-Swc3 and Arp6-Swc6 were absent. In the following section, I will use the presented data to validate the structural findings.

2.3.5 Pulldowns with *Chaetomium therompilum* to identify additional SWR1 subunits

I performed pulldown mass spectrometry proteomics to identify a *C.t.* orthologue of Swc7 and to investigate whether *C. t.* SWR1 contains any additional, species-specific subunits. Pulldowns of chromatin factors are especially challenging, as they might not end up in the soluble fraction that is accessibly for mass spectrometry analysis. In addition, the occupancy of chromatin factors can be low and cell state dependent. I used the SWR1 subunit Bdf1 and recombinant SWR1 as bait proteins. The latter was produced as elaborated in Section 2.3.3. The choice fell on Bdf1 as it was described in the literature, that the assembly of Bdf1 into the SWR1 complex is dependent on the presence of Swc7 implying a direct interaction of the two [228]. Also, Swc3 and Swc5 were tested as bait proteins for the pulldown as both of them are specific to the SWR1 complex. However, the C-terminus of Swc3 was not accessible for its purification and Swc5 turned out not to be specific for SWR1 (see Section 2.3.6). For an initial trial pulldown, I purified SWR1, Bdf1 and eGFP as a negative control in triplicates from 0.5 g of insect cell pellets using anti-FLAG affinity beads (see procedure in Figure 2.24a). Subsequently, beads were incubated with lysate from *Chaetomium thermophilum* and subjected to extensive washing. SDS eluted samples were submitted to the proteomics core facility for tandem mass tag labelling (TMT) and subsequent mass spectrometry for quantitative analysis comparing the different bait proteins directly. In an optimized set-up, I used the to homogeneity purified bait proteins for immobilization on the beads.
In the initial pulldown experiment using SWR1 as bait, the SWR1 complex pulled down the canonical histones as well as H2A.Z (Figure2.24b). Interestingly, H2A.Z is among the enriched hits while H2A was detected with lower abundance indicating a higher affinity of SWR1 for H2A.Z compared to H2A. This finding is in line with previous reports as SWR1 can bind H2A.Z via its Swc2 and Swr1 subunits [233][234]. Of note is, that Bdf1 is present in the SWR1 pulldown, even though it was not part of the SWR1 bait as it is lost in the used purification strategy (see Section 2.3.3). However, it is an integral part of the SWR1 complex according to my pulldown result. While the presence of the canonical histones indicates the detection of nuclear components, most of the enriched hits represent proteins of the translation machinery. Arp9 and Taf14 were among the enriched hits, even though they are part of the SWI/SNF and INO80 complexes respectively. In summary, the data does not allow to conclude if any additional species-specific subunits were found. Interesting hits were further analyzed whether they have a Swc7-like fold, which also was not the case. To improve the data quality, I optimized the pulldown protocol by using pre-purified eGFP and SWR1. I purified SWR1 and eGFP to homogeneity and added purified proteins to the anti-FLAG beads in equimolar amounts. In these experiments, I was however not able to fully saturate the beads, due to the limited amounts of purified, recombinant SWR1.

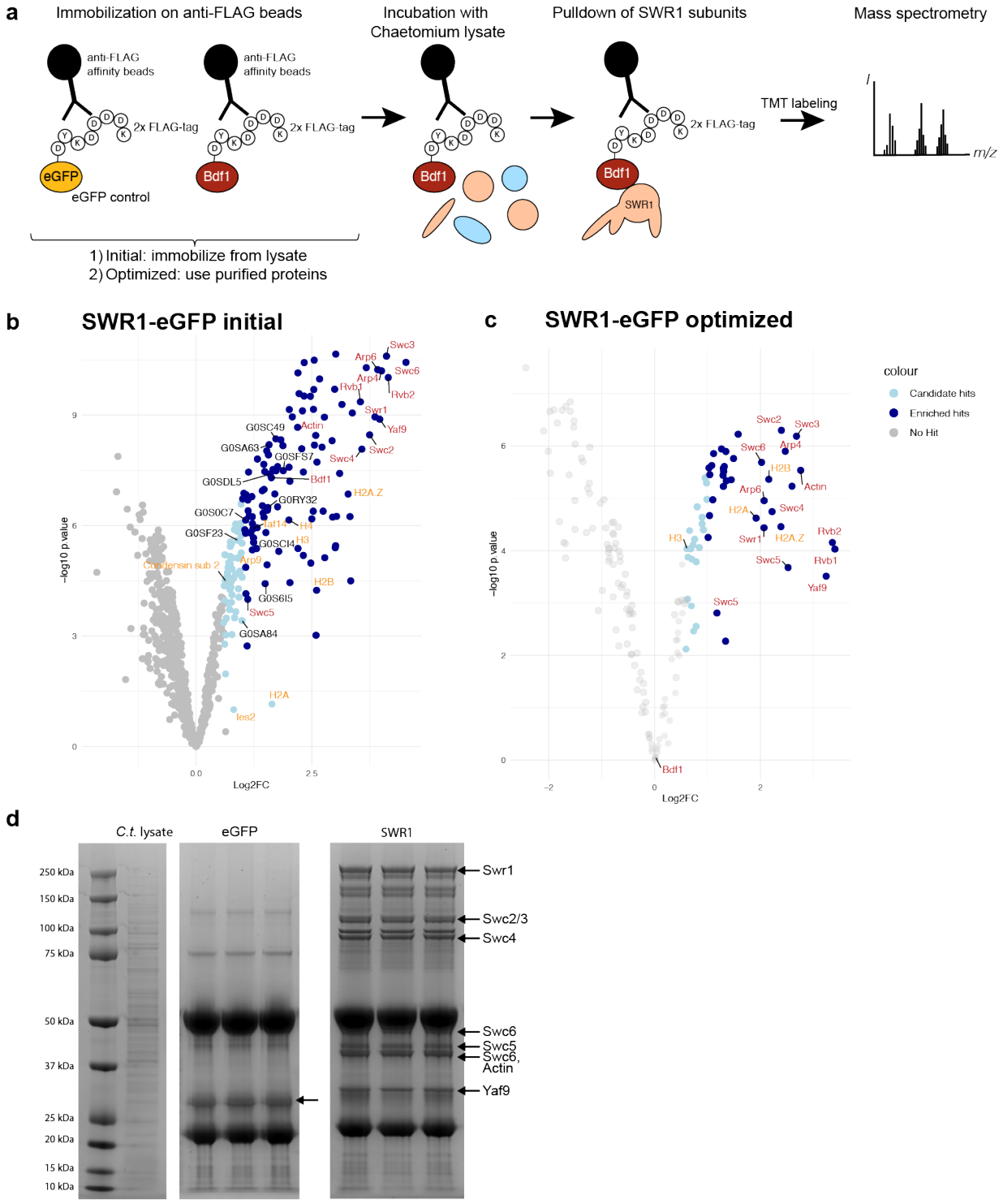


Figure 2.24: ***Chaetomium thermophilum* pulldowns using SWR1 as bait protein.** **a** Experimental procedure. Bait proteins were either purified on the fly on the beads prior to incubation with Chaetomium lysate or recombinantly purified bait proteins were immobilized on the beads. **b** Vulcano plots of SWR1 pulldowns. SWR1 subunits are labeled in red, while histone and other interesting hits are labeled in orange. Hits that were followed up by comparing their AlphaFold predictions with Swc7 are labeled with their respective uniprot identifiers. R scripts for generation of Vulcano plots were kindly provided by Thomas Dahlet. **c** Optimized SWR1 pulldown showing overall less proteins. Bdf1 is no longer pulled down efficiently by the remaining SWR1 complex. **d** Control SDS-PAGEs showing Chaetomium lysate input and boiled beads of the final samples. Arrow indicate were bait proteins run. The two thickest bands represent the light and heavy chains of the anti-FLAG antibodies respectively.

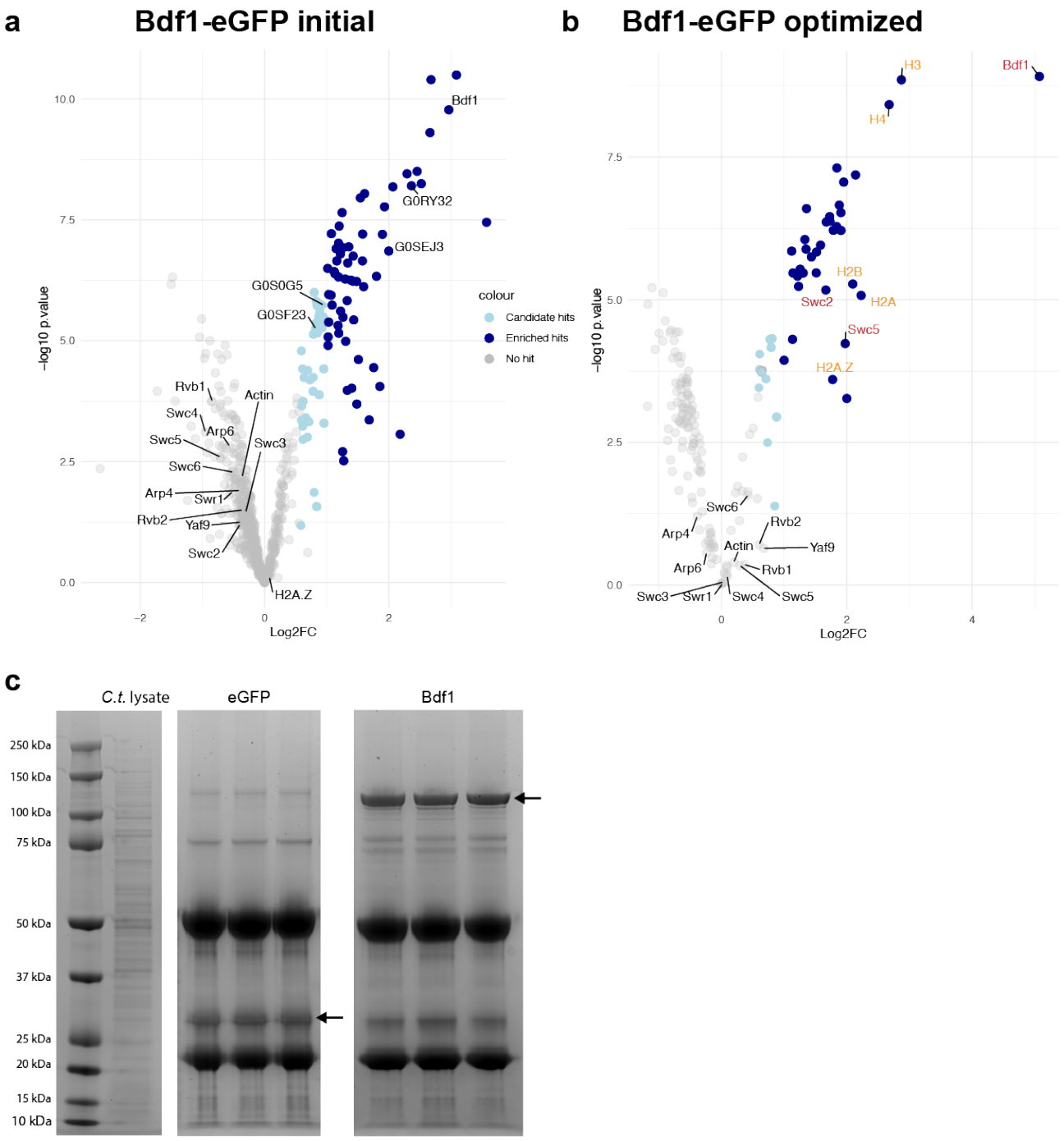


Figure 2.25: **Pulldowns using Bdf1 as bait protein. a** Vulcano plot of initial Bdf1 pulldown. In the presented initial pulldown, Bdf1 did not pull down any of the SWR1 subunits. **b** In the optimized condition, Bdf1 pulled down Swc2, Swc5 as well as some histone proteins (in orange). **c** Control SDS-PAGEs showing Chaetomium lysate input and boiled beads of the final samples. Arrows indicate were bait proteins run.

Generally less proteins were detected with the optimized setup (Figure 2.24c, d) indicating that the beads were more saturated when immobilizing SWR1 from lysate. Also, Bdf1 was not among the enriched hits. Again, I could not identify any additional species-specific subunits nor the Swc7 candiate G0SFI0 described in Section 2.3.2.
In case of the intial Bdf1 pulldown, none of the other SWR1 components were detected as hits but

were even less abundant compared to the eGFP control (Figure2.25a). This indicated a problem in the relative amounts of Bdf1 protein in comparison to the eGFP control. However, the optimized conditions using pre-purified Bdf1 improved the analysis (Figure 2.25). Bdf1 pulled down the canonical histones as well as H2A.Z.The H3 and H4 histones are of almost similar abundance compared to Bdf1 which is in line with previous reports that Bdf1 binds to acetylated H3/H4 [211]. In this experiment, the SWR1 subunits Swc2 and Swc5 were detected as hits. Of note is that, those two SWR1 subunits are histone binding proteins: While Swc2 binds the incoming H2A.Z/H2B dimer [233], Swc5 binds the canonical H2A/H2B dimer which had just been ejected by the SWR1 complex. This puts them in close contact to the nucleosome where also Bdf1 binds via its two bromo domains. In summary, my pulldown experiments were not sufficient to identify any species-specific subunits or the Swc7 candidate. A solution to this could be the usage of more bait protein or to use crosslinking conditions to capture weak interactions. Nevertheless, the identified 13 *C. t.* SWR1 subunits are sufficient to study the structure and function of the Core and Arp modules. Further, the pulldown experiments provided no evidence for a Tip60-like super complex in *C. t.*

2.3.6 Swc5 is a putatively shared subunit between the SWR1 and INO80 complexes

Chaetomium thermophilum pulldowns were used to identify the repertoire of SWR1 subunits in *C. t.* For this purpose I looked for SWR1-specific subunits that can be used as bait proteins to identify additional SWR1 subunits especially in its species-specific module. Swc5 was a good candidate as it is a subunit unique to SWR1 and is located somewhere between the Core and Arp modules of SWR1. The experiment was carried out likewise to the pulldowns described in Section 2.3.5.
The quantitative mass spectrometry result (Figure2.26a) displays proteins that are enriched in comparison to the eGFP control. Among these enriched hits, none of the 12 expected SWR1 subunits were detected. Based on published crosslinking data on SWR1 [222], it was expected that at least the subunits Swr1, Yaf9 and actin are pulled down. Apparently, these interactions are not strong enough to enable Swc5 to pull down these expected subunits and the remaining SWR1 subunits from the cell lysate. Alternatively, the SWR1 subunits might not be present in the soluble part of the lysate. Instead of the expected SWR1 subunits, Swc5 pulled down a large proportion of the INO80 subunits, namely: Iec3, Taf14, Arp8, Arp5, Ies4, Ies1, HMG, Ies2, FHA and ZnF. Of note is that the subunits shared between INO80 and SWR1 – Rvb1, Rvb2, Arp4 and actin – were not detected as hits as they were of too low abundance and/or not enriched compared to the eGFP negative control. The presented result, indicates that Swc5 is a shared subunit between SWR1 and INO80. Another pulldown was performed likewise to the optimized conditions described in Section 2.3.5. This time, I detected Bdf1 as an interaction partner of Swc5 which is in line with the Bdf1 pulldown as Bdf1 is vice versa capable of pulling down Swc5. However, no other SWR1 peptides were detected by any means indicating that I had used less Swc5 bait protein compared to the first approach. Also less INO80 subunits are among the hits, yet this time Arp4, Rvb1 and Rvb2 were detected as hits. Overall, the repetition of the pulldown confirms what I had seen before: Swc5 can interact with INO80. On the side, I want to point out that Importin and Importin α were detected as enriched hits indicating that Swc5 and potentially associated chaperons associate with Importin to enable shuttling to the cell nucleus. This is in accordance with [295] that identified also Ino80 itself to be targeted by Importin.

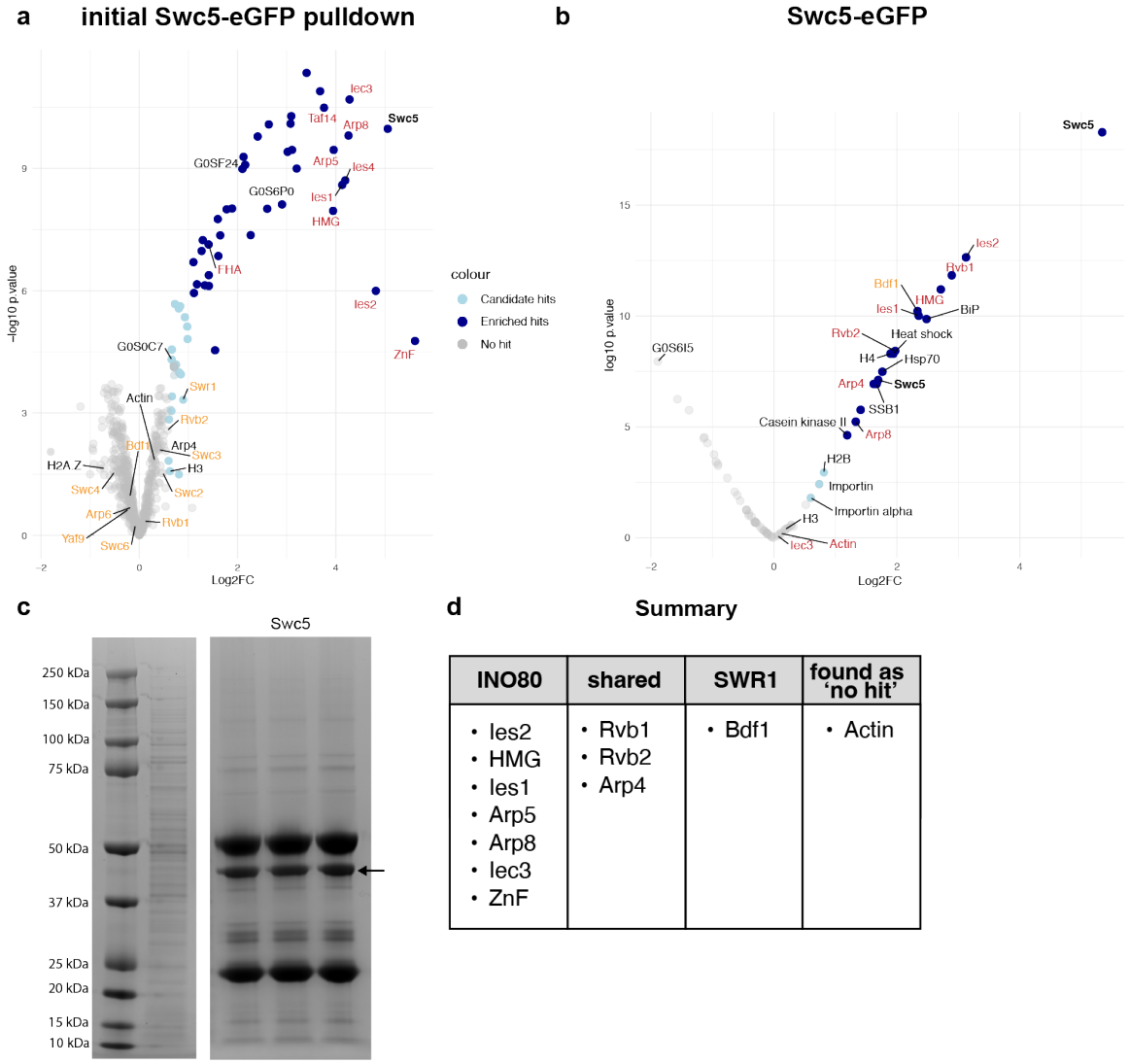


Figure 2.26: ***Chaetomium thermophilum* pulldown with Swc5 as bait protein. a** Vulcano plot of initial Swc5-eGFP pulldown and **b** optimized conditions. INO80 subunits are labeled in red while SWR1 subunits are in orange. **c** SDS-PAGE with lysate input and eluted triplicates of Swc5 pulldown. Arrow indicates where Swc5 runs. **d** Summary showing which SWR1 and INO80 subunits were detected.

In order to probe the Swc5 interaction with INO80 *in vitro*, I performed an MST experiment together with Karine Lapouge from the EMBL protein expression and purification core facility. For this purpose, I fluorescently labeled a subset of lysines of Swc5. Thereby, one can measure protein-protein interactions and determine the K_D of the underlying interaction. When I titrated increasing concentrations of INO80 to Swc5, I started to see a change in fluorescence indicating binding of INO80 and Swc5 (data not shown). However, due to the low concentrations of recombinant INO80 in comparison to Swc5, I was not able to reach the saturation point which is required for the determination of the K_D. Instead, I performed crosslinking MS experiments using INO80 and Swc5 to identify the subunits involved in the interaction. For this purpose, the evolutionary-conserved 11-subunit INO80 complex was used that contains Ino80$^{718\text{-}1848}$, Arp5, Arp8, Ies2, Arp4, Rvb1, Rvb2, Ies4, actin, Taf14, Ies6. Notably, Swc5 migrates similarly as Ies4 and is therefore not visible as a separate band on the gel, as shown in Figure 2.27a. To optimize crosslinking conditions, I tested different BS3 crosslinker

concentrations (Figure 2.27b). At 0.25 mM BS3 minor bands of non-crosslinked INO80 subunits were still observed, while the majority of protein was crosslinked. Consequently, this condition was used to proceed with the MS anaylsis. To this end, a larger version of the recombinant *C. t.* INO80 complex was utilized that contains five additional (species-specific: FHA, Ies1, Iec3, ZnF and HMG) subunits and was kindly provided by the Hopfner lab in Munich.

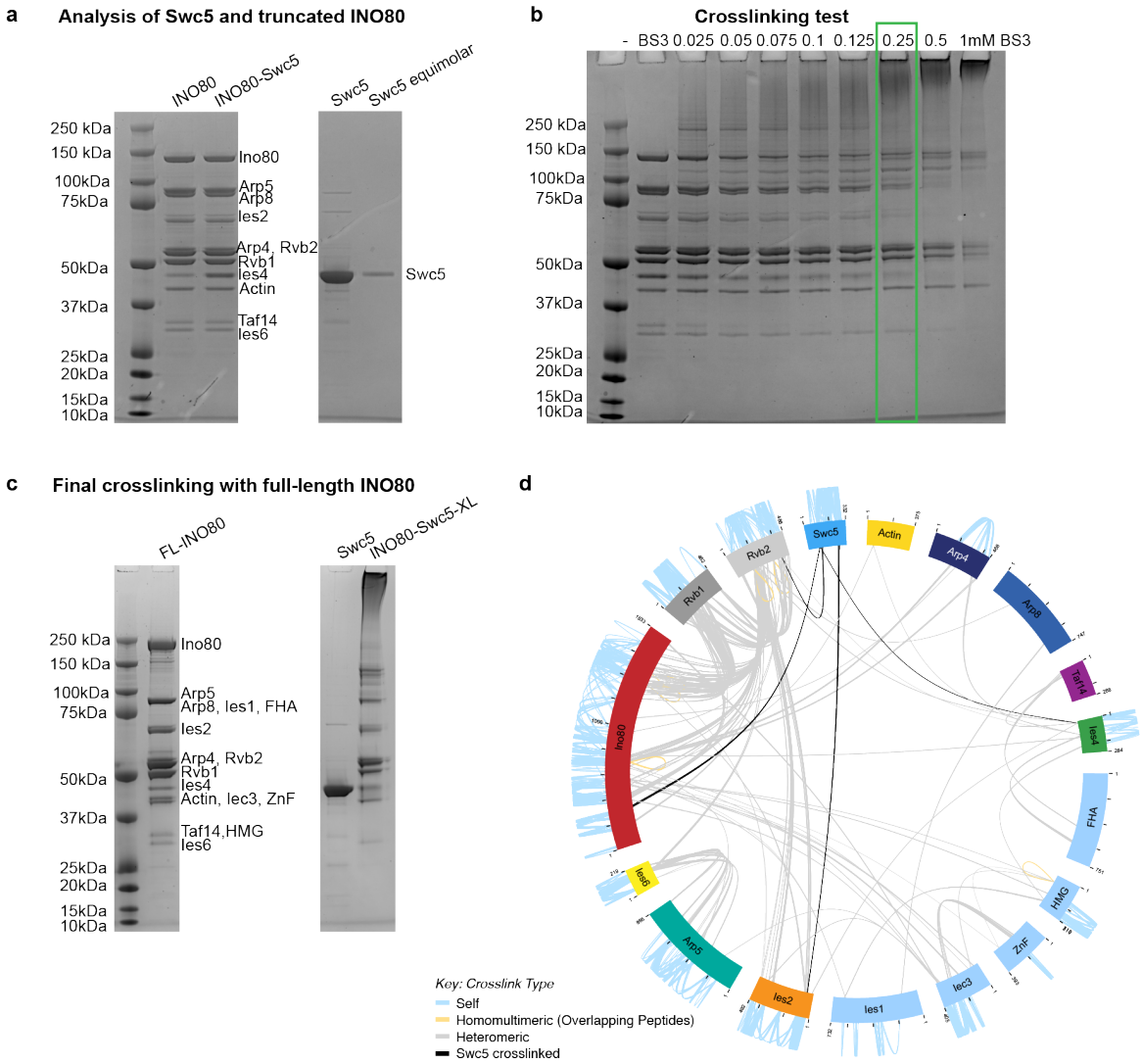


Figure 2.27: **XL-MS of INO80 and Swc5.** **a** Running behaviour of INO80 subunits and Swc5 on SDS-PAGE. **b** Crosslinking test of INO80 and Swc5 at different BS3 concentrations. A truncated 11-subunit INO80 construct was used. Smeared out bands at high molecular weight resemble crosslinked peptides. BS3 concentration that was used for final crosslinking is marked by a green square. **c** SDS-PAGE of full-length INO80, Swc5 and final crosslinked sample at 0.25 mM BS3. **d** XL-MS result depicting crosslinks by xiView tool [293]. Crosslinks between Swc5 and INO80 subunits are colored in black.

The analysis of the mass spectrometry data shows crosslinks between Swc5 and four INO80 subunits: the N-terminus of Ino80 itself, Ies2, Ies4 and Rvb2. Interestingly, these subunits belong to different modules of INO80: While Ies2 and Rvb2 are part of the INO80 Core module, Ies4 is located in the INO80 Arp module and the N-terminus of Ino80 is placed in the INO80 species-specific

module. This places Swc5 at a similar location as in the SWR1 complex. Here, Swc5 is placed between the Core and the Arp module of SWR1. In more detail, Nguyen et al. [222] detected crosslinks for yeast SWR1 between Swc5 and Yaf9, actin as well as Swr1 close to the HSA helix and the very C-terminus of the motor domain. I used recombinant *C. t.* SWR1 in a similar approach in which I found crosslinks of Swc5 to Rvb1, Rvb2, actin and the C-terminal region of Swr1 (Section 2.3.4). Nonetheless, the possibility remains that Swc5 artificially pulls down the INO80 subunits from the lysate. Michaelis and colleagues [296] recently created an in-depth yeast interactome based on high-affinity purifications and mass spectrometry. Here, they did not detect an interaction of Swc5 with any of the INO80-specific subunits. Still, the Swc5-INO80 interaction could be species-specific and is therefore only detected in *C. t.* However, more generally I might have detected an artefact that arises from the similar folds of INO80 and SWR1 especially around the shared subunits Rvb1, Rvb2, Arp4 and actin. Together they form the interface for Swc5 binding in SWR1. As described in [297] it is possible that during post-translational folding and complex formation the binding site interface of the Rvb1/Rvb2/Arp4/actin subunits of INO80 is not occupied by any of the INO80-specific subunits. As a consequence binding of Swc5 to INO80 is possible. In accordance, I detected a large proportion of proteins of the translation and protein folding machinery. *In vivo* the assembly of the SWR1 and INO80 complexes is likely more regulated preventing Swc5 from falsely associating with the SWR1 complex. Regardless, the performed MST measurement implies that the fully assembled INO80 complex is still capable of binding Swc5 *in vitro*. To finally answer the open question of the observed INO80-Swc5 interaction, I suggest to perform size exclusion chromatography on INO80-Swc5 and test if they co-migrate. Finally, the question remains what the function of Swc5 would be within the INO80 would be. In context of the SWR1 complex, Swc5 binds to H2A/H2B when it gets replaced by H2A.Z/H2B [226]. In this manner, it is prevented that the outgoing H2A/H2B dimer is immediately incorporated back into the nucleosome core particle. Therefore, the kinetics of the histone exchange reaction is pushed in the direction of the successful incorporation of H2A.Z/H2B. INO80 was suggested to perform the counter-reaction of SWR1 namely the exchange of H2A.Z/H2B by the canonical H2A/H2B [76]. This has been however a debate in the field [133][134] also because no histone chaperons had been described for INO80 which could bind the incoming H2A/H2B dimer and/or the outgoing H2A.Z/H2B copies. Based on my findings, it is tempting to speculate that Swc5 may play a role as a histone chaperone in complex with INO80.

2.3.7 Recombinant protein production of yeast SWR1

While I utilized *C. t.* SWR1 for structural studies in this thesis, I cloned and purified SWR1 from *Saccharomyces cerevisiae* to probe the histone exchange by SWR1 on a genome wide scale. I performed most of the cloning myself but got help from Olga Kolesnikova in the last few cloning steps. Additionally, Olga Kolesnikova continued my efforts of recombinant protein production and optimization of purification conditions while I had already started writing my thesis. However, we discussed closely the results and next steps.

Cloning of yeast SWR1 was performed likewise to *C. t.* SWR1. I again distributed the gene expression cassettes for the subunits on two plasmids (Figure 2.28a) with the core subunits on one plasmid and Arp module and potentially species-specific module on the other. However, for technical reasons, I had to put Swc4 that belongs to the Arp module, on plasmid one. Despite the correct DNA sequences and the presence of all subunits in the Bacmid DNA (Figure 2.28b), the subunits Swc3, Arp6 and Swc6 were missing from the first purification trials (data not shown). While for Swc3, this could be explained by a non-specific primer used for the underlying PCRs, it is unclear why the two others were missing from the purifications. Also, non-stoichiometric bands on Agarose gel not necessarily corresponded to non-stoichiometry protein expression, highlighting the challenges and

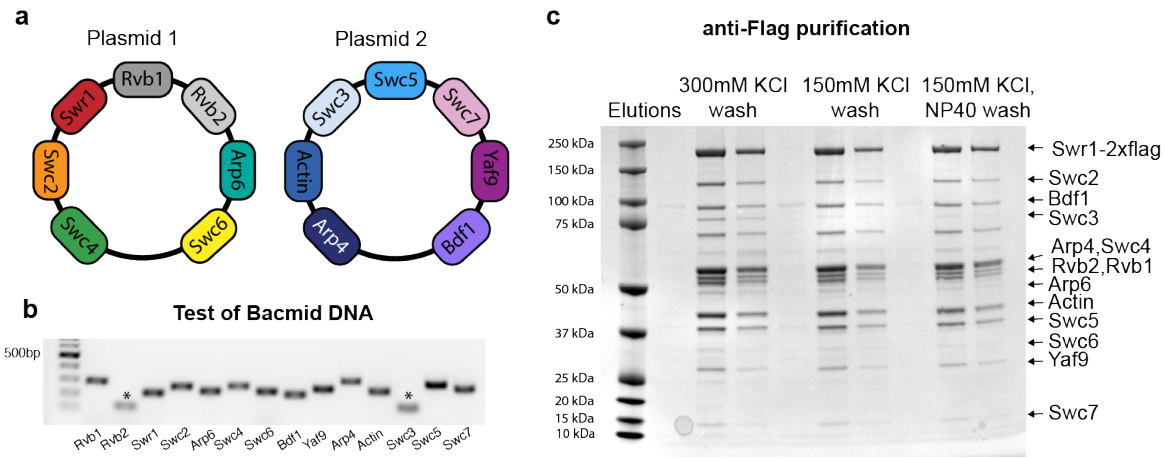


Figure 2.28: **Cloning and recombinant protein expression of *S. c.* SWR1. a** Plasmid overview for recombinant yeast SWR1 cloning and protein expression. **b** Agarose gel with analysis of Bacmid DNA for the two plasmids. PCRs were performed to check for the presence and stoichiometry of subunits on the Bacmid DNA. For this purpose primer pairs were used that generated PCR products of about 250 bp for better comparability on the gel. All subunits are present but Rvb2 and Swc3 bands are less intense indicating that their amounts are non-stoichiometric (marked with an asterisks). However, Rvb2 was expressed in stoichiometric amounts while the band for Swc3 was a false positive one. **c** SDS-PAGE of anti-FLAG affinity purification step. Illustrated are respecitively two elutions of different washing and elution buffer conditions: 1) 300 mM KCl 2) 150 mM KCl and 3) 150 mM mM supplemented with NP40 detergent. Identity of subunits is labeled respectively.

difficulties in recombinant protein expression of multi-subunit complexes. After preparing fresh Bacmid DNA, all subunits were detectable both on SDS-PAGE and in mass spectrometry analysis. In parallel, Olga Kolesnikova optimized the purification protocol. I observed that the small scale anti-FLAG purifications done in high-salt buffer (500 mM) looked promising on SDS-PAGE while the band pattern as well as protein yields turned out worse in large scale purification. For the large scale purifications I performed the anti-FLAG purification step at lower salt concentrations (150 mM) likewise to the INO80 purification. This implied that the difference in salt concentrations explains the different outcomes. Therefore, Olga Kolesnikova tested different washing and elution protocols for the anti-FLAG purification step to improve the yeast SWR1 elution. She compared 300 mM KCl with 150 mM plus/minus NP40 detergent for better solubility (Figure 2.28c). A larger extend of unspecific proteins were bound to the beads at the lower salt concentrations compared to the 300 mM KCl condition, indicating protein aggregation on the beads. Therefore, less protein could be eluted from the anti-FLAG affinity beads. The large scale purification was continued for the 300 mM KCl sample. As binding of SWR1 to the anion exchange column is salt dependent, the sample was diluted to 150 mM KCl prior to loading it to the column. Yeast SWR1 however elutes during a salt gradient at only around 500-600 mM salt. I therefore suggest to abstain from diluting SWR1 to lower ionic strength, to maintain complex stability. As mentioned before, the sample obtained from this purification, contains the 14 expected yeast SWR1 subunits as verified by mass spectrometry. Additionally, all subunits are visible on SDS-PAGE (Figure 2.28c) including Bdf1 and Swc7 which are absent in the *C. t.* SWR1 purifications. Bdf1 can be also co-purified with the remaining yeast SWR1 subunits in the lower salt condition which resembles the buffers used for *C. t.* SWR1. Therefore I conclude that it is not the purification conditions that lead to the loss of Bdf1 in the *C. t.* SWR1 purifications shown in Section 2.3.3. It is rather the absence of Swc7 in comparison to the yeast SWR1

complex in accordance to [221] where it is described that Bdf1 assembles only in the presence of Swc7. Despite the optimized purification protocol for yeast SWR1, the obtained protein yield and concentration after ion exchange chromatography remain low (about 0.5 µM). Through the usage of an ÄKTA micro, we hoped to obtain higher protein concentrations around at least 1-2 µM. In order to be able to adjust the buffer conditions in the following activity and genome wide assays, yeast SWR1 was concentrated using a spin concentrator which I usually tried to avoid with the mega-Dalton sized SWR1 complex.
In a next step, I suggest to test the biological activity of yeast SWR1 that can be assessed in histone exchange experiments similar to [235]. In brief, a native gel based assay is used in which SWR1 acts on H2A/H2B containing nucleosomes in presence of ATP. The H2A.Z/H2B substrate is triple FLAG-tagged allowing to differentiate between H2A and H2A.Z containing nucleosomes based on the different running behaviour on Native PAGE. Besides the assurance of having biological active yeast SWR1 at hand, this assay will be of particular use for the assessment of structure-function studies based on the structural studies of this thesis.
It took four years from cloning to obtaining the recombinant yeast SWR1 complex that contains all 14 subunits. Therefore, I could not perform the genome wide study in collaboration with the Korber lab in Munich as planned. To date, it is unclear what drives the specificity of SWR1 for +1 nucleosomes and whether there are DNA features that guide SWR1 to its substrate. We want to address these questions by testing the activity of yeast SWR1 on a DNA library that is used to reconstitute recombinant nucleosomes which are positioned by INO80 prior to adding SWR1. Using ChIP-Seq in context of anti-H2A.Z antibodies, one can study the specificity of SWR1 for DNA sequences and/or DNA shape features similar to the studies on INO80 [169][36]. With the help of INO80, the nucleosome positioning and more importantly the correct positioning of +1 nucleosomes is ensured, enabling SWR1 to incorporate H2A.Z/H2B into the +1 nucleosome. We therefore study at the same time, the interplay of INO80 and SWR1 around transcription start sites.
In summary, I established protein expression and purification of *S. c.* and *C. t.* SWR1. This success is a major advancement towards whole genome reconstitutions to reveal the molecular mechanism and functional requirements of histone exchange by SWR1.

2.4 Biochemical and structural characterization of the SWR1 Arp module

As previously described, the Arp module of SWR1 is not well characterized especially in respect of a possible second actin molecule. Truncation experiments on the Swr1 subunit however suggest the presence of a second actin copy in the SWR1 complex located next to the first actin copy [221]. Likewise, the performed mass photometry of full-length *C. t.* SWR1 showed the presence of a second actin molecule (Figure 2.23a). In order to study the Arp module more closely by biochemical and structural methods, I designed different Arp module constructs. This approach has the advantage that different nucleotide states of the actin-fold proteins – Arp4 and actin – can be studied separately of the overall SWR1 complex. This is paramount as SWR1 contains 10-11 ADP/ATP nucleotide binding sites. In addition, structural data on the Arp module alone will be of use in context of the full-length SWR1 complex given that the Arp module is presumably flexibly tethered to the core of the complex. This will prevent high resolution structure determination of the full complex. I got some help for this part of my PhD project from my master intern Lena Gottschalk, who joined me in my efforts for four months. She especially helped me with cloning of the constructs, protein production, mass photometry measurements and initial cryoEM trials.

2.4.1 Purification and biochemical characterization of the Arp module

The Arp module of SWR1 contains the subunits Arp4, actin and Swc4 which assemble onto the HSA helix of SWR1. I identified the amino acid sequence for the HSA helix based on *S. c.* Swr1. For biochemical and structural studies, I designed three different constructs that vary in the length of the Swr1 subunit. One construct – Arp$^{\text{HSA minimal}}$ – contained a minimal HSA version that theoretically should allow the assembly of the full Arp module including a second copy of actin. The second construct contains the HSA extension and is therefore termed Arp$^{\text{HSA FL}}$ (full-length). The third construct is complemented by the Swr1 domain at the N-terminus of the HSA helix (Arp$^{\text{N-HSA FL}}$). The rational behind the design of multiple constructs was to find a construct that is biochemically stable for structural studies. In addition, the length of Swr1 might influence the affinity for DNA binding as I expect Swr1 to behave similarly to Ino80 which binds with its HSA helix to DNA [150]. It was possible to purify the three constructs using the protein purification protocol from *C. t.* SWR1 (SDS-PAGE for Arp-module$^{\text{HSA FL}}$ construct in Figure 2.29b).

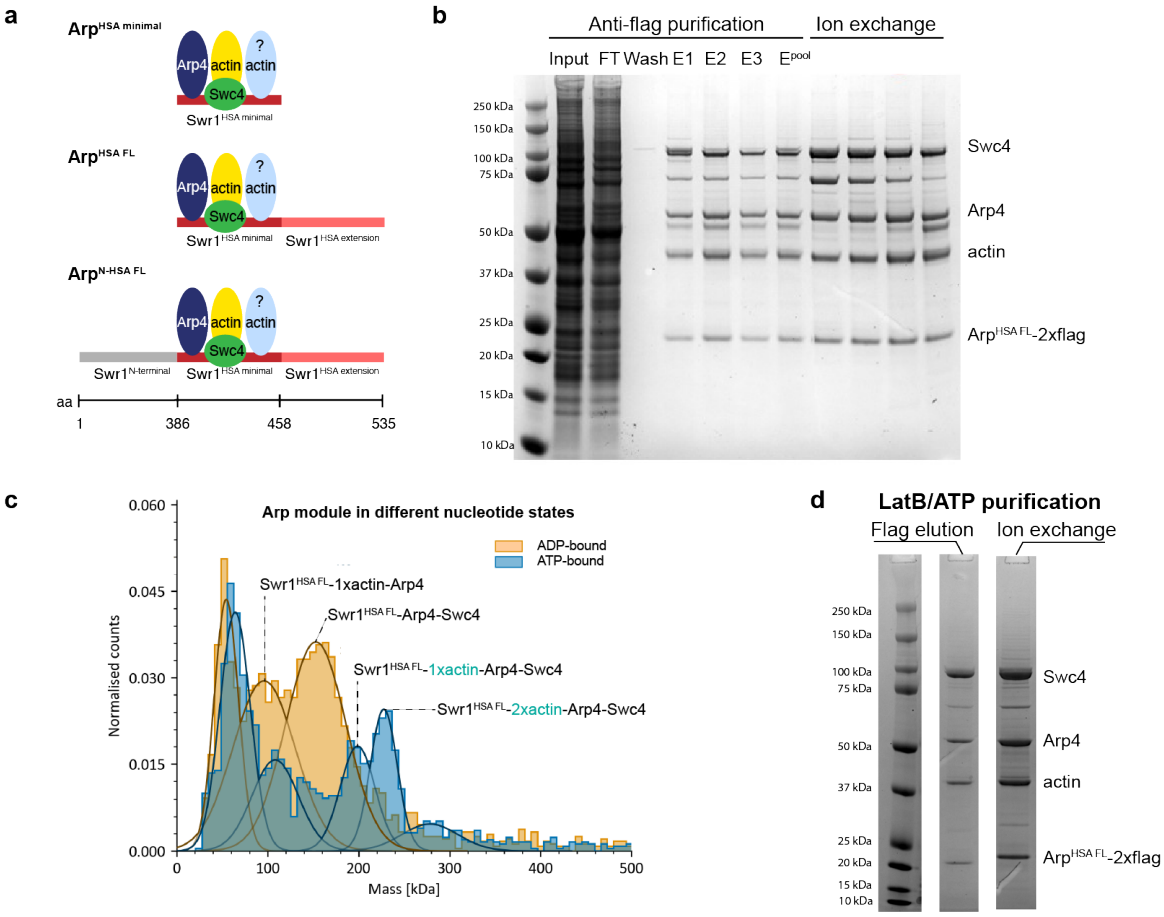


Figure 2.29: **Purification of Arp$^{\text{HSA FL}}$ and role of ATP on stabilizing a second copy of actin.** **a** Subunit composition of the three different Arp modules. **b** SDS-PAGE of Arp$^{\text{HSA FL}}$ showing only one copy of actin. **c** Mass photometry of Arp$^{\text{HSA FL}}$. A shift in presence of ATP is observed towards an Arp module with two actin copies. **d** SDS-PAGE of Arp$^{\text{HSA FL}}$ purification. Purified Arp module contains still only one actin copy despite the presence of 10 µM LatB and 1 mM ATP during the purification.

However, the constructs differed in stability: The Arp-module$^{\text{HSA minimal}}$ construct showed no peak at the expected size in mass photometry indicating its instability during the dilution process for the measurement. As it turned out later in the AlphaFold predictions, the chosen 'HSA minimal' is not sufficient to harbor a second actin copy, explaining the inherent instability of the construct. CryoEM grids produced from the Arp-module$^{\text{N-HSA FL}}$ construct did not look promising compared to Arp-module$^{\text{HSA FL}}$. Lena Gottschalk and I therefore focused on the Arp-module$^{\text{HSA FL}}$ construct for cryoEM that lacked the N-terminal region of Swr1.
Analysis of the purified Arp-module$^{\text{HSA FL}}$ construct on SDS-PAGE showed only one copy of actin (Figure 2.29b) based on its band intensity. The purification was performed in presence of calcium chloride which stabilizes ATP. The effect of ATP on the Arp-module$^{\text{HSA FL}}$ was tested by mass photometry showing a shift towards a complex with two actin molecules in the presence of ATP (Figure 2.29c). This is in line with data published on INO80 [154][150] where it was shown by X-ray crystallography and cryoEM that the Arp4 and actin molecule of the INO80-Arp module are ATP bound. In order to stabilize the second actin molecule, I purified the Arp module in presence of ATP and LatB. The latter stabilizes ATP, preventing its hydrolysis and ensuring that actin stays ATP bound throughout the purification procedure [298]. Despite the usage of LatB, the stoichiometry of the complex remained the same – no second actin molecule is visible on the respective SDS-PAGE (Figure 2.29d). In summary, a stabilization of a second actin molecule was observed in presence of ATP in single molecule analysis. However, this was only a fraction of the sample and a second actin copy could not be purified together with the remaining Arp module components. I therefore conclude that the presence of a second actin is possible but is not bound to the Arp module with high affinity which prevents its purification. Further evidence comes from the crosslinking data presented in Section 2.3.4. Here, at least one of the actin copies interacts with Swc5, Rvb1 as well as Rvb2 which are absent in the used Arp module construct. Therefore, it seems that a second actin copy requires the presence of the other Core module subunits. One actin copy dimerizes with Arp4 and binds to the Swr1 minimal HSA helix forming a stable complex that remains intact in the absence of other subunits [225]. I will discuss the presence of two actin molecules again in context of the full-length SWR1 complex in Section 2.5.

2.4.2 Structural characterization of the Arp module using cryoEM

Initial cryoEM data

Based on the presented data of the Arp module, the Arp-module$^{\text{HSA FL}}$ construct was used to determine its 3D structure by cryoEM. Lena Gottschalk and I prepared cryoEM grids of the recombinantly purified Swr1$^{\text{HSA FL}}$-Swc4-Arp4-actin complex in the presence of the detergent β-octylglycoside to prevent preferred orientation on the EM grids and the adherence of particles to the hydrophobic air water interface [262]. Despite the usage of detergent, only top and bottom views of the Arp module were observed in 2D classifications (Figure 2.30a). As a consequence, we collected another dataset at a tilt angle of 30° and merged the two datasets for collective data processing. However, the resulting 3D reconstructions still showed preferred orientations. The reconstructed cryoEM density is highly anisotropic (Figure 2.30b), which hindered accurate model building of the protein structure. Instead, I predicted a Swr1$^{\text{HSA FL}}$-Swc4-Arp4-actin model using AlphaFold2 Multimer [273] and fitted the model by rigid-body fitting into the cryoEM density. (Figure 2.30c, d). In the predicted structure, the core of the Arp module is formed by the actin-Arp4 dimer that assembles with its barbed end onto the Swr1 HSA helix (residues 395-435) as also shown experimentally in a previously published crystal structure [225]. C-terminal of the HSA helix resides the Swr1 HSA extension helix that I termed in accordance with Ino80 HSA$^{\alpha 1}$ [154]. Furthermore, AlphaFold predicted the interaction of

Swc4: Two β-sheets at its very N-terminus fold against the interface of the Arp4-actin dimer. From there, Swc4 folds into a joint consisting of multiple α-helices that pack against the unstructured N-terminal region of Swr1. The joint causes a kink within Swc4 so that Swc4 overall resembles the shape of the letter L. The AlphaFold prediction score of the Swc4 C-terminus is of low confidence (Figure 2.30d) and is therefore not interpreted. Overall, the core of the Arp module including the N-terminus of Swc4 with its joint domain, fits into the anisotropic 3D reconstruction (Figure 2.30e). This is apparent besides the preferred orientation.

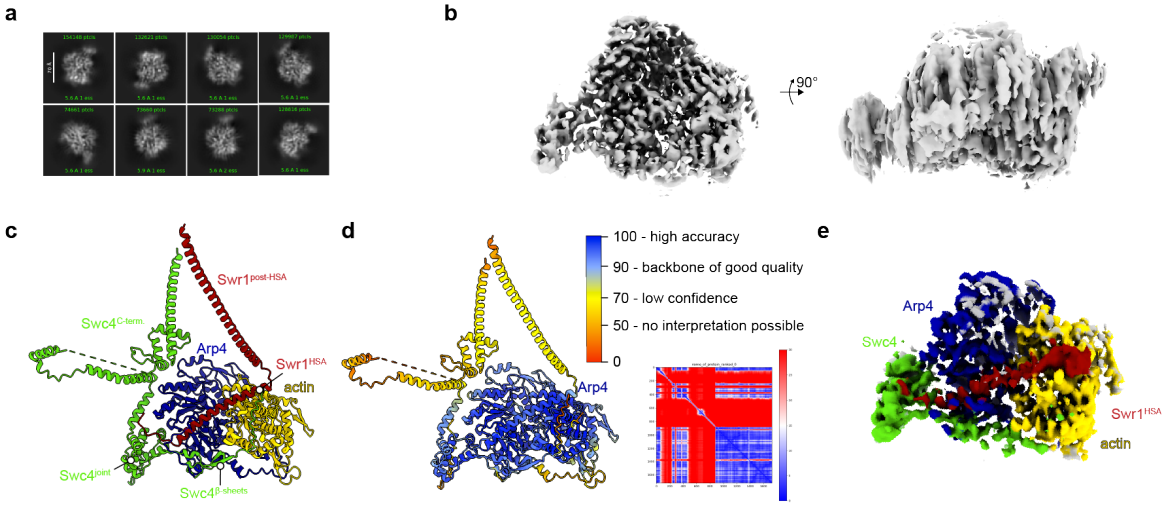


Figure 2.30: **Initial cryoEM data and AlphaFold models on the SWR1 Arp module. a** 2D classes from CryoSPARC showing only top and bottom views of SWR1. Scale bar on the left equals 70 Å. **b** 3D reconstruction of the Arp module in two views show anisotropic cryoEM density indicative of preferred orientation to the air water interface. **c** AlphaFold prediction of Swr1$^{\text{HSA-FL}}$-Swc4-Arp4-actin colored by residue and in **d** colored by prediction confidence score. On the right, PAE plot of the model. **e** 3D reconstruction colored based on fitted AlphaFold model.

Yaf9 is an integral part of the SWR1 Arp module

The crosslinking data on SWR1 in Section 2.3.4 indicated that Yaf9 is part or at least in close proximity to the Arp module. In more detail, Yaf9 was crosslinked to the N-terminus of Swr1 and to Swc4. For further investigation, I included Yaf9 into the AlphaFold prediction of the Arp module (Figure 2.31a). Indeed, AlphaFold predicts Yaf9 to interact with the Arp module: The Yaf9 helix protruding from its YEATS domain interacts with the helical architecture of the Swc4 C-terminal region. The N-terminal region of Swr1 folds against Swc4 and continues to the YEATS domain. Overall, the AlphaFold prediction is in agreement with the crosslinking mass spectrometry data. For further experimental validation, I performed a co-expression of the Arp module with untagged Yaf9, testing if Yaf9 can be pulled down when using the Arp module as bait protein. In depth, I used the Arp module construct containing the N-terminus of Swr1 (Swr1$^{\text{N-HSA FL}}$). Secondly, I generated viruses from plasmid number two of the full-length *C. t.* SWR1 expression as it contains among Bdf1, Swc3, Swc4 and Swc5, untagged Yaf9 and hence avoided the need for re-cloning. I prepared baculoviruses from the Arp module and Yaf9 Bacmid DNA and co-infected the insect cells for co-expression.

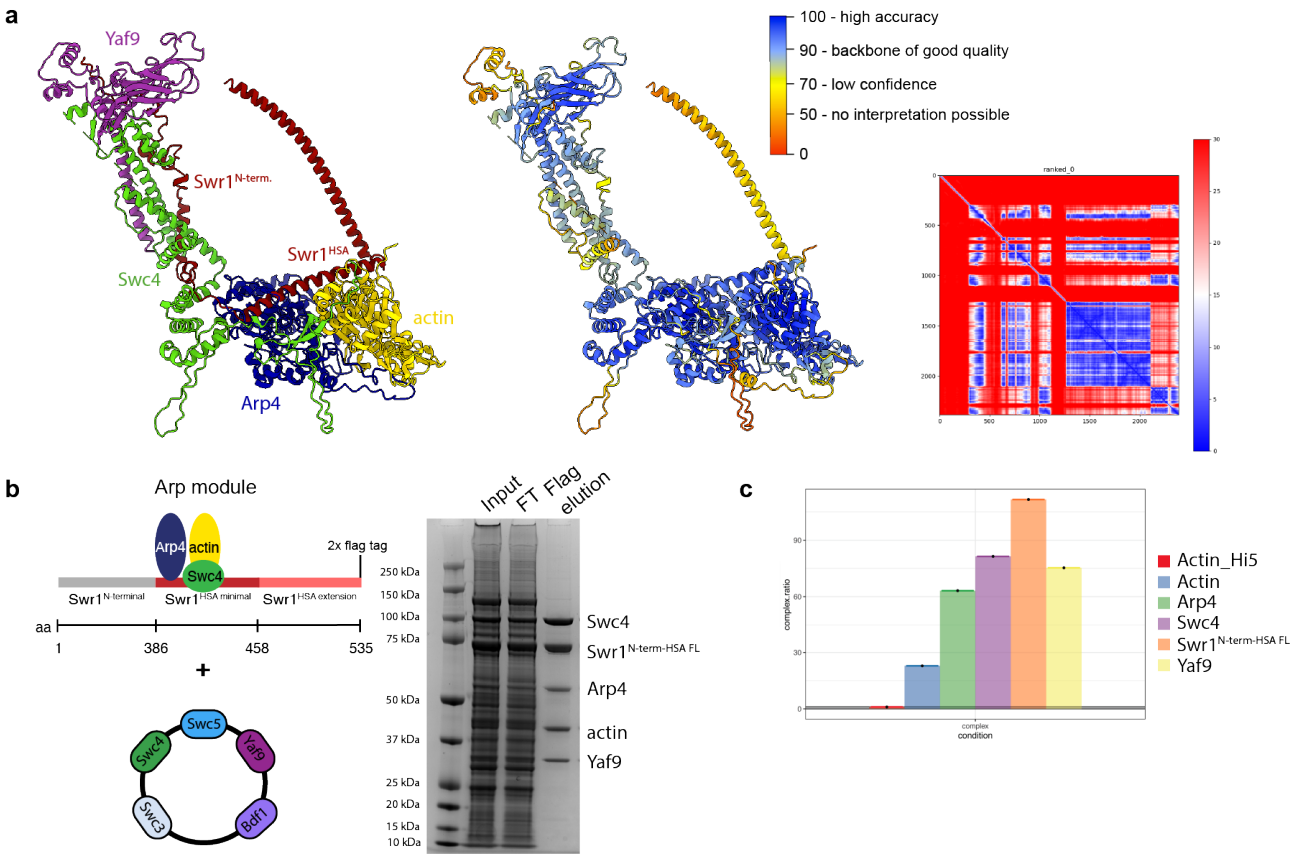


Figure 2.31: **Interaction of Yaf9 with the Arp module.** **a** AlphaFold Multimer prediction of the Swr1$^{\text{N-term-FL}}$ construct and one copy of actin, Arp4, Swc4 and Yaf9. Middle panel: prediction colored by confidence score and underlying PAE plot on the right. **b** Overview of co-expressed subunits (*left*) and anti-FLAG purification gel (*right*). **c** Stoichiometry of the Arp module-Yaf9 complex by mass spectrometry.

Subsequently, the Arp module was purified by anti-FLAG affinity purification by the 2*x* FLAG tag on the C-terminus of Swr1. If Yaf9 interacts with the Arp module, it should be pulled down alongside with the other Arp module subunits. The according SDS gel of the FLAG elution is shown in Figure 2.31b. The Yaf9 subunit co-eluted with the Arp module subunits Swr1-Swc4-Arp4-actin. The Arp module-Yaf9 complex also stayed intact during subsequent ion exchange chromatography. I therefore conclude, that Yaf9 is an integral part of the SWR1 Arp module. Following purification, an analysis by mass spectrometry confirmed the co-elution with Yaf9 in a one to one ratio to the Arp module subunits. Actin was underrepresented even when calculating with only one copy. As actin is well-conserved, I included the Hi5 host actin in the stoichiometry analysis. The amount of Hi5 actin was however negligible (Figure 2.31c). Of note is that Swc5 was not pulled down with the Arp module even though it is linked to actin based on the performed crosslinking mass spectrometry analysis (see Section 2.3.4). Presumably, Swc5 interacts only with the second actin copy that is absent in the used Arp module construct and requires other core subunits like Rvb1/Rvb2 to assemble to the Arp module.

CryoEM structure of the *C. t.* SWR1 Arp module

Next, I prepared cryoEM samples of the Arp module-Yaf9 complex. While the Arp module alone showed preferred orientation (Figure2.32a), the presence of Yaf9 led to a more diverse particle dis-

tribution (Figure 2.32b). With the presence of Yaf9, also side views of the Arp module were captured. However, the particle size itself did not differ in comparison to the previous Arp module sample indicating that the Arp module-Yaf9 complex had fallen apart. Accordingly, there were additional 2D classes present (Figure2.32c) that resembled the shape of the separate Yaf9 subunit. Of note is, that Yaf9 has only a molecular weight of 32 kDa, yet CryoSPARC picked it up in the 2D classification.

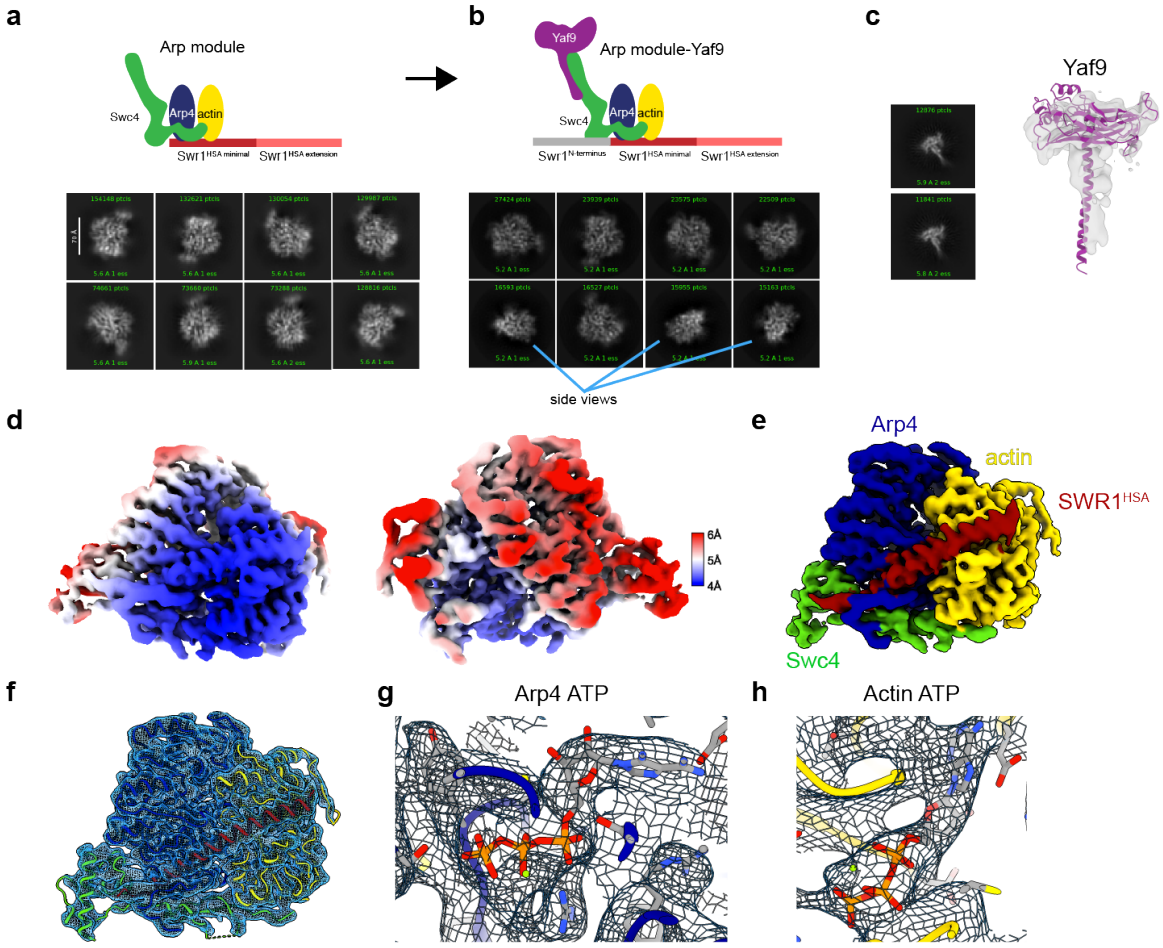


Figure 2.32: **Structure of the SWR1 Arp module. a** 2D classes of Arp module comprising the HSA helix, Arp4, actin and Swc4. Particles show only top and bottom views. **b** 2D classes of the Arp module comprising Yaf9 in addition. Now, also side views are visible. **c** 2D classes and 3D reconstruction with fitted Yaf9 AlphaFold prediction. **d** Arp module map colored by resolution. 3D refinement comprises an overall resolution of 3.8 Å. **e** Map colored by subunit. **f** Built models inside the final map in mesh representation. **g** ATP fitted in the nucleotide binding site of Arp4 **h** and actin.

Subsequently, I calculated a low resolution 3D reconstruction from 20000 particles into which the AlphaFold prediction of Yaf9 fitted. Attempts to increase the number of Yaf9 particles using the neural network based TOPAZ picker [268] failed. It was therefore not possible to obtain a 3D reconstruction of Yaf9 at higher resolution. In addition, classes with both the Arp module and Yaf9 were not observed. Nonetheless, I continued processing the Arp module as the observed particles were of good quality. While I pursued the analysis, RELION 5.0 was released [252] which implements Blush. Blush is a denoising convolutional neural network, that with regularization during the iterative refinement procedures has been described to work especially well for small particles in 3D classifications and 3D refinements. The Arp module with the ordered heterodimeric Arp4-actin

core of around 100 kDa can be still considered small. I performed two rounds of 3D classifications followed by a masked 3D refinement. Blush refinements resulted in better quality compared to a volume obtained in RELION 4.0 [263]. However, the blush tool showed a tendency to hallucinate especially in 3D classification. This means that density parts were connected while the AlphaFold predictions did not show an interaction at these sites. 3D classes had hence to be carefully selected. After post-processing, I obtained a map at overall 3.4 Å resolution. As the map was fragmented, I preferred working with the 3D refined model from the step before which was at 3.8 Å. Local resolution mapping within RELION (Figure 2.32d) revealed a heterogeneous resolution ranging from 4-6 Å. The latter was present especially at the rim of the volume and the bottom view, resulting from a still present preferred orientation. The quality of the map was nonetheless sufficient for model building (Figure 2.32f) which I performed based on the AlphaFold model presented in Figure 2.30c. The EM map represents the core of the SWR1 Arp module which consists of Arp4, one copy of actin, the SWR1 HSA helix and the Swc4 joint and 2W hairpin domains (Figure 2.32e). Arp4 and actin are clearly ATP bound highlighting again the quality of the obtained 3D volume (Figure 2.32g, h). In more detail, Arp4 and actin assemble via their barbed ends as a hetero-dimer to the HSA helix of SWR1 (residues 724-775) (Figure 2.33a). The Swc4-2W hairpin is anchored onto actin (proline residue 367 of the actin molecule) through its two tryptophans on the two β-sheets that give rise to the underlying domain name. The loop connecting the two β-sheets folds alongside the HSA. The Swc4-2W hairpin domain is connected to the Swc4-joint domain via an extended loop that runs in parallel to the C-terminal region of the Swr1 HSA helix.

Comparison of the SWR1 and INO80 Arp modules

The Arp module of SWR1 resembles the Arp module of INO80 in regard to its architecture and subunit composition. For comparison, I used the presented SWR1 Arp module structure as well as the INO80 cryoEM structure from [150] while there is also a crystal structure available lacking Ies4 [154]. Both Arp modules consist of one copy of actin and Arp4 that assemble onto the Swr1 and Ino80 HSA helices respectively (Figure 2.33a,b). In more detail, the Arp4 and actin molecules harbor an ATP molecule in their nucleotide binding pockets both in the SWR1 and INO80 complexes. Actin is anchored via tryptophan anchors onto the HSA helices of Swr1 and Ino80 respectively. In addition, the Swc4 (form SWR1) and Ies4 (from INO80) subunits both comprise a so called 2W-hairpin domain [150]. Two conserved tryptophans anchor Swc4 and Ies4 to the nearby proline of actin (Figure 2.33e) while the 2W-hairpins themselves are not conserved in terms of amino acid sequence. A more obvious difference between the Arp modules, is their composition: The Arp module of INO80 is completed by Arp8 that assembles with its barbed end onto the C-terminus of the Ino80 HSA helix. Furthermore, Arp8 extends through an unstructured loop domain all the way to the N-terminus of the Ino80 HSA helix where it folds into the so called Arp8 hook [150]. Instead of Arp8, SWR1 contains a second copy of actin which I did not capture in the presented structure and is therefore not further discussed at this point (but in Section 2.5). More differences become apparent when comparing the Arp module structures in more detail. The HSA helices differ in their N-terminal region as the Ino80 HSA helix is ten residues and thereby three helical turns longer (Figure 2.33c, d). The helical architecture is further stabilized by the nearby Arp8-hook which is absent in SWR1. Remarkably, the two structures differ at the N-terminus of Arp4. In SWR1, it is resolved from residue 3 onwards, while it was only captured form residue 13 in the INO80 complex indicating a higher degree of flexibility.

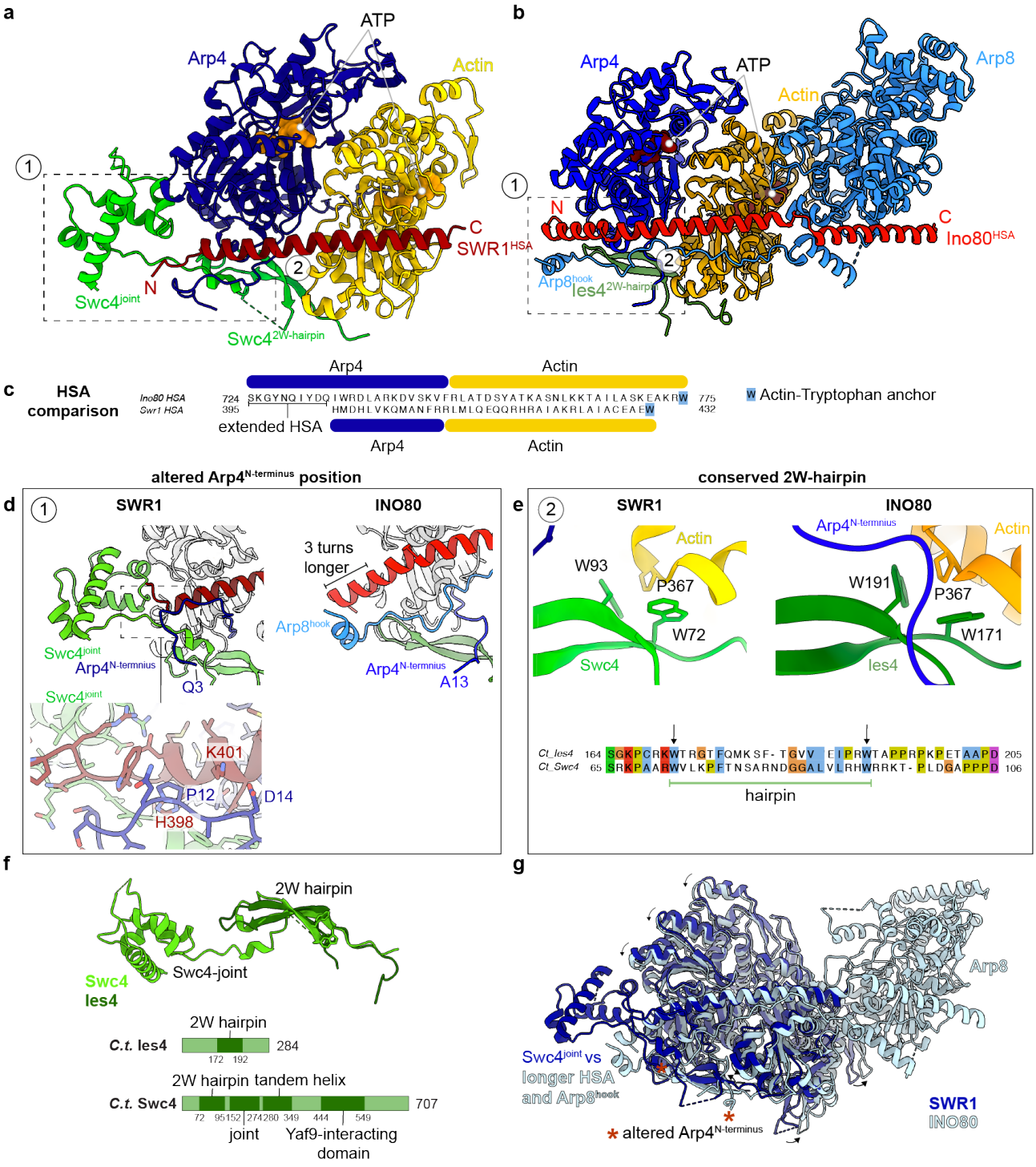


Figure 2.33: **Structural comparison of the SWR1 and INO80 Arp modules.** **a** Structure of SWR1 Arp module. Enlarged view of (1) is shown in **d** and (2) in **e**. **b** Structure of INO80 Arp module (PDB 8A5D). **c** Amino acid sequence of Ino80 and Swr1 HSA helices. Interactions with Arp4 and actin are indicated. Actins are anchored to the HSA via tryptophan residues. **d** Comparison of location of Arp4$^{\text{N-terminus}}$ (dark blue). Bottom: Close-up of Swr1 HSA-Arp4 interaction. **e** Interaction of 2W-hairpin of Swc4 and Ies4 with actin residue P367. Bottom: Snippet of sequence comparison of Swc4 and Ies4. The two tryptophans are highlighted with arrows. **f** Domain architecture of Swc4 and Ies4. **g** Overlay of SWR1 (dark blue) and INO80 (light blue) Arp modules on HSA helices. Rotation of INO80 Arp module in relation to SWR1 is indicated by arrows. Altered position of Arp4$^{\text{N-terminus}}$ is marked with red asterisks.

Furthermore, the location of the N-terminus itself differs: In INO80 it folds onto the Ies4 2W hairpin close to the two tryptophans, while in SWR1 it extends alongside the N-terminal region of the Swr1 HSA helix (Figure 2.33d). Two of the interacting residues are Arp4-P12 and Swr1-H398 that stack on top of each other. In the Ino80-HSA helix such an aromatic residue is lacking at this location and it is hence the difference in sequence between the HSA helices that results in the altered Arp4-N-terminus location. In addition, The Arp4-N-terminus within SWR1 gets stabilized by Swc4 features that are not shared with Ies4. Ies4 as such is a smaller protein in comparison to Swc4 with less features (Figure 2.33f). Swc4 contains in addition to the shared 2W hairpin, the joint domain present also in the structure, a tandem helix and a Yaf9-interacting domain. On a global scale, the structural similarities and differences become apparent when overlaying the two structures on the HSA (Figure 2.33g) helix. The Arp4-actin dimer of INO80 is shifted anti-clockwise in relation to SWR1. Yet, it remains unclear what causes this minimal rotation on a molecular level. The longer Ino80-HSA helix would clash with Swc4 explaining why the Swr1-HSA helix is shorter at its N-terminus.

2.5 cryoEM on the SWR1 core complex

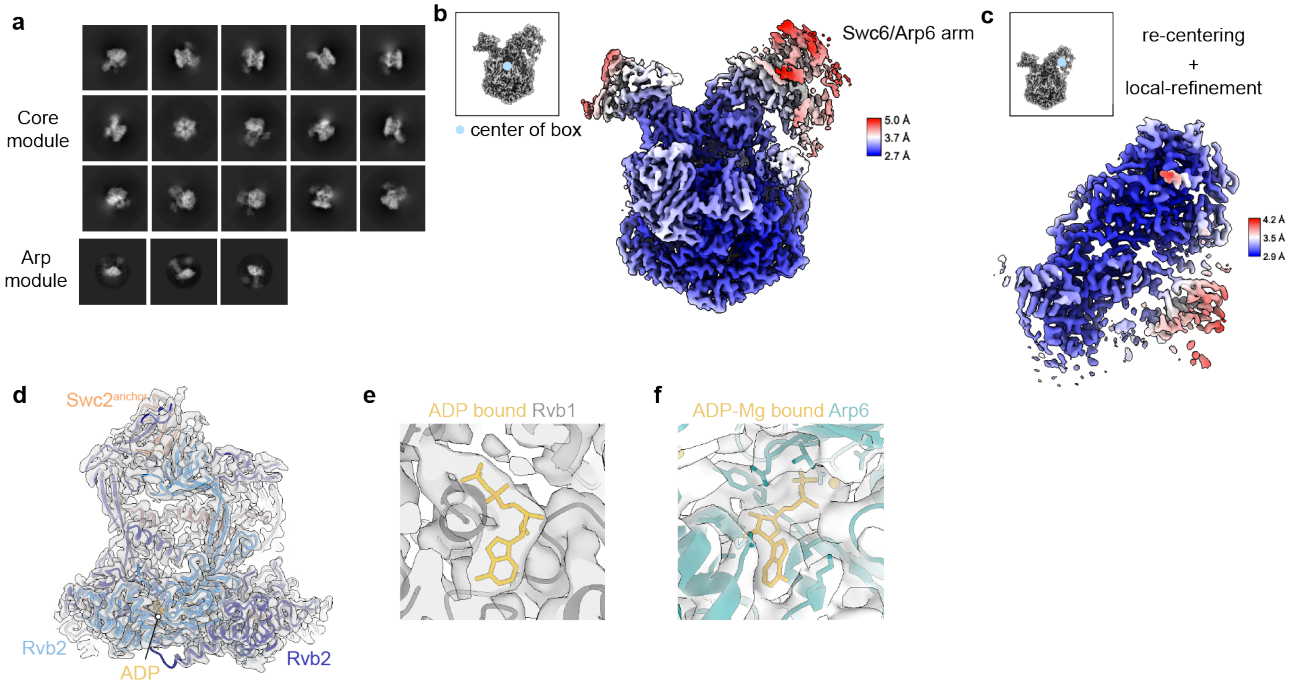


Figure 2.34: **Data processing and model building of the SWR1 Core module. a** 2D classes of Core module and Arp module particles with a box size of 360 Å. **b** 3D refined map of the SWR1 Core module colored by local resolution. The center of the box lays close to the Rvbs as shown on the left. The Swc6/Arp6 arm is of worse resolution. **c** 3D refinement after re-centering and local refinement on Swc6/Arp6; colored by local resolution which is significantly better compared to **b**. **d** Example map of the Rvb1/Rvb2 hetero-hexamer showing the good quality of the obtained data. **e** Nucleotide binding site of Rvb1 occupied by a molecule of ADP. **f** Arp6 is bound by an ADP molecule that is coordinated by a magnesium ion (yellow sphere).

The structural studies on the Arp module of SWR1 laid the foundation to determine the apo structure of the full-length SWR1 complex. For this purpose I prepared cryoEM grids of the recombi-

nantly purified *C. t.* SWR1 at a concentration of 600 nM and in a buffer with high ionic strength including 500 mM sodium chloride. I collected a dataset of 5000 micrographs which I pre-processed in RELION 4.0 [263]. 2D classes in CryoSPARC showed two major types of classes (Figure 2.34a). The majority of classes were projections of the Core module of SWR1. A minor fraction of classes showed a similartiy with the Arp module 2.4.2, yet indicated additional densities which were not observed for the isolated module. A more in depth analysis of the Arp module classes will follow at the end of this chapter. Neural network based particle picking by TOPAZ increased the number of particles [268]. The resulting 900000 particles were 3D classified in three consecutive rounds resulting in a final class of 166166 particles which was refined in RELION 4.0 [263] to 2.97 Å resolution (Figure 2.34b). The density around the Swc6/Arp6 subunits showed however a lower resolution of 3.7-5 Å, most likely due to structural heterogeneity of the Swc6/Arp6 subunits in relation to the Rvb1/Rvb2 hetero-hexamer. To refine this part of the structure to higher resolution, I re-centered the Swc6/Arp6 region and performed a new 3D refinement. In addition, I utilized a local refinement by restarting the 3D refinement at an intermediate iteration and by providing a mask around the Swc6/Arp6 region. This resulted in a significantly improved map depicted in Figure 2.34c with an overall resolution of 3.6 Å. In addition, the rim regions of the map that correspond mainly to Swc6, are more connected which is important for the following model building process. The underlying focused map contains in addition to the Arp6 and Swc6 subunits, the Swr1 anchoring helices and the OB fold of the close by Rvb2 copy which is of poor resolution in the overall map and was hence built in the focused map instead (Figure 2.35c). The obtained maps were sufficient for model building which was performed based on either published *C. t.* structures or alternatively generated AlphaFold multimer models [273]. Model building of Swc6/Arp6 as well as the Swr1 anchoring helices and a small part of one Rvb2 copy were first built into the re-centered and locally refined map. The built models were subsequently fitted into the overall map, to obtain a full structure of the SWR1 Core module allowing better interpretation and figure presentation. The good quality of the maps is highlighted in Figure 2.34d. Furthermore, the Rvb1 and Rvb2 subunits are bound to ADP (Figure 2.34e) likewise to the Rvb1/Rvb2 hetero-hexamer from INO80 (this study and [77][150]). Arp6 is bound to ADP that is coordinated by a magnisium ion (Figure 2.34f) likewise to yeast Arp6 in the yeast SWR1 complex [142].

Structural features of the apo SWR1 complex

Based on the published yeast SWR1-nucleosome structure [142], I interpreted the obtained cryoEM map of the SWR1 core and assigned the map to the different subunits. This enabled me to perform model building based on the Rvb1/Rvb2 AAA^+ ATPase from the presented INO80-hexasome complex as well as AlphaFold predictions of the respective *C. t.* subunits. The map of SWR1 comprises the Rvb1/Rvb2 AAA^+ ATPase, the Swr1 insert domain, Arp6, Swc6 as well as parts of the Swc2 subunit. The Swr1 motor domain was not captured due to its flexibility in absence of a nucleosome. Nonetheless, it was beginning to appear at high counter level. Besides the identified densities, there was an additional density on top of Swc2 that could not be assigned and is absent from the *S. c.* SWR1 structure from [142] (Figure 2.35a). Early biochemical studies of *S. c.* showed an interaction of Swc2 with the Swc3 subunit [224] that was also captured in the presented SWR1 crosslinking mass spectrometry experiment (Figure 2.23). I therefore created an AlphaFold Multimer model of the Swc2/Swc3 dimer (Figure 2.35c) that fitted well into the additional density on top of Swc2 and that complements the apo structure of the SWR1 Core module.

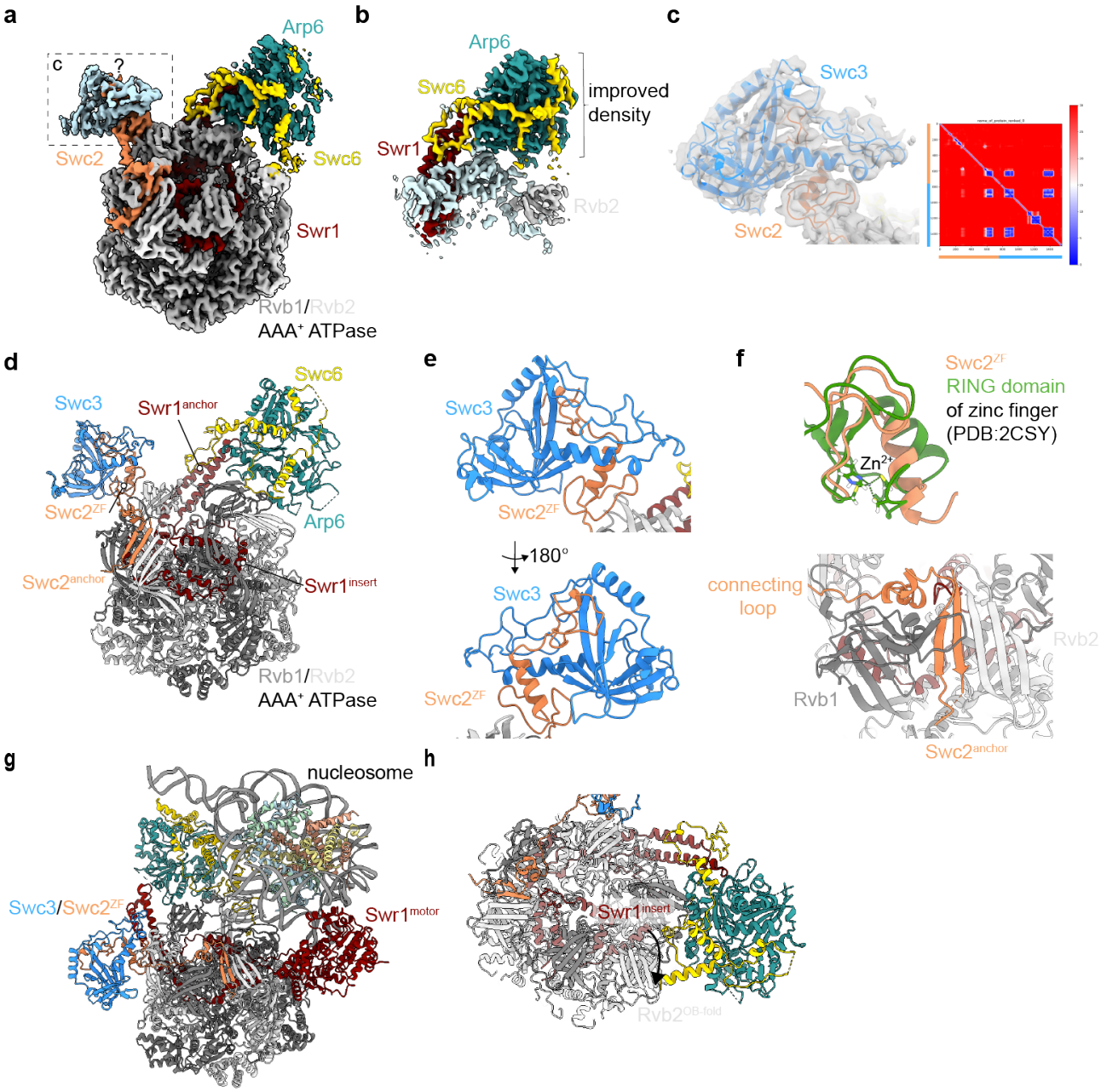


Figure 2.35: **Core module structure of SWR1. a** Map of overall Core module colored by subunit. In light blue, unassigned density that is further characterized in **c**. **b** Re-centered and focused refinement around Swc6/Arp6 dimer colored by subunit. In red Swr1 anchoring helices of the Swr1 insert domain and in gray OB fold of the close by Rvb2. **c** Fitting of an AlphaFold prediction identifies cryoEM density as Swc2 and Swc3. PAE plot of the prediction on the right. **d** Structure of the apo SWR1 Core module. **e** Two views of Swc3 bound to the Swc2 zinc finger (ZF). **f** *Top*: Comparison of Swc2 zinc finger with RING domain zinc finger (PDB: 2CSY). *Bottom*: Swc2 anchor folds against β-sheets of Rvb2 OB fold. **g** Swc2/Swc3 dimer in context of the *S. c.* SWR1 structure bound to a mono-nucleosome. Swc2 and Swc3 from this study, remaining structure PDB 6GEJ [142]. **h** Swr1 insert causes displacement of Rvb2 OB fold close to Swc6/Arp6 dimer; indicated by a black arrow.

At the base of the core complex lays the Rvb1/Rvb2 AAA$^+$ ATPase that is intersected by the Swr1 insert domain (Figure 2.35d). It is also one of the Swr1 insert helices that causes a displacement of the

Rvb2 OB fold close to Swc6. As a consequence, the hetero-hexameric Rvb1/Rvb2 barrel is asymmetric (Figure 2.35h). The Arp6/Swc6 dimer is located on top of the Rvb1/Rvb2 hetero-hexamer and makes additional interactions with anchoring helices of the Swr1 insert domain. Swc3 sits on top of the Swc2 zinc finger domain on the left to the Swc6/Arp6 arm. Besides the zinc finger, a loop region of Swc2 C-terminal to the zinc finger folds against Swc3 as illustrated in Figure 2.35e. Sequence alignment shows that the Swc2 zinc finger fold is degenerated and no longer contains the necessary cysteine or histidine residues to coordinate a zinc ion which is consistent with the absence of a zinc ion in the obtained cryoEM map. Figure 2.35f shows a comparison with another zinc finger protein that contains cysteine and histidine residues to coordinates a zinc ion. The degenerated zinc finger domain of Swc2 serves as an anchor for Swc3. A short disordered loop connects the Swc2 zinc finger domain to the Swc2 anchoring hairpin that packs in between the OB fold of Rvb1/Rvb2. (Figure 2.35f). Superimposition of the Swc2/Swc3 dimer onto the published *S. c.* SWR1 structure (PDB 6GEJ) [142] places Swc3 away from the mono-nucleosome and the Swr1 motor domain. This is in line with biochemical data that showed that the deletion of Swc3 does not impair nucleosome binding nor ATPase activity [228]. In addition, the deletion of Swc3 in context of the *S. c.* SWR1 complex does not result in the depletion of any other SWR1 subunit [224]. Swc3 is an evolutionary conserved component of SWR1 and its funciton remains to be identified.

The Arp module is flexibly tethered to the SWR1 Core module

Next, I focused on analyzing the putative Arp module observed in the SWR1 dataset. Initial 2D classification identified approximately 34000 particles of this module. I therefore performed TOPAZ training and picking as described previously, which enhanced the number of Arp modules to 183000 (Figure 2.36a). 2D class averages indicated that the particles aligned well on the core of the Arp module, while the two protruding densities are smeared out. This degree of structural flexibility prevented me from determining a high resolution map, as it could not be circumvented by further 3D classifications. Trying other picking methods like template based algorithms to center the particles on a common feature should be therefore tried in the future. Despite the limiting resolution of the obtained cryoEM density for the Arp module, I could interpret the density based on my structural study of the isolated Arp module. In addition, I made use of extensive AlphaFold multimer predictions. Besides the core of the Arp module, I fitted the Arp module foot that is composed of the N-terminal region of Swr1, the C-terminal helices of Swc4 as well as the Yaf9 subunit (Figure 2.36b). The AlphaFold prediction of the Arp module core and foot components presented previously in Figure 2.31 had to be adapted as the angle between the core and foot part was not correctly predicted by AlphaFold. The junction of the two parts is hence not accurate and can only be improved by obtaining a cryoEM map at higher resolution. The second arm that protrudes from the Arp module consists of another copy of actin bound to the $\alpha 2$ helix of the Swr1 HSA domain which is hereafter refer to as actin^{B} while the actin that dimerizes with Arp4 is actin^{A} (Figure 2.36b,c). Looking at the HSA domain in more detail, it is composed of the two helices $\alpha 1$ and $\alpha 2$ that are connected by a short loop of 8 residues. Arp4 and actin^{A} together with the Swc4 subunit assemble onto $\alpha 1$ while it is only actin^{B} that binds to the $\alpha 2$ helix (Figure 2.36d). To further map the interaction of actin^{B}, I performed an additional AlphaFold prediction of just the $\alpha 2$ helix and actin^{B}. Based on these predictions I was able to assign and locate the actin^{B} molecule in the low-resolution cryoEM map and to derive a more complete model of the Arp module. Interestingly, this revealed that two actin molecules bind to the HSA helix with their barbed end and are further anchored through a common tryptophan residue (Figure 2.36e). Similarly to SWR1, also within the INO80 Arp module, its actin and Arp8 subunits are anchored by a tryptophan residue at the end of their interaction site (Figure 2.36g).

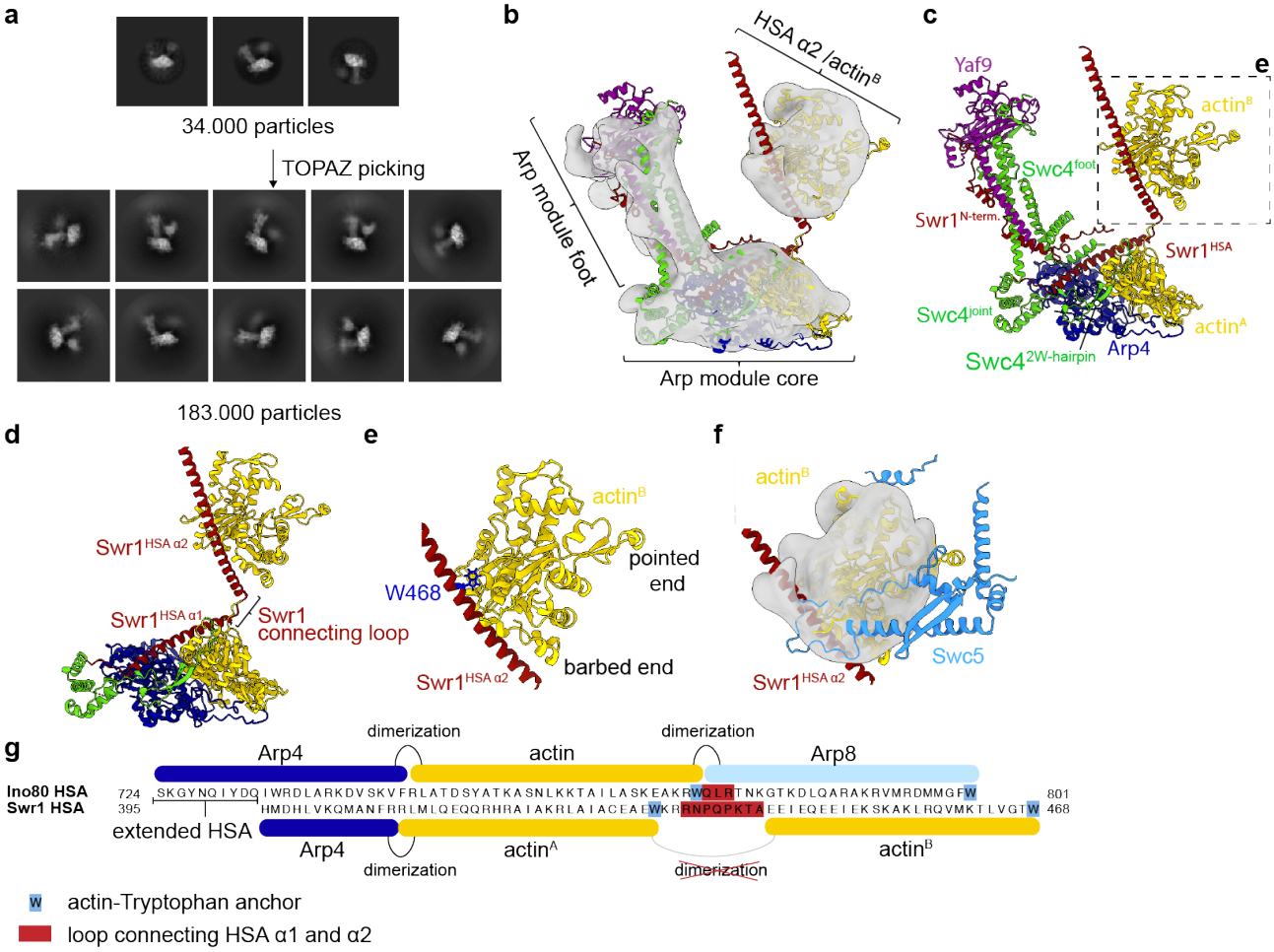


Figure 2.36: **Structural characterization of the Arp module in context of apo SWR1. a** 2D classes of the Arp module prior and after TOPAZ based particle picking (box size 360 Å). **b** Obtained low resolution 3D refined map from RELION with fitted AlphaFold models. Models of Arp module foot, core and HSA α2/actin^{B} were fitted separately. **c** AlphaFold model of the SWR1 Arp module. Dashed square indicates close-up view detailed in **e**. **d** The SWR1 HSA domain consists of the SWR1 HSA α1 and α2 helices that are connected by a loop. **e** Close up of actin^{B} that is anchored by the Swr1 tryptophan anchor (residue W468). **f** Prediction of Swr1 HSA α1/actin^{B}/Swc5 does not fit into the density. **g** Comparison of Ino80 and Swr1 HSA domains similar to Figure 2.33. In blue the common tryptophan residues that anchor actin and Arp8 molecules respectively. Loop connecting HSA α1 and α2 helices in red.

In both complexes, the HSA helix is interrupted in between the actin and Arp8/actin binding sites by a short loop. The loop within Ino80 is however only 3 residues long while it is comprised of 8 amino acids in Swr1. In addition, the Swr1 HSA α1 helix contains one more helical turn following the tryptophan anchor. As a consequence, the two binding sites of the two actin molecules are further away compared to the actin/Arp8 pair in INO80. Therefore, the dimerization of the two actin copies is prevented in SWR1. In conclusion, all SWR1 subunits could be assigned to a density of the apo SWR1 structure, except for Swc5. Based on the crosslinking mass spectrometry data from Section 2.3.4, I suggested an interaction of Swc5 with actin^{B} that sits at the HSA α2 helix. The low resolution cryoEM density around actin^{B} is large enough to harbor another protein component. However, an AlphaFold prediciton of Swr1$^{HSA\alpha2}$, actin^{B} and Swc5 did not fit into the map

(Figure 2.36f). Nonetheless, AlphaFold did predict an interaction of Swc5 and actin[B]. Possibly, Swc5 is too flexible to be captured by cryo electron microscopy.

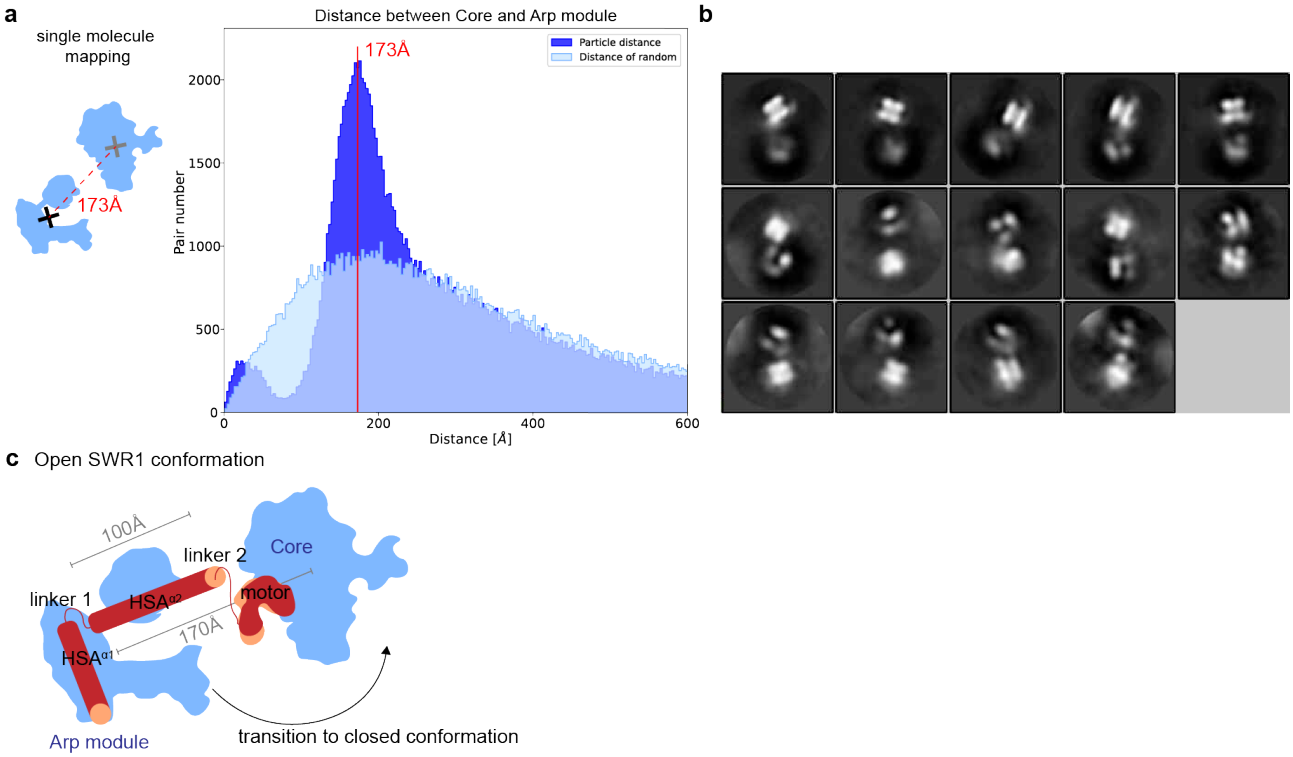


Figure 2.37: **Single molecule mapping of Arp and Core modules. a** Histogram of nearest neighbor pairs of Core and Arp module particles based on their center of mass. Pictogram on the left explains the method. Center of mass of the two modules is marked with a cross respectively. **b** 2D classes obtained through newly determined coordinates based on the single molecule mapping allowing to illustrate the Core/Arp module complexes. **c** Schematic illustration of the SWR1 complex with the Core and Arp module flexibly tethered. Swr1 subunits are colored in red. 2D classes showed an open conformation of the SWR1 complex while in a closed conformation the Arp module would fold back onto the Core module as indicated by an arrow.

The separate 2D classes of SWR1 Core and Arp modules, raises the question if the complex has fallen apart during the cryoEM sample preparation. To investigate this possibility, I collaborated with Luis Hauptmann, who performed a single molecule mapping analysis of the apo SWR1 dataset. He mapped nearest neighbors of Core module and Arp module pairs and plotted them in a histogram (Figure 2.37a). The histogram indicates a non-random peak at a particle distance of 173 Å. Additionally, Luis Hauptmann obtained new particle coordinates that lay in between the two modules allowing me to re-extract the particles with a larger box size. Subsequently, I performed 2D classifications of the Core/Arp modules (Figure 2.37b). The 2D classes show low resolution densities which resemble the Core and Are modules. Some of these classes show a connection between the two modules. AlphaFold predictions of the Swr1 protein show that the HSA α2 helix and motor domains are intersected by an unstructured region of 150 residues. Based on this data, I conclude that the Arp and Core modules of the SWR1 complex are flexibly tethered, however at a preferred distance of about 173 Å (Figure 2.37a,c). Due to the unstructured region of Swr1 it is not possible to map and confirm the 173 Å distance. It is however in a similar range compared to the INO80 complex [281]. Interestingly, the 2D classes of the Arp/Core SWR1 complex indicate an open conformation.

Previous negative staining EM studies reported a mixture of open and closed conformations of *S. c.* SWR1 [189].

2.6 Towards a SWR1-nucleosome structure

The presented work on the SWR1 Arp and Core modules lays the basis for the structural understanding of the histone exchange activity of SWR1 in complex with nucleosomal substrates. Many questions remained unresolved form the previous structural studies of *S. c.* SWR1 [142]. First, this SWR1 structure in complex with a nucleosome resolved only the minimal core which lacks the histone chaperone Swc5 and the Swc3 subunits. Additionally, the Arp module and N-terminal modules are absent from the structure and it remains unclear how these modules are linked to the core of SWR1. Only a limited amount of SWR1-histone contacts was resolved, thus, it remains unclear how SWR1 incorporates H2A.Z/H2B dimers while it removes H2A/H2B and why SWR1 does not translocate nucleoosmes. To tackle these open questions, I used cryoEM to determine a structure of the thermostable *C. t.* SWR1-nucleosome complex.

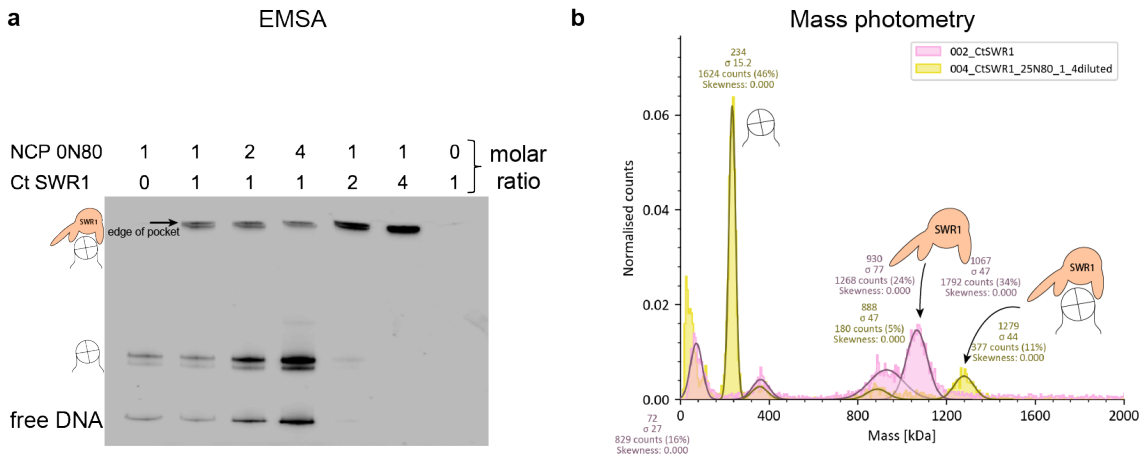


Figure 2.38: **Binding studies of SWR1 to nucleosomes. a** Electromobility shift assay (EMSA) at different molecular ratios. **b** Mass photometry comparing apo SWR1 with SWR1 bound to 25N80 nucleosomes. Identity of peaks are labeled with pictograms.

Before starting to screen optimal conditions for the grid preparation of the SWR1-nucleosome complex, I biochemically probed the binding of the recombinantly purified *C. t.* SWR1 to different nucleosomal substrates. First, I tested the binding to 0N80 nucleosomes which have been used for the structural characterization of the INO80-nucleosome complex (Section 2.1) using an electrophoretic mobility shift assay (EMSA). It allows to probe binding by detecting a shift of SWR1-bound nucleosomes in relation to free nucleosomes on Native PAGE (Figure 2.38a). While the EMSA showed nucleosome binding, it was challenging to assess the homogeneity of complex formation as the SWR1-nucleosome complex was hardly entering the gel. Consequently, I applied mass photometry which can be used to determine the molecular mass of proteins-nucleic acid complexes in solution. This method detects the reflection of light at the single molecule level from which their molecular weights can be derived. Compared to EMSA experiments, I used 25N80 which comprise extra-nucleosomal DNA on both sites of the 601 sequence. Previously, it was reported that extra-nucleosomal DNA on both ends enhances SWR1 binding [221][223]. Performing mass photometry, I detected a population of complexes with a peak at 1279 kDa, which is consistent with a fully assembled SWR1-nucleosome

complex (Figure 2.38b). Notably, a range of other, lower-molecular weight peaks indicated the presence of unbound nucleosomes as well as partially assembled SWR1 complexes. To improve complex formation, I decided to include H2A.Z/H2B histone dimers from *Chaetomium thermophilum* in the reconstitutions of the complex. I speculate that a pre-loading with the dimer is required for efficient nucleosome binding. In addition, a structure including H2A.Z/H2B enables a better understanding of the subunits coordinating the histone exchange.

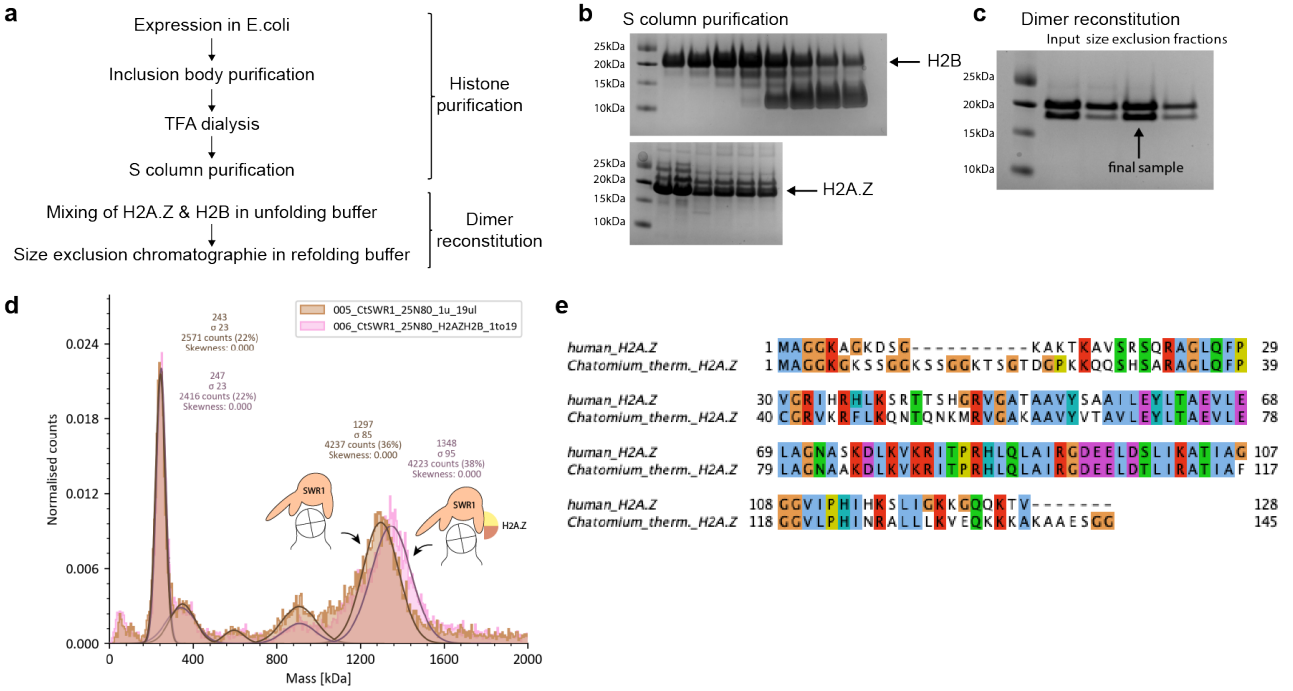


Figure 2.39: **H2A.Z/H2B dimers. a** Expression and purification strategy for *C. t.* H2A.Z and H2B. **b** SDS-PAGEs of S column fractions for H2A.Z and H2B. **c** SDS-PAGE of reconstituted H2A.Z/H2B dimers. Shown are the size exclusion fractions. **d** Mass photometry of SWR1-25N80 nucleosomes in absence and presence of *C. t.* H2A.Z/H2B dimers. **e** Sequence alignment of human and *C. t.* H2A.Z created with Jalview.

to this end, I established the recombinant expression of *C. t.* H2A.Z. Histones are generally very conserved proteins. *C. t.* H2A.Z could be easily identified by Blast® sequence searches as it shares 72 % of its amino acid sequence with its human counterpart (Figure 2.39e). Nonetheless, I am expecting that *C. t.* H2A.Z/H2B dimers have a higher affinity for *C. t.* SWR1 compared to the human dimer. I therefore produced the dimer from *C. thermophilum*. For the protein production of the H2A.Z/H2B dimers, I used the same protocol as for human histones (schematic illustration in Figure 2.39a). In brief, I purified the histones from E.coli inclusion bodies using a TFA dialysis and S column purification. I analysed all fractions from the S column step to identify the purest and most similar fractions. Of note is, that the gels showed a multitude of bands at similar molecular weight as the expected histone sizes. Especially, for H2B a second thick band at about 10 kDa was detected which could be separated by ion exchange chromatography (Figure 2.39b); mass spectrometry confirmed that this band is also full-length H2B. Next, I reconstituted H2A.Z/H2B dimers by unfolding and refolding of the histone pair. Using size exclusion chromatography I was able to isolate stoichiometric H2A.Z/H2B dimers (Figure 2.39c). Finally, I used mass photometry to probe binding of H2A.Z/H2B to the SWR1-nucleosome complex (Figure 2.39d). Indeed, I observed a new

population that was of 30 kDa corresponding to H2A.Z/H2B binding. This small shift is close to the limit of what one can distinguish by mass photometry. To provide further evidence for H2A.Z/H2B complex formation with SWR1, I analyzed the interaction also by crosslinking mass spectrometry as shown in Section 2.3.4.

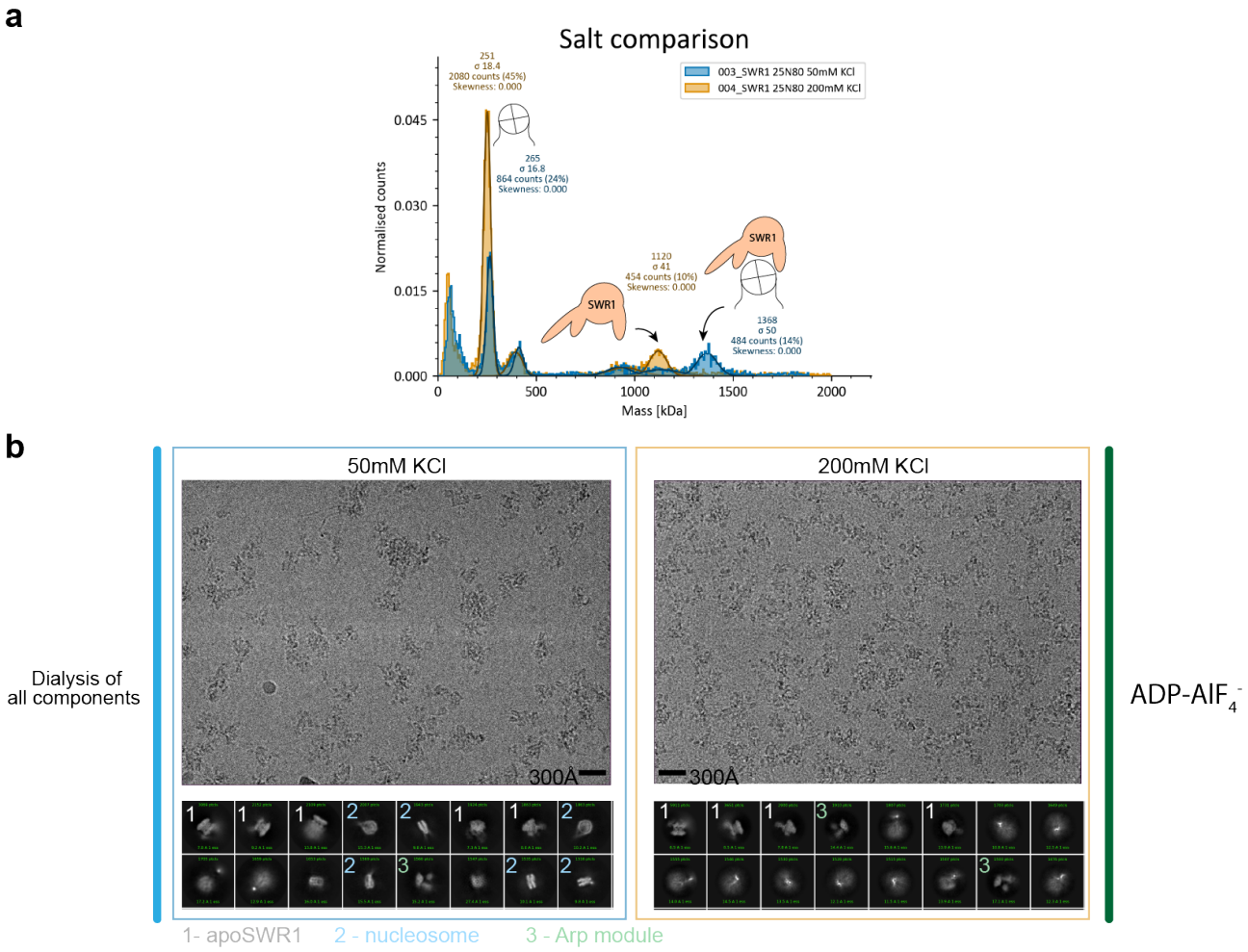


Figure 2.40: **Influence of salt concentration on complex integrity. a** Mass photometry of SWR1-25N80 complex in presence of 50 (blue curve) and 200 mM (orange curve) KCl. Only apo SWR1 observed at 200 mM KCl. **b** cryoEM samples of SWR1-25N80 at 50 (*left*) and 200 mM (*right*) KCl. Samples had been dialysed to respective salt concentration and ADP-AlF_4^- was added prior to sample vitrification. Top: micrographs, Bottom: 2D classes labeled to illustrate identity of classes.

Having probed SWR1-nucleosome binding and having the *C. t.* H2A.Z/H2B dimers at hand, I was able to prepare cryoEM grids for the structural studies of the SWR1-nucleosome complex. I first started to prepare grids by using only SWR1 and nucleosomes. MonoQ ion exchange purification provides micro-molar concentrations of SWR1, yet a high salt buffer including about 500 mM KCl. To enhance binding of SWR1 to nucleosomes, I dialysed the complex to low salt (50 mM KCl). I collected a small cryoEM dataset and analysed the data by generating first 2D averages of the picked particles. Lowering the salt, was seemingly not sufficient to stabilize the complex, as I observed only 2D classes for either apo SWR1 or nucleosomes (data not shown). In order to stabilize the complex further, I added ADP-AlF_4^- which causes a closure of the motor domain around the nucleosomal DNA (see also in Section 2.2.4). However, the addition of the ATP transition state analogue did not

result in an improvement of the sample. More extensive screening was necessary. However, the fresh preparation of SWR1 for grid preparation as well as the grid screening and data analysis are very labour intensive. To circumvent going this, I utilized mass photometry as a pre-screening method to probe the complex stability in different buffers and additives. First, I tested the effect of different salt concentrations. The mass photometer showed intact SWR1-nucleoesome complexes at 50 mM potassium chloride, while at 200 mM salt (Figure 2.40a) there was no complex detectable. Seemingly the presence of too many salt ions disrupts the electrostatic interactions of the nucleosome with SWR1. At the same time, high salt seems to be harmful for the nucleosomes themselves as I observe less nucleosome views in cryoEM at high salt concentrations (Figure 2.40b). Comparing the samples in cryoEM, revealed that the particles at 200 mM salt are well separated, while they start to form amorphous oligomers at 50 mM salt (Figure 2.40b). Particles being not well separated, prevents the usage of automated particle picking tools as the software cannot distinguish oligomerized particles from aggregates. Also, the extracted particle images would contain information of other particles preventing successful structure determination. It was hence a goal to improve the distribution of the particles. I therefore tested the complex integrity also at 100 mM and 150 mM to find the cut-off when the complex becomes destabilized. As the cut-off lays close to 100 mM potassium chloride, I decided to dialyse the complex in future grid preparations to 75 mM salt where the particle spreading on EM grids should be better while the complex is still intact.
Besides the different salt concentrations, I tested the effect of ADP-AlF_4^- on the SWR1-nucleosome complex using again mass photometry. To my surprise, I did not observe any peak for neither Apo SWR1 nor the SWR1-nucleosome complex indicating that the ADP-AlF_4^- destabilizes SWR1 (data not shown). As a consequence I prepared the next grids without the ATP transition analogue. Instead, I added H2A.Z/H2B dimers that I had produced in the meantime. Now, at 75 mM of potassium chloride, I still observed particle clusters on cryoEM images (Figure 2.41a) of copper 2/1 grids. However, the particles were nicely spread on gold grids. Despite the promising looking micrographs of the gold grid, (Figure 2.41b) no SWR1-nucleosome complexes were observed in the subsequent data analysis. As there were no nucleosome classes but only free DNA, I changed the sample preparation once more. Even if the nucleosomes are exposed to high ionic strength conditions, they should reassemble when dialysed to low salt. As this seems to have not worked in my case, I instead dialysed SWR1 bound to H2A.Z/H2B to low salt prior to assembling the SWR1-Z/B-nucleosome complex (Figure 2.41c-f). In this condition, the particles were well distributed on copper grids but not on grids coated with gold. The pre-dialysis had no positive effective on neither higher nucleosome numbers nor the stability of the SWR1-nucleosome complex (Figure 2.41c,d) in this case. Besides the presented data, I tested Tween20 as detergent replacing the otherwise used β-octylglycoside and I froze the sample on copper 2/1 grids with a 2 nm carbon support. The Tween20 grids did not differ from the ones prepared with β-octylglycoside and the particles were sticking even more together on the carbon coated grids compared to ones without the carbon support layer.

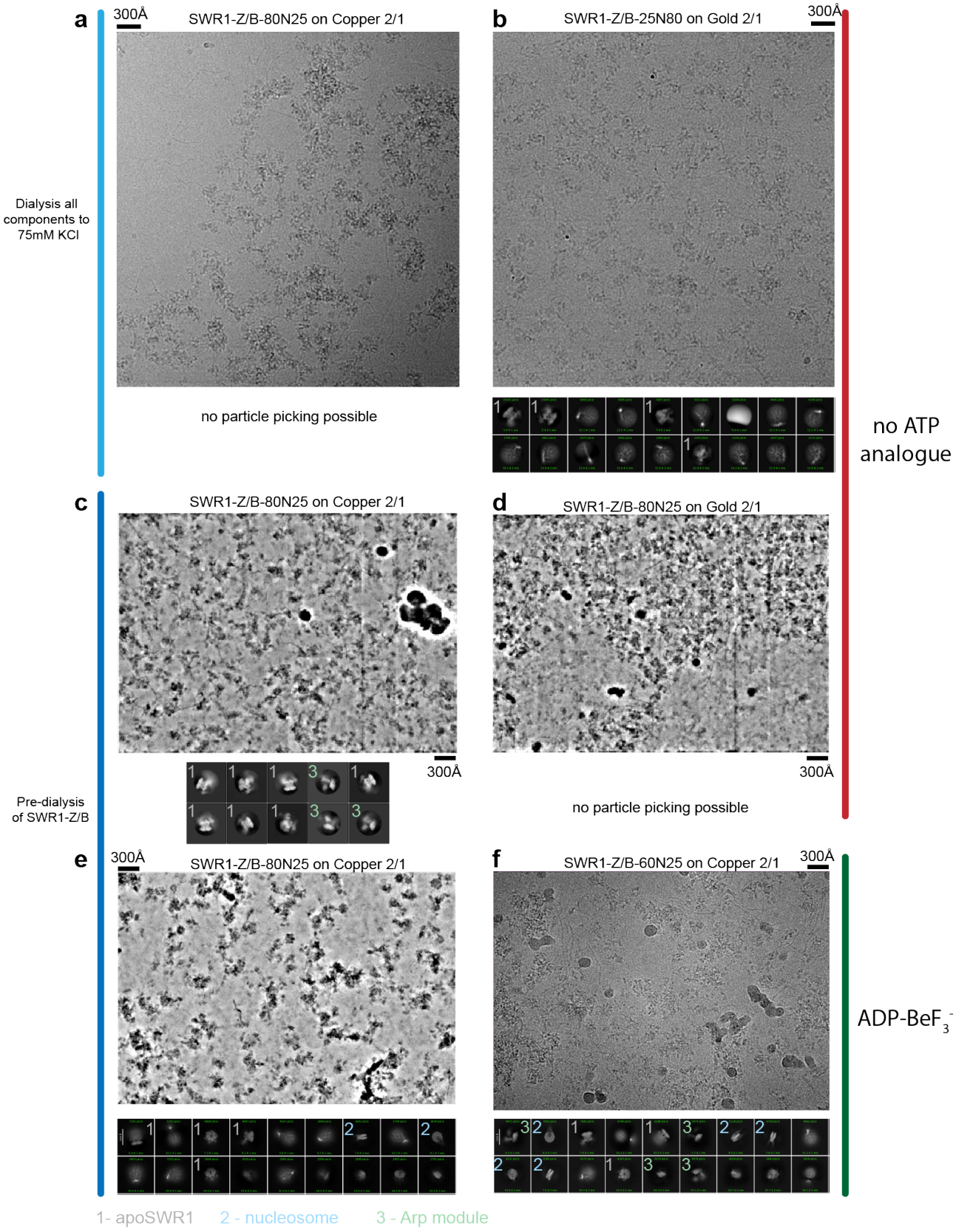


Figure 2.41: **Screening for SWR1-nucleosome complex. a-f** EM images and 2D classes are shown where available. Legend for 2D classes at the very bottom. SWR1-nucleosome-H2A.Z/H2B (Z/B) complex was either dialysed to 75 mM KCl or SWR1/H2A.Z/H2B was dialysed alone prior to complex formation. Grids were prepared in absence of any ATP analogues or in presence of ADP-BeF_3^-.

In a last trial, I mixed the SWR1-Z/B-nucleosome complex with the toxic ADP-BeF_3^- which had also been used in the Willhoft publication [142] as ATP analogue. In this condition, the particles were still sticky (Figure 2.41e,f) but separated enough to allow for its analysis by cryoEM. In this case, I saw a positive effect of the pre-dialysis of SWR1 as I observed multiple nucleosome classes indicating that they are intact throughout the sample preparation. Yet, the supplementation of ADP-BeF_3^- had not the desired effect of stabilizing the SWR1-nucleosome interaction. Therefore, separate 2D classes of apo SWR1 and nucleosomes were observed. In summary, I used a combination of mass photometry and cryoEM screening to identify conditions that stabilize the SWR1-nucleosome complex. Despite the extensive screening of multiple conditions, I could not stabilize the SWR1-nucleosome complex for cryoEM analysis. The conditions included low salt buffers which enhance electrostatic interactions or the usage of detergent to prevent association with the air-water interface. The latter was described to have a degrading effect on proteins through its hydrophobic characteristic [262]. At the same time, I tested gold and copper grids with and withough extra carbon support which can have a positive effect on an even particle distribution [299]. Finally, also the usage of ATP analogues did not prevent the complex from falling apart under conditions of cryoEM grid preparation.
Possibly, the used nucleosome with its non-native and generic 601 sequence and flanking DNA, might just not be a substrate good enough for SWR1 binding leading to its destabilization on cryoEM grids. Instead, I speculate that SWR1 has a preference for certain DNA sequences which guides SWR1 *in vivo* to +1 nucleosomes as well as to histone acetylation. Collectively, these factors may contribute to its substrate specificity. Preferred DNA sequences of SWR1 can be identified by genome wide studies. I therefore established a collaboration with the Korber lab to perform genome wide studies with yeast SWR1 for the identification of promoter DNA sequences that SWR1 preferably targets.
Complementary, I propose to test a nucleosome variant that harbors a single base mutation close to the binding site of the SWR1 motor domain at SHL+2. This mutant was used for structural studies of the Chd1 chromatin remodeler that binds the nucleosome at the same location as SWR1 [300]. Nucleosome reconstitutions of the mutant were already tested but unsuccessful to date. In addition, one can test acetylated nucleosomes that enhance binding of the Bdf1 and Yaf9 components of SWR1 [211][209]. Finally, the cryoEM screening that I performed in the context of this thesis for the SWR1-nucleosome complex will be a good starting point for future studies tackling this challenging task.

3 Discussion

3.1 Hexasomes as another layer of genome regulation

I presented the structure of *C. t.* INO80 bound to a hexasome that lacks the proximal copy of the H2A/H2B dimer in relation to linker DNA (see Section 2.2). The absence of the proximal H2A leads to a restructuring of the nearby copies of H3 and H4 resulting in a β-α switch of the C-terminal H4' tail while leaving the H3'αN helix unstructured. On top of that, the H3-H4 interface gets exposed which creates a novel binding platform and is specifically recognized by INO80 via multiple contact sites. Unwrapping of 40 bp of DNA from the hexasomal histone core results in a spin-rotated binding mode of INO80. Due to spin-rotation, INO80 binds SHL-3 in context of hexasomes. Thereby, INO80 has a dual binding model: the motor lobes sit at SHL-6 on nucleosomes wheras they recognize SHL-3 in the context of hexasomes. Spin-rotation of INO80 by 145° enables INO80 to adapt to the changed histone surface. As a consequence, the Arp5-foot helix binds to the exposed H4α2 helix. On nucleosomes, on the contrary, the Arp5-foot binds the acidic patch of the proximal H2A/H2B on nucleosomes. As a consequence, hexasome remodeling becomes independent on the acidic patch as I showed by INO80 sliding assays (Section 2.2.8). Mutation of the Arp5-foot leads to an increased sliding activity showing that the Arp5-foot has a dual function: While it is essential for successful nucleosome sliding, it has a regulatory function in the context of hexasomes. Lastly, I showed that the nucleosome is the preferred substrate of INO80 when an equal amount of 60 or 80 bp of extra-nucleosmoal DNA is available (in results Section 2.2.10). However, if the amount of extra-nucleosomal DNA is limited to about 20 bp, a hexasome becomes the preferred substrate compared to nucleosomes. The reason lays in the unwrapping of an additional 40 bp of DNA upon hexasome formation. This scenario occurs for example in yeast gene bodies where nucleosomes are tightly packed and spaced by 18 bp of linker DNA [283]. However, many chromatin factors including INO80 require longer stretches of linker DNA [192][194] for their respective activities on chromatin (Figure 3.1a). In this regard, INO80 remodeles nucleosomes if at least 60 bp of linker DNA are available [194]. Hexasomes can be generated through ATP-dependent nucleosome remodeling [69][55][301] or through active gene transcription by RNA Polymerase II [58][55][62]. RNA Pol II removes in context of transcription elongation the distal copy of H2A/H2B in relation to the nucleosome free region around gene promoters [63]. Here, I describe an example where the +1 nucleosome becomes a hexasome. I refer to the promoter region as nucleosome-free region (NFR) while the DNA in between the +1 and +2 nucleosome is the linker. Formation of a hexasome leads to the unwrapping of about 40 bp of DNA increasing the linker to 64 bp in between the newly generated hexasome and the +2 nucleosome (Figure 3.1a). Now, INO80 can bind and remodel the formed hexasome. INO80 is tethered through its Arp module to the linker DNA, while the INO80 core complex recognizes the hexasome (Figure 3.1b). The linker DNA represents the entry DNA which gets pulled into the hexasome in 1 bp steps per ATP hydrolysis [70]. Thereby the hexasome is slid towards the +2 nucleosome. Overall, this scenario is consistent with RNA Polymerase II removing the distal H2A/H2B copy from nucleosomes (Figure 3.1a). Looking again at the +1 and +2 nucleosomes downstream of promoter DNA, this means that the distal H2A/H2B copy is removed in relation to the NFR. INO80 recognizes the +1 hexasome as the proximal H2A/H2B is missing in relation to the linker DNA that separates the +1 hexasome and +2 nucleosome (Figure 3.1c). Why the hexasomes need to be slid by INO80 re-

mains to be investigated. However, it was shown that hexasomes exist *in vivo* thereby contributing to the genome-wide nucleosome pool [60][67] making them a target for energy-driven remodeling. Of note is, that hexasome formation with its underlying DNA unwrapping and H3/H4 exposure change the chromatin landscape. Spin-rotation alters the orientation of histone modifications and thereby how they are read. In this manner, hexasomes represent another layer of gene regulation.

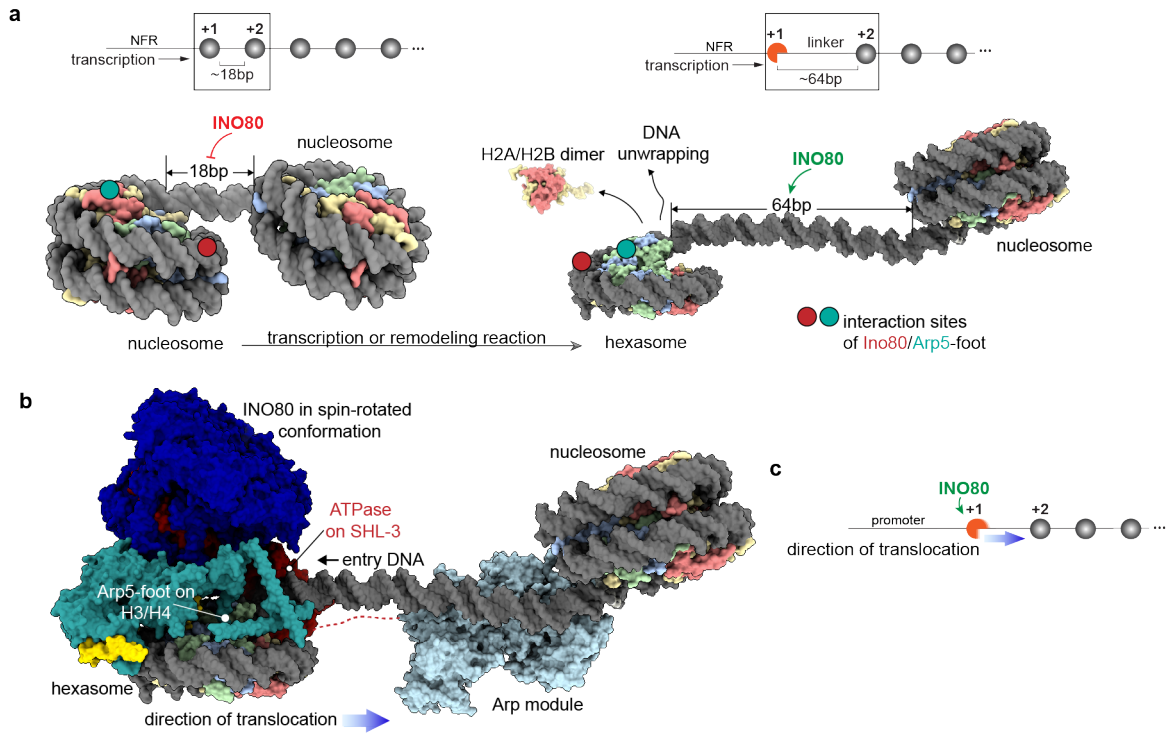


Figure 3.1: **Hexasome remodeling in context of gene transcription a** Two nucleosomes are separated by app. 18 bp within yeast gene bodies. As 18 bp of extra-nucleosomal DNA are not enough for INO80 to bind, INO80 is not active on these sites. Upon formation of a hexasome through transcription or other remodeling reactions, 64 bp become available for INO80 to bind and remodel the hexasome. Min Zhanag generated di-nucleosome models by combining two copies of PDB 7OHC separated by a 18 bp linker. Hexasome-nucleosome dimer was prepared from 7OHC and 8AV6. Interaction sites of Arp5-foot and Ino80 motor lobes are labeled by cyan and red dots respectively. **b** INO80 bound to a hexasome and merged with the published 8AV6 INO80 nucleosome structure. The unwrapped DNA is bound by the INO80 Arp module (light blue) which connects through the post-HSA (red dotted line) to the INO80 core complex. The hexasome gets translocated towards the nucleosome as the linker DNA gets pulled into the hexasome by the ATPase. **c** Illustration of nucleosomal array downstream of promoter DNA. RNA polymerase II removes the distal H2A/H2B copy in relation to nucleosome free region. This is consistent with INO80 binding to the +1 nucleosome as shown in **b** and translocation of the hexasome towards the +2 nucleosome.

3.2 Hexasomes as a possible intermediate of the INO80 sliding mechanism

The presented structural findings are supported by another study of the Narlikar and Cheng labs (Hsieh et al.) [67] which was published back to back with the here presented results [281]. The

Narlikar lab used endogenously purified INO80 from *Saccharomyces cerevisiae* and reconstituted likewise to our study 0H80-INO80 complexes for cryoEM studies. They observed three major classes which mainly differ in the positioning of the INO80 motor domain on the hexasomal DNA: SHL-2, SHL-2.5 and SHL-3. While I only observed binding to SHL-3, the main class of INO80 in Hiseh et al. [67] is bound to SHL-2. Notably, SHL-2 is a binding hub for ATP-dependent chromatin remodelers including CHD [72][302][303], ISWI [107][304][106], RSC [122][305][306], Snf2 ATPase [307] and SWR1 [142] which is structurally related to INO80. The same lab described that hexasomes are the preferred substrates of INO80 [67] when comparing sliding activity of *S. c.* INO80 on nucleosomal and hexasomal substrates. However, in my hands nucleosomes are the preferred substrates of *C. t.* and *S. c.* INO80 when comparing nucleosomes and hexasomes with the same amount of linker DNA and in competition assays. This difference can possibly occur as the Narlikar lab was working with endogenously purified material while I was using recombinantly purified INO80. On top of that, I compared nucleosomes and hexasomes with the same linker lengths taking into account the 40 bp of DNA unwrapping in context of the hexasome. Therefore, I could exclude the possibility, that it is rather the linker length driving an increase in sliding activity than the composition of the histone core. Nonetheless, the combined studies of the Narlikar lab and our lab issue the possibility that hexasomes are an intermediate of the INO80 sliding reaction [67]. This would mean that INO80 starts DNA translocation on a nucleosome to remodel the nucleosome into a hexasome. Subsequently, INO80 spin-rotates to bind the formed hexasome at SHL-2/-3 and starts hexasome sliding which is more efficient due to the reduced amount of DNA-histone contacts compared to a nucleosome. This hypothesis is supported by FRET studies showing an initial, ATP-dependent pause step [194] which is presumably equal to INO80 spin rotation to the freshly formed hexasomal substrate consuming thereby ATP. At the same time, Hsieh et al. [67] detect an increasing amount of DNA unwrapping at the entry site in the transition from SHL-3 to SHL-2 binding. A similar detachment of the entry DNA off the histone core is observed by FRET studies [76] indicating that the unwrapping of DNA is required for successful nucleosome remodeling. One fact contradicting the hypothesis of a hexasome intermediate is shown in the Hsieh paper [67] itself: the usage of a dimer dislodgement mutant, retained INO80 activity. Overall, the exact sliding mechanism of INO80 remains elusive and requires further investigation. For this purpose, the reaction of INO80 needs to be followed over a multitude of ATP hydrolysis steps. Studies in the presence of different ATP analogues (as compared in Section 2.2.4) is not sufficient as INO80 functions in larger step sizes, accumulating about 10 single base pair translocation steps before a net DNA translocation around the histone core occurs [76]. In its first DNA translocation steps, INO80 moves 10 bp around the nucleosomal histone core towards the Arp5 grip, while the INO80 motor sits close to the DNA entry site at SHL-6. Thereby, the interface close to the proximal H2A/H2B is loosened up, suggesting that a dislodgement of H2A/H2B is in principle possible [77]. The Arp5 grip prevents a net DNA translocation at this point, building up a DNA strain which leads to DNA underwinding and possibly DNA loop formation in between the anchored INO80 motor domain and the Arp5 grip (also suggested in [77]). Backtracking is prevented in this scenario by the upstream sitting Arp module which senses the entry DNA. In order to shed light on the remodeling mechanism of INO80, I propose to trap the suggested loop intermediate in a roadblock study by crosslinking the DNA close to the Arp5 grip to the histone core. By that, the loop intermediate is accumulated as the DNA propagation through the crosslinked histone core is not possible. I started to develop a crosslinking and sample preparation strategy in my PhD work. In brief, I suggest to crosslink a histone cysteine mutant to an NHS ester modified nucleotide using a bis-maleimide crosslinker. In more detail, H3 residue D82 and H4 residue T81 are close to the Arp5 grip and the DNA and can be mutated to a cysteine as they do not directly contribute to DNA binding. I planned to have the NHS ester at 601 position 98 (24 bp downstream of dyad) as this residue is close enough to the introduced H3 and H4 cysteines. In a pilot experiment, bis-maleimide

crosslinkers of different lenghts (BMOE (bis-maleimidoethane) 8 Å crosslinking length, BMB (bis-maleimidobutane) 10.9 Å, BMH (bis-maleimidohexane) 13 Å) (according to Thermofisher) can be tested in combination with H3-D82C or H4-T81C containing nucleosomes. Unspecific crosslinking can be further prevented by mutating the only inherently present cysteine residue in H3 (C111) to an alanine.

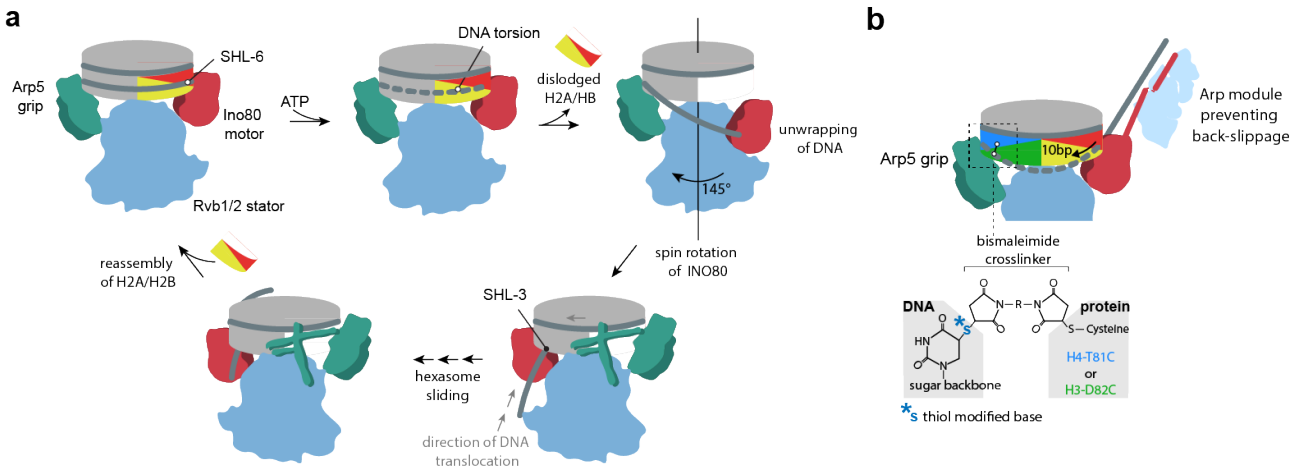


Figure 3.2: **a Proposed mechanism for INO80 nucleosome sliding**. INO80 translocates DNA towards the Arp5 grip which blocks a net DNA translocation. The accumulated DNA torsion results in hexasome formation. INO80 then spin-rotates to bind the formed hexasome at SHL-3 and slide the DNA along the hexasomal histone core. Lastly, the second H2A/H2B copy is reintroduced. **b Roadblock study**. The histone core is crosslinked to the DNA close to the Arp5 grip (upper part, dotted square). INO80 translocates about 10 bp towards the crosslink while the Arp module prevents backslippage of the formed DNA loop. Crosslinking strategy on the bottom shows that a thiol modified base close to the Arp5 grip is crosslinked via a bismaleimide crosslinker to either H4-T81C or H3-D82C.

In the final setup one of the two histone mutants is used with the identified optimal crosslinking reagents. The reaction is initiated with ATP and cryoEM grids are prepared after a small waiting time facilitating the formation of the INO80-nucleosome loop complex. Modern data processing software can be applied to determine structures of different reaction states. With this setup, a possible hexasome intermediate can be captured and imaged by cryo electron microscopy.
Complementary, I suggest using FRET studies labelling the proximal H2A/H2B copy and the DNA in between the Ino80 motor and the Arp5 grip. In addition, a simplified time-resolved cryoEM procedure can be used where one mixes the INO80-nucleosome complex with ATP and a non-hydrolysable ATP analogue. Thereby, the reaction gets statistically trapped whenever a copy of the non-hydrolysable ATP is bound to INO80. Overall, the introduced procedures are expected to answer how the INO80 motor domain and its accessory subunits are concerted to accomplish nucleosome sliding. Finally, one can achieve an understanding what drives the remodelers' substrate specificity and what determines the reaction outcome. The latter point is of special interest considering the similar architecture of the INO80 and SWR1 remodelers that yet catalyse distinct reactions.

3.3 Hexasomes are a substrate of other multi-subunit chromatin remodelers

In light of the dual substrate specificity of INO80, I asked whether other multi-subunit chromatin remodelers, are capable of acting on hexasomal substrates. Modeling of SWI/SNF and RSC onto a hexasome [306][121] suggests that these remodelers could also undergo a 145° spin-rotation likewise to INO80 (presented in Section 2.2.7). This places the Sfh1 finger helix of RSC and the Snf5 subunit of SWI/SNF onto the exposed H3/H4. Thereby they recognize the H3α1L1 elbow – a feature that is also recognized by the foot helix of the INO80 subunit Arp5. Interestingly, the spin-rotated movemenet of RSC and SWI/SNF places their motor domains onto SHL-6 suggesting that these multi-subunit remodelers also have a dual binding mode on nucleosome-like particles. In addition, the unwrapped DNA is bound by the accessory subunits of RSC and SWI/SNF. In summary, the binding modes to hexasomes are analogous among the different multi-subunit chromatin remodelers. Evidence for hexasome binding by RSC was found recently as it was shown that RSC dislodges H2A-H2B dimers in context of histone acetylation [69]. A similar mode of action is possible for INO80 as illustrated in the previous paragraph. A H2A/H2B ejection seems plausible as INO80 pumps DNA along the H2A/H2B interface in context of nucleosomes. However, this hypothesis needs further investigation by a combination of *in vitro* structural characterisation as well as studies in context of transcription, other remodelers and histone chaperones which are likely required for hexasome generation and re-installation of the nucleosome landscape.

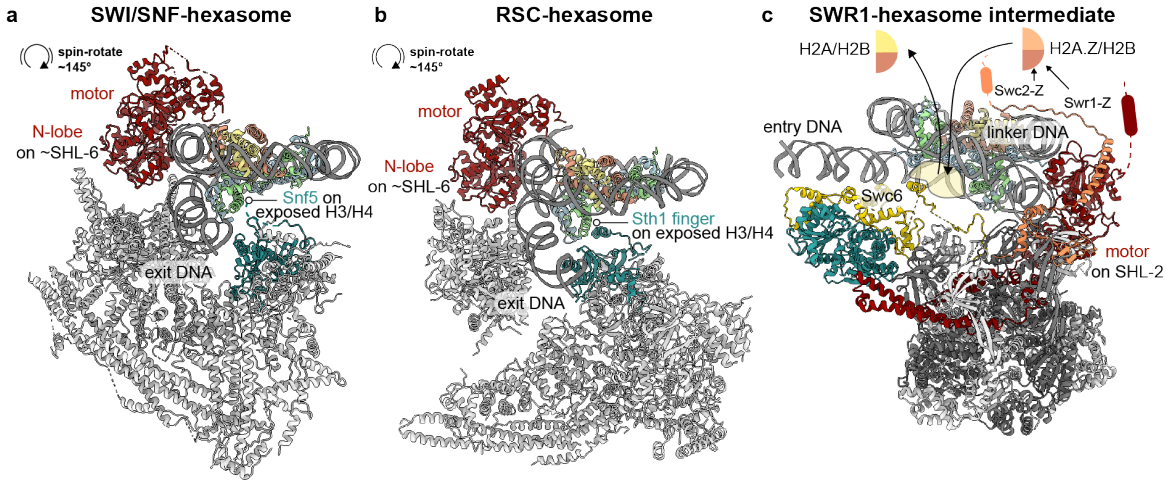


Figure 3.3: **Hexasomes in context of other chromatin remodelers.** **a** Model of SWI/SNF (PDB 6UXW) on a hexasome. SWI/SNF can accommodate a hexasome if undergoing a 145° rotation likewise to INO80-hexasome binding. This places the Snf5 subunit close to the exposed H3/H4 interface. **b** Similar representation for the RSC complex (based on PDB 6TDA). Here the predicted Sth1 finger would interact with the exposed H3/H4. In both cases the scaffolding subuntis (in gray) contact the extra-nucleosomal DNA at the DNA exit site. **c** Hexasome model of SWR1 based on PDB 6GEJ. The proximal H2A/H2B is replaced in relation to the motor (marked by yellow sphere underneath the histone core). The Swc6 subunit is close by. The outgoing H2A/H2B is bound by Swc5 while the incoming H2A.Z/H2B is delivered by the Swr1-Z and Swc2-Z domains.

Finally, the question remains if SWR1 binds and acts on hexasomes. In regard to the stepwise histone exchange reaction of SWR1 [235], I rather suggest that hexasomes present short-lived intermediates of the underlying reaction. In more detail, the SWR1 complex anchors the nucleosomal substrate at

a fixed position through at least three components as captured in the *S. c.* SWR1 structure [142]: the Swr1 motor lobes, Swc6/Arp6, and the Swc2 subunit that interacts with the acidic patch of the distal H2A/H2B. Potentially the Arp module with its Yaf9 component represents a fourth anchor point. The motor domain destabilizes histone-DNA contacts as shown for example by single molecule studies [229] leading to the dislodgement of the motor proximal H2A/H2B copy [142][214]. This is the moment where a hexasome might exist. I propose that SWR1 does not undergo a spin-rotation for its binding to the hexasome but rather stays bound to SHL-2 – fixed at the nucleosome through the described interactions points. The dislodged H2A/H2B copy stays bound to the nucleosome for about 10 s [229] through its Swc5 interaction [226]. Next, the H2A.Z/H2B dimer gets delivered to the previous H2A/H2B binding site by the Swc2-Z and Swr1-Z domains [233][234]. The Swc2-Z and Swr1-Z are located at unstructured regions at the N-terminus of Swc2 and in between the HSA and motor domain of Swr1. Interestingly, they are potentially closer to the distal instead of the proximal H2A/H2B copy. Therefore, it remains unclear how Swc2 and Swr1 deliver the H2A.Z/H2B copy. For this purpose, I suggest to solve a cryoEM structure of SWR1 bound to a hexasome and in the presence of H2A.Z/H2B dimers. In order to verify a SWR1 created hexasome intermediate, dedicated FRET studies could be performed.

3.4 New perspectives on the SWR1 architecture

Combining previous and the presented structural studies with known biochemical data, provides a new perspective on the SWR1 complex summarized in Figure3.4. The Core module and its binding mode was already well described in [142], however lacking the Swc3 core subunit. Swc3 is anchored to the Core module by the degenerated Swc2 zinc finger positioning the Swc3/Swc2 arm close to Arp6 (as presented in Section 2.5). Thereby, Swc3 is positioned away from the nucleosome which is in line with biochemcial data showing that Swc3 does not contribute to nucleosme binding [228]. The function of Swc3 in yeast and fungi remains however unclear. A human homologue of Swc3 in the SRCAP complex is not characterized indicating that Swc3 is dispensable and was hence lost in the evolutionary advancement of the SWR1 complex. The Arp module of SWR1 is flexibly tethered to the Core module through the Swr1 linker region that connects the Swr1 motor domain with its HSA helix. The length of the Swr1 linker varies among different species. Further, I showed by single molecule analysis that the mean distance of the Core and Arp modules is 173 Å which was calculated based on the center of mass of the two modules. The Arp module displays the shape of the letter U: Its subunits assemble onto the Swr1 HSA helix which is split in *C. t.* into the helices α1 and α2. The core of the Arp module is composed of the HSA α1 helix, Arp4, actin^{A} and the Swc4 2W hairpin domain. In line with [221], I identified a second actin molecule within the Arp module (actin^{B}). However, actin^{B} does not dimerize with its neighboring actin molecule as one would assume when comparing the SWR1 Arp module to the INO80 Arp module. In more detail, the two actin binding sites on the Swr1 HSA helix are too far apart thus preventing dimerization. The two HSA helical parts that also separate the two actin molecules, are connected by a linker that causes a kink. Yet, the two HSA helices might be straightened upon DNA binding which would however still not lead to the dimerization of the actins. Furthermore, crosslinking mass spectrometry data places the Swc5 histone chaperone between the actin^{B} on HSA α2 and the Rvb1/Rvb2 hetero-hexamer. AlphaFold predicts Swc5 to be mostly disordered. Swc5 functions as a histone chaperone for the replaced H2A/H2B dimer [308][309] and its localization on the Rvb1/Rvb2 is therefore plausible. Further, my integrated structural biology study could identify Yaf9 as an integral component of the Arp module. The joint formed by Swc4 is located adjacent to Arp4 and introduces a 90° kink into the Arp module. Thereby the Arp module foot formed by Yaf9, the Swc4 C-terminus and the Swr1 N-terminal region,

can fold backwards towards the Core module of SWR1. In context of nucleosome binding, the 90° kink of the Arp module could enable its binding to the histone core. In more detail, Yaf9 would be placed on top of H3/H4 which is in agreement with its recorded binding of acetylated histone tails of H3K27ac and H4K9ac [209][210]. In addition, Bdf1 recognizes other acetylated histone tails however in this case through its two bromo domains [211]. Yet, Bdf1 was absent in the purified *C. t.* SWR1 complex due to the missing Swc7 homologue. Nonetheless, biochemical data [228] places Bdf1 at the very N-terminus of SWR1 which would be presumably close to Yaf9 and hence also close to the nucleosome. Future studies may include different patterns of histone acetylations to capture SWR1 in different nucleosome bound states.

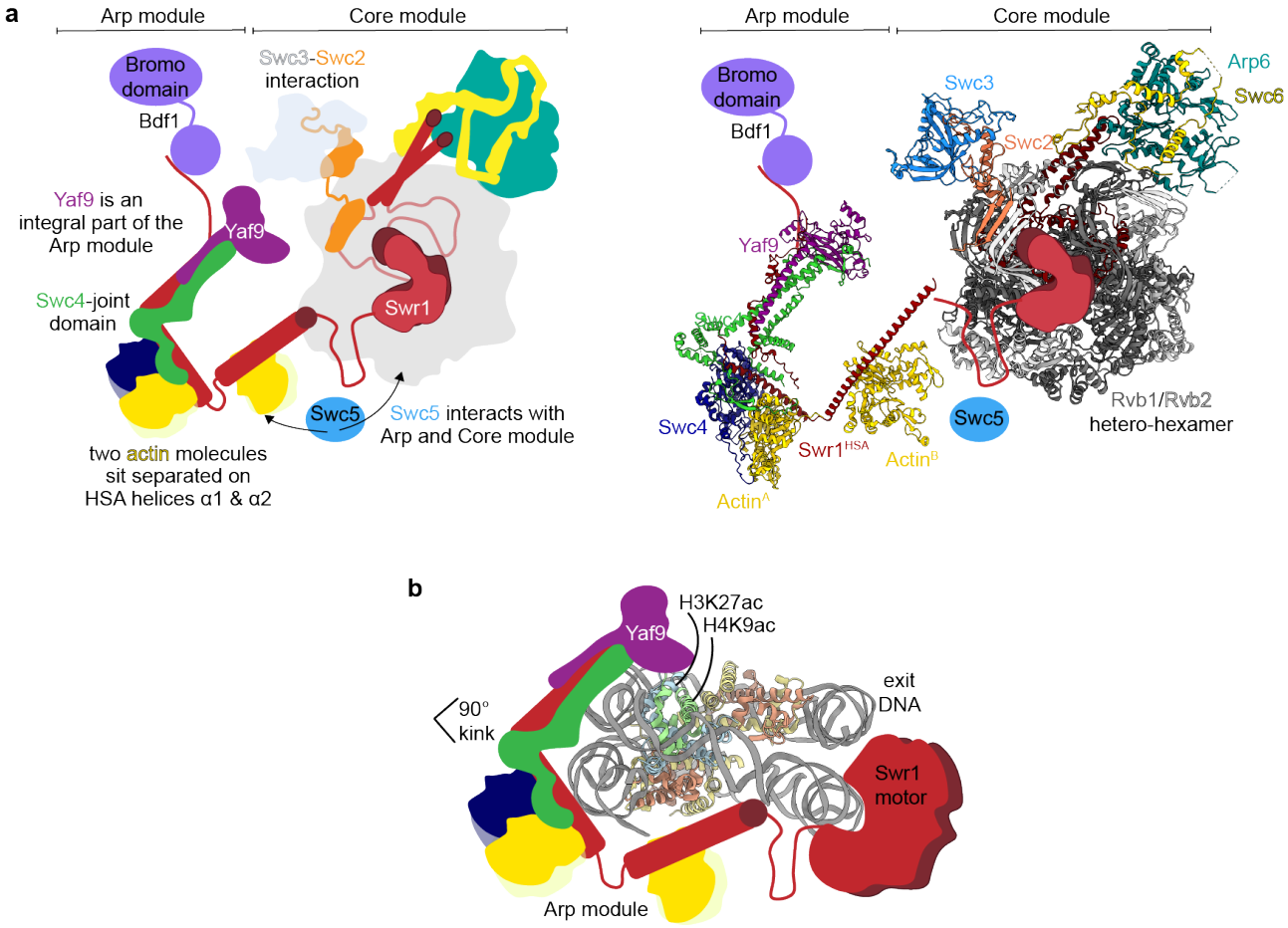


Figure 3.4: **SWR1 architecture based on this thesis.** **a** *Left:* Scheme of SWR1 architecture based on its arrangement into the Arp and Core module. *Right:* Same architecture but with structures from this thesis. **b** Model of the Arp module binding to a nucleosome. The 90° kink in the Arp module allows Yaf9 to fold back onto the nucleosome and interact with H3K27ac and H4K9ac.

3.5 SRCAP and SWR1 share a common Arp module architecture

While writing my PhD thesis, Yu et al. [143] published a cryoEM structure of SRCAP in complex with a mono-nucleosome which is the human homologue of the yeast and *C. t.* SWR1 complexes. Unlike yeast, no homologues for Bdf1, Swc3 and Swc5 were present in the used human SRCAP construct; partly because they could not be identified. While Bdf1 and Swc3 were shown to be not essential

for ATPase and histone exchange activity in yeast, Swc5 is required with its function as histone chaperon for successful H2A.Z incorporation [221]. In accordance, no histone exchange activity was observed for the recombinantly purified SRCAP complex [143]. Therefore, the structures obtained need to be considered to be a non-active state of the SRCAP complex. In addition, the authors utilized crosslinking to stabilize the complex for cryoEM studies which can cause unnatural protein-protein and protein-DNA interactions. Nonetheless, their structural findings are of interest to the field and I will use them to compare my results on *C. t.* SWR1 and its Arp module. Besides the core-subunits of SRCAP, the authors captured the Arp module consisting of a long HSA helix, two copies of actin as well as DMAP1 (Swc4) and ACTL6A (Arp4). In addition, they observed a low-resolution density at the tip of the pre-HSA helix that they interpret as YEATS4 which resembles Yaf9 in yeast.

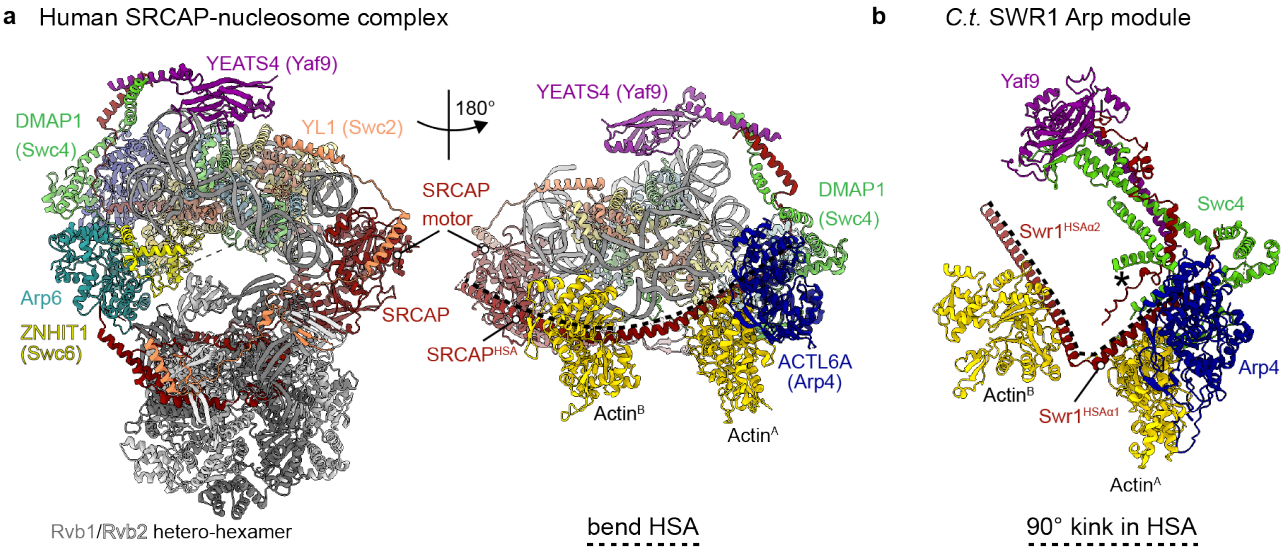


Figure 3.5: **Comparison of human SRCAP and *C. t.* SWR1 Arp module. a** SRCAP structure bound to a mono-nucleosome based on PDB 8X19 [143]. On the right, close up on the Arp module. Structure is turned by 180°. HSA helix, labeled by dashed black line, is bend. **b** Similar view of *C. t.* Arp module. Homologue subunits are colored likewise to human SRCAP. HSA shows a kink of 90°.

Looking at the architecture of the human Arp module, the long HSA helix protrudes from the motor lobes and binds about 80 Å apart from the motor actin^{B} that does not interact with any other of the SRCAP subunits. The rest of the Arp module assembles as a quadruple complex onto the HSA helix – from the C to the N-terminus – actin^{A}, ACTL6A and DMAP1. In more detail, actin^{A} and ACTL6A form a dimer in between which the DMAP1 subunit is anchored. DMAP1 continues to interact with SRCAP N-terminal to the HSA helix. Overall, the architecture of the human Arp module resembles the structural findings of *C. t.* SWR1 presented in this thesis (Section 2.5) (comparison in Figure 3.5). Especially the core of the Arp module around Arp4, actin, Swc4 and the HSA resemble each other. However, the HSA helix of human SRCAP is continuous but slightly bent while the *C. t.* HSA is segmented by a short loop region into the HSA helices α1 and α2. Nevertheless, the kink at the HSA linker residues could form a continuous helix in the presence of DNA likewise to the SRCAP structure. The presence of a second copy of actin (actin^{B}) in the SRCAP structure is in agreement with my structural characterization of the *C. t.* Arp module from Section 2.5. The SDS-PAGE analysis of the recombinant SRCAP purification indicates one copy of actin similarly to my *C. t.* SWR1 purifications. Conclusively, either the intensity of actin is misleading on SDS-PAGE or a sub-fraction of SWR1/SRCAP complexes exist with only one copy of actin.

The actin^{B} is anchored by a seemingly conserved tryptophan residue likewise to the *C. t.* structure, while the tryptophan anchor of actin^{A} is located at an altered position laying more towards the N-terminus of the HSA. Furthermore, the DMAP1 (Swc4) subunit facilitates a similar kink of the Arp module's foot part. This places YEATS4 on top of the H3/H4 interface where it is ready to interact with acetylated histone residues as I suggested in Figure 3.4. While the densities for Yaf9 are of low resolution in both the SRCAP and SWR1 structures, I could confirm by co-purification of the Yaf9-Arp module complex that Yaf9 is an integral part of the underlying module (see Section 2.4.2).

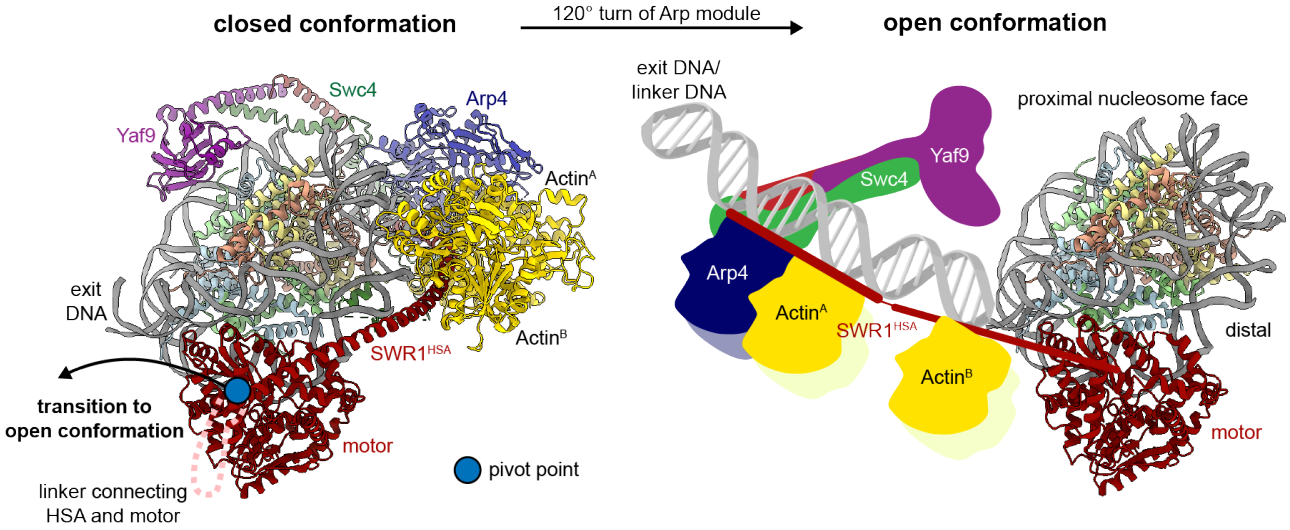


Figure 3.6: **Molecular switch of SWR1 Arp module from closed to open conformation.** SRCAP structure on the left is in closed conformation with the Arp module placed on nucleosomal DNA. Yaf9 folds back on top of the nucleosome to interact with acetylated H3/H4 histone tails. SWR1 can presumably switch similar to the RSC complex into an open conformation in which the Arp module is turned by about 120° around the pivot point. The latter sits at the C-terminal tip of the HSA helix. As a consequence, the Arp module can bind to the close by linker DNA. Proximal and distal faces of the nucleosome are labeled in relation to linker DNA. The presented structures are based on PDB 8X19.

On a global scope, the Arp module in human SRCAP extends from the motor lobes that sit at SHL +2 and folds alongside the nucleosomal DNA. The Arp module folds against the proximal gyre of DNA in relation to the motor lobes. Thereby, the Arp module is placed close to the Rvb1/2 hetero-hexamer and the Arp6/ZNHIT1 (Swc6 homolog) dimer. The pre-HSA-DMAP1-YEATS4 subunits extend from the proximal gyre of DNA up to the distal one placing YEATS4 at the distal histone core facilitating its interaction with acetylated histone tails [212]. Furthermore, the authors present structures in three distinct nucleotide states of the motor domain: apo-SRCAP, ADP-BeF_x and ADP bound. The location of the Arp module on the nucleosome remains the same while the Rvb1/2 hetero-hexamer and the Swc6/Arp6 homologue undergo conformational changes that presumably facilitate histone exchange. Accordingly the authors suggest, that the Arp module promotes through its anchoring role that the ATP hydrolysis is translated into H2A/H2B destabilisation and overall histone exchange. Further, the Swc6 homologue ZNHIT1 was identified as key component for histone exchange as it contacts the proximal H2A that will be replaced by H2A.Z. The Arp module is in close proximity. A similar location is possible in *C. t.* SWR1 as I observe crosslinks between Swc6 and the Arp module subunit Swc4. On top of that, crosslinking data presented in this thesis, showed interactions between actin and Rvb2 as well as of Swc5 contacting actin and Rvb1. Interestingly, this would place

Swc5 close to the proximal face of the nucleosome where the H2A/H2B dimer is removed and taken over by Swc5 [309][308]. Recent structural data resolved Swc5 however on the opposite site of the nucleosome which would be distal to the Swr1 motor and proximal to linker DNA resembling for example promoter DNA [227]. Of interest is, that the authors showed that an interaction of Swc5 with the acidic patch on H2A/H2B is essential for *S. c.* SWR1 to perform its histone exchange activity. However, they present a structure of solely Swc5 and only a structure of the full SWR1 complex including Swc5 will help to understand the mechanism of histone exchange.

Looking at the position of the Arp module in SRCAP, it becomes apparent that the complex was captured in a closed conformation. This places the Arp module at an unusual position as it is not interacting with linker DNA but the nucleosomal DNA. In contrast, INO80 was captured in an open conformation [150]. Here, the Arp module consists of an HSA helix, Arp4, one copy of actin, Arp8 and Ies4 which is comparable to SWR1/SRCAP as discussed in Section 2.4.2. Yet, these subunits are placed upstream of the Ino80 motor domain onto the entry DNA which resembles linker DNA for example at gene promoter sites. As a consequence, INO80 can read out DNA shape features of the linker DNA, function as a ruler for nucleosome spacing and couple ATP hydrolysis to nucleosome sliding [167][192]. Also in case of SWR1, interactions with linker DNA are reported by biochemical studies and even more so, linker length next to the nucleosomal DNA stimulates SWR1 activity [221][223] and is scanned by Swc2 [213]. Yet, linker DNA interactions of SRCAP are absent and the presented structure therefore fails to provide a molecular basis for the biochemical data. More plausible would be an interaction of the SRCAP/SWR1 Arp modules with the linker DNA likewise to INO80 which would require a switch to an open conformation. The literature provides evidence that INO80 and SWR1 can be captured in open and closed conformations bound to a nucleosome [187][189]. The underlying EM data is however of low resolution. Nonetheless, higher resolution data exists for the RSC complex from the SWI/SNF family [310] in which two distinct conformational states of the Arp module were resolved. Its motor domain Sth1 connects through an HSA helix to an Arp module of an altered composition: it contains the actin-related proteins Arp7 and Arp9 as well as the fungi-specific Rtt102 subunit. The Arp module is rotated through a pivot point on the motor domain by 120° upon assembly of the full RSC complex [310]. Taken together, I propose a similar mechanism for SWR1 and SRCAP as depicted in Figure 3.6. The structure of the inactive SRCAP complex on the nucleosome can undergo a switch-like reconfiguration in which the Arp module engages linker DNA and, thereby, activates SWR1. The pivot point would lay on the motor domain where the HSA helix interacts with the N-lobe of the motor domain. The HSA domain and the motor are separated by an unstructured linker of species-specific length (350 residues in SRCAP contrasting 150 residues in *C. t.*). Yet, the C-terminal region of the HSA helix interacts directly with the motor lobe. Based on the RSC study, I speculate that the post-HSA helix as well as the Arp module undergo a rotation by about 120°. Notably, the kink in the Arp module facilitated by the Swc4 joint, allows Yaf9 to fold back onto the nucleosome and recognize acetylated histone tails. As a consequence, it is plausible that SWR1/SRCAP reads out DNA sequence information by the promoter bound Arp module thereby facilitating substrate specificity of SWR1/SRCAP for +1 nucleosomes. It would be also in agreement with findings that the distal H2A/H2B copy is exchanged in relation to linker DNA [230]. Further, I suggest a dual substrate recognition mode for SWR1 to specifically bind to +1 nucleosomes: 1) recognition of DNA sequence or shape features unique to promoter DNA 2) binding to acetylated +1 nucleosomes by the reader subunits Bdf1 and Yaf9. The fixed connection of the HSA helix to the motor domain further provides the possibility of coupling correct substrate recognition to ATP hydrolysis and thereby to histone exchange as also addressed in the next section. Overall, these insights provide a framework to understand how the multi-subunit SWR1 complex enables histone exchange adjacent to promoter sites. How SWR1 acts on DNA replication sites and DNA double strand breaks remains to be determined.

3.6 Comparison of Arp modules from the INO80/SWR1 family

The Arp module is an evolutionary conserved feature of both the INO80 and SWI/SNF family of the ATP-dependent chromatin remodelers as well as of NuA4-like histone acetyltransferase complexes. Common features within their Arp modules become apparent when looking at their subunit composition and molecular structures [150][143][144][145][146]. The Arp modules harbors the actin-fold proteins Arp4 and actin, which are anchored through barbed end interactions with conserved tryptophan residues of the HSA domain (illustrated in Figure 3.7b) [154] (see sequence comparison in Figure 3.8c). In yeast NuA4, the HSA helix is part of the Eaf1 subunit, which however does not contain the typical Snf2-type motor domain. Arp4 and actin are connected by the Swc4 or Ies4 subunits respectively which share a common 2W hairpin feature that anchors Swc4/Ies4 onto the actin molecule [150]. Furthermore, Swc4 is a shared subunit of the NuA4 and SWR1 complexes.

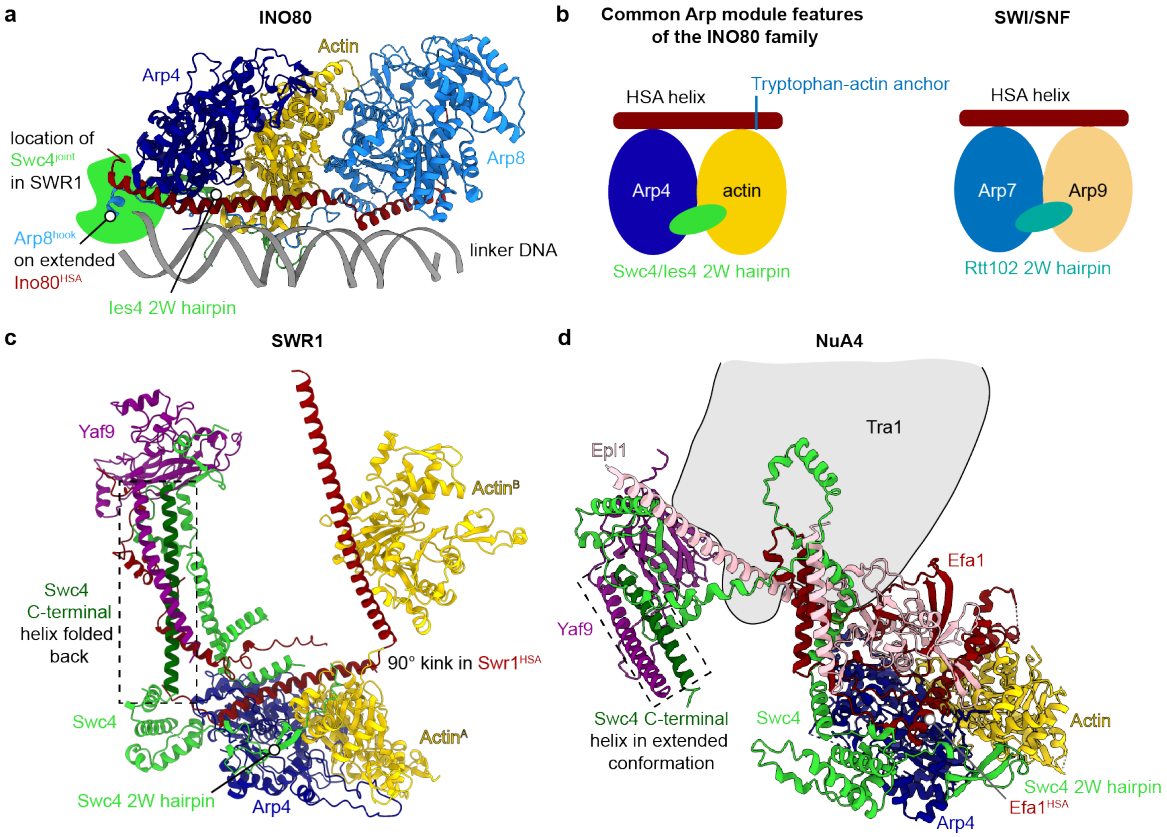


Figure 3.7: **Arp modules of the INO80/SWR1 family.** **a** INO80 Arp module (PDB 8A5P). Green sphere represents where Swc4 is located in SWR1. **b** *Left:* Common Arp module features of the INO80 family contrasting the Arp module in the SWI/SNF complexes on the right. **c** Arp module of *C. t.* SWR1 (this study). In dark green: C-terminal helix of Swc4 that interacts with Yaf9 (interaction highlighted by dotted box). **d** Arp module of NuA4 composed of PDB 7ZVW and my AlphaFold Multimer prediction of Eaf1, Epl1, Swc4 and Yaf9. C-terminal Swc4 helix is further away from the Arp module core compared to SWR1.

Such a minimal Arp module is also part of the SWI/SNF complexes comprising a HSA helix and actin-related proteins. However, SWI/SNF complexes harbor Arp7 and Arp9 [311] (illustrated in Figure 3.7b). In yeast, the 2W hairpin feature is conserved in form of the Rtt102 [122][150] subunit which also binds in between the Arp7 and Arp9 subunits. Coming back to the INO80 family, its

members comprise remodeler-specific features. While INO80 and SWR1 harbor three actin-fold proteins, NuA4 contains only one copy of actin and Arp4 due to its shorter HSA helix (see sequence comparison in Figure 3.8c). The third actin-related protein in INO80 and SWR1 are Arp8 and a second copy of actin respectively. Arp8 comprises a unique N-terminal extension that folds all the way back to Arp4 and ends in the Arp8^{hook} [150]. It sits where the Swc4^{joint} is located in SWR1 (Figure 3.7a). The Swr1 HSA helix is shorter at this position compared to the Ino80 HSA to make room for the Swc4^{joint} as elaborated in Section 2.4.2. The Ino80 subunit Ies4 is shorter compared to Swc4 that folds into the Arp module foot alongside with the Swr1 N-terminal region and Yaf9 (Figure 2.33). Also, the Ino80 subunit extends at the N-terminal end of the Arp module to assemble the N-terminal module and species-specific subunits. The subunits Ies1, Ies3, Ies5 and Taf14 assemble onto the N-terminus of the Ino80 subunit, but the flexibility of the underlying features prevented its structure determination. In addition, their function remains elusive. Yet, the INO80 subunit Taf14 shares the YEATS domain with Yaf9 which likewise enables the recognition of acetylated histone tails [195]. If Taf14 plays a role in nucleosome recognition and read out for nucleosome positions remains however elusive, especially as Taf14 is only flexibly tethered to the remaining complex. The N-terminal Arp module regions of SWR1 and NuA4 are better defined as they represent integral features of the Arp module. Both not only share Swc4 but also the Yaf9 subunit (see Figure 3.7c, d). However, Yaf9 was absent from the recent structural advances on the *S. c.* NuA4 complex [144][145][146]. Therefore, I prepared AlphaFold Multimer predictions of the Eaf1, Swc4, Yaf9 and Epl1 subunits which showed some interactions with the Swc4 subunit in the published NuA4 structures. The prediction revealed a conserved interaction of Yaf9 with a C-terminal helix of Swc4 in the SWR1 and NuA4 complexes. While the C-terminal Swc4 helix folds back onto further N-terminal Swc4 helices in SWR1, Swc4 is in an extended conformation in NuA4. As a consequence, the Yaf9/Swc4 domain is further away from the Arp module core in NuA4. This prevents clashes with Eaf1 and Epl1 features. Furthermore, the presented AlphaFold prediction is in line with the published structures and does not clash with the nearby Tra1 subunit. Overall, the Arp modules of NuA4 and SWR1 function as scaffolds bridging Yaf9 with the remaining complexes. The recognition of acetylated histones H3K27 and H4K9ac by Yaf9 drives the complex assembly of NuA4 and SWR1 [312]. As a consequence, Yaf9 regulates transcription initiation of metabolic genes [312]. In contrast, the Arp module of INO80 together with the yeast-specific Nhp10 module functions as a ruler element to allow for even spacing of nucleosomes [192][194] (depicted in Figure 3.8a). To enable its length-sensing function, the Arp8 subunit dimerizes with the close by actin molecule. Instead of Arp8, the SWR1/SRCAP complexes harbor a second copy of actin which does not dimerize with the other actin molecule. The SWR1/SRCAP complexes do not slide nucleosomes and consequently do not require a compact and well-connected length-sensing module like INO80. Instead, SWR1/SRCAP exchange the canonical H2A/H2B copies with the histone variant H2A.Z/H2B (Figure 3.8b). For this purpose, they harbor dedicated H2A.Z binding features, namely the Swc2-Z [233] and Swr1-Z domains [234]. The Swc2-Z domain is located at the N-terminus of Swc2 and was not captured in either of the available nucleosome-bound structures despite the addition of H2A.Z/H2B dimers upon sample preparation [142][143]. The Swr1-Z domain sits in the long insertion (150 residues in *C. t.* versus 350 in human) that intersects the Swr1 N-lobe and the C-terminal tip of the HSA helix. The latter folds onto the N-lobe of the motor domain despite the long insertion domain [143]. A similar architecture is observed in INO80 where the so called post-HSA with its QTELY motif sits on top of the Ino80 N-lobe (Figure 3.8c) [150]. The QTELY motif is a common feature in INO80 complexes of different species [154]. In fact, the post-HSA with its QTELY is a continuation of the HSA helix but can undergo a conformation switch. The conformational change of the post-HSA helix is supposedly linked to the activation state of the motor lobes that enclose the DNA [71][150]. The Ino80 post-HSA sits at the same location as the auto-N domain of ISWI and post-HSA of SWI/SNF [107][71].

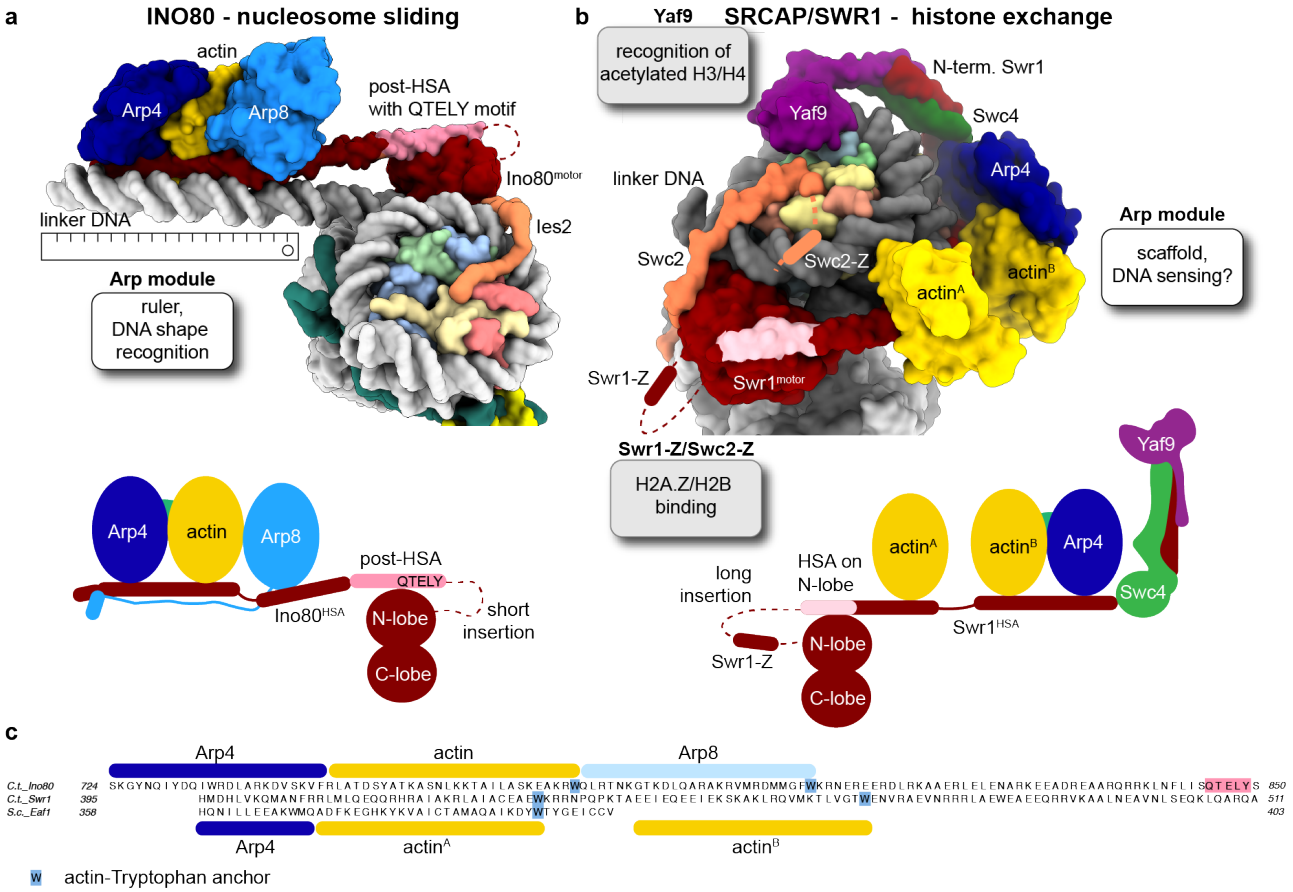


Figure 3.8: **Structure-function relationship of the INO80 and SWR1 Arp modules. a** INO80 Arp module connected by the Ino80 post-HSA helix to the Ino80 motor. Arp module functions as a ruler and recognizes DNA shape features of linker DNA. **b** SRCAP structure (PDB 8X19) highlighting function of different domains and illustrating locations of H2A.Z interacting domains Swc2-Z and Swr1-Z. HSA folds likewise to Ino80 back onto the motor's N-lobe, however lacks a QTELY motif. **c** Amino acid sequences of Ino80, Swr1 and Eaf1 HSA domains. QTELY motif in Ino80 highlighted.

The SWI/SNF family of chromatin remodeler contains the comparably QTXX(F/Y) motif within its post-HSA [154]. This is of interest as the Arp module of the SWI/SNF family couples likewise to INO80 the DNA translocation of its motor domain to nucleosome sliding and ejection [2]. Depending on the nucleotide state of the motor, the post-HSA gets unstructured [150] highlighting again its function in coupling and its involvement in regulation of motor activity. Furthermore, the Arp module of INO80 was shown to couple ATP hydrolysis to nucleosome repositioning [154]. It remains unknown, if a similar coupling mechanism occurs in SWR1. Yet, deletion of Swc4 and Arp4 abolishes histone exchange activity while the ATPase activity of the SWR1 complex remains similarly high [221]. Therefore and due to the HSA-motor interaction, a coupling mechanism similar to the INO80 and SWI/SNF complexes is plausible. Correct substrate recognition through acetylated histone tails by Yaf9 and Bdf1 or DNA recognition by the Arp module might be similarly transferred to the motor domain and therefore convey substrate specificity. In human, the Yaf9 homologue YEATS4 (Gas41) associates with H3K14ac and H2K27ac on active genes which increases SRCAP activity [212] and functions as recruiting mechanism for the SWR1/SRCAP complex. Strikingly, SWR1 requires about 50 bp of DNA on the linker DNA which fits to the binding of the Arp module to the extra-nucleosomal DNA [202]. Overall a tight regulation for the H2A.Z incorporation can be expected

due to its specificity for e.g. +1 nucleosomes or DNA double strand breaks. The motor domain hence needs careful activation which could be linked to a coupling mechanism of the HSA domain. However, the current structural model of SRCAP shows a closed conformation of the Arp module and thereby does not explain such sensing. In order to understand the histone exchange mechanism an integrative approach that combines further structural, biochemical and genomic experiments is needed in the future.

3.7 Future directions

The advances of cryoEM in the last decade enabled the structural characterization of single and multi-subunit chromatin remodelers including INO80 and SWR1 in complex with mono-nucleosomes [72][77][107][106][121][122][123][142][303][306]. However, we are only beginning to understand how chromatin remodelers function on a molecular level, how they are regulated and what defines their substrate specificity as well as reaction outcome. The structural and functional knowledge of these intricate molecular machines started to shed light on the mechanism of the underlying nucleosome remodeling reactions. However, these structures only present glimpses into the reaction cycle. While the motor domains of chromatin remodelers were captured in different nucleotide states, they fail to shed light on the complete reaction cycles. This is the case, as chromatin remodelers accumulate multiple ATP hydrolysis steps to acquire enough DNA underwinding or DNA strain to remodel nucleosomes [76]. As a consequence intermediate and late reaction stages are unknown and it remains unclear how chromatin remodeler subunits facilitate larger conformational changes upon release of DNA strain. As a possible solution, one could prepare cryoEM samples where one adds ATP just prior to sample vitrification to start the respective nucleosome remodeling reaction. This would create complex cryoEM samples where one captures a multitude of reaction stages depending on the reaction time between ATP addition and sample vitrification. Therefore, large data sets and excessive classifications in the following data processing would be required to enable the determination of high-resolution structures. To lower the complexity, one should make use of the arising specialised equipment for time-resolved cryoEM (recently reviewed in [313] and [314]). These are dedicated vitrification devices that allow for a precise control of the reaction time that elapses from start of the reaction to vitrification. Classically, the reaction is started by mixing the protein of interest with its substrate. In case of chromatin remodelers, I suggest to mix them with ATP to start and control the nucleosome remodeling. Mixing is either performed on the grid by two subsequent sample applications or by dedicated microfluidic setups. For fast sample application, multiple labs have developed mixing-spraying devices that are not commercially available yet [315][316][317]. The Spotiton device that uses one or two piezo elements for sample application is available on the market [318]. With these setups small reaction times of around 10 ms are feasible depending on the device. Therefore, also short-lived intermediates can be captured as for example illustrated in ribosome assembly [319]. Yet, the usage of such dedicated devices presupposes the knowledge of the reaction duration. The usage of single molecule FRET studies can be a help in this regard as also performed on chromatin remodelers [229][320]. Here, it became clear that the reactions rather happen on the scales of a few seconds. Therefore a dedicated time-resolved vitrification device might even not be needed. As a consequence, one could simply play around with a mixture of active versus non-active substrates to slow down the reaction (e.g. ATP vs ATPγS).
More studies are required to understand what kind of (epi)genetic information is required to regulate nucleosome remodeling and how this information is allosterically transferred to the motor domains. Remodeler release off the DNA was shown to be also ATP dependent indicating that one last ATP hydrolysis step is required at the end [196]. DNA shape features and sequences have been identified to

regulate nucleosome positioning which would be one possible end point of nucleosome remodeling [36][169]. Additionally, collision with neighboring nucleosomes might for example regulate nucleosome sliding as postulated for the ISWI remodeler [75][108]. The same features together with histone modification readers function in recruitment of chromatin remodelers [312]. In addition, regulatory proteins or transcription factors can target chromatin remodelers to their nucleosomal substrate [321][322]. Yet, in cryoEM mostly artificially positioned nucleosomes are used that lack the context of genomic DNA and histone modifications. Genome wide studies using techniques like ChIP sequencing can inform on better cryoEM samples to simulate remodeler-promoter interactions or nucleosome positioning for example. Further, studies on di-nucleosomes or nucleosmal arrays with possibly also native DNA can be informative. Instead of purified protein samples one could use lysates that are incubated with pre-formed nucleosome arrays. This would open the possibilities of integrative approaches combining sequencing, mass spectrometry and structural methods to understand what factors regulate and interact with chromatin remodelers. Ultimately, the usage of *in situ* structural biology methods can aid to capture the action of chromatin remodelers in their native cellular environment.

4 Materials and Methods

4.1 Subunit identification of *C. t.* SWR1

Subunits were identified using Blast®® protein-protein alignments. Amino acid sequences of *S. c.* SWR1 subunits were used as starting point to identify their homologues in *Neurospora crassa* (*N. c.*). Based on *N. c.*, hits respective *C. t.* SWR1 subunits were characterized. Swc7 homologue in *N. c.* as well as *C. t.* could not be identified. Blast® searches of Swc7 returns yeast entries only. Entries for *C. t.* SWR1 subunits were validated against respective subunits of *Homo sapiens*, *Drosophila melongastser*, *Mus musculus*, *Saccharomyces cerevisiae* and *Saccharomyces pombe*. In addition, sequences for the ten closest fungi entries of the respective subunit were identified. All of these amino acid sequences were aligned against the identified *C. t.* subunits by performing ClustalO [287] within the Jalview software [284][285]. Swc5, Arp6 and Swc6 were annotated incompletely. The N-terminus of Swc5 was found as separately identified protein: both sequences were therefore merged (uniprot entries: G0SHD9 as N-terminal region, G0SHE0 as C-terminal region). Arp6 was missing an internal conserved stretch of amino acids. Therefore, Arp6 was amplified from cDNA (generated by Olga Kolesnikova) and sequenced. Further, Swc6 missed some amino acids at its C-terminus. The underlying amino acids could be identified when searching for them in the scaffolds making use of the database from the Bork lab [286]. The identified sequences for Swc5, Arp6 and Swc6 are as follows:

```
>Swc5
MTAEFEDQDDNYVSEEDSDFAPDDAAEEDESLSSDDDEAAEGESAPAKPDRKRPAADSSADDVGFENSGDEAIIER
GKKNKRDIRKKLAPTRIGRKRTGHMRAAEKEGKYAKIPKGPVTIDVDKLWQDLISAPVVRVPGQDTTATTQPQPNG
DAQKPSQSKPVPNAADESEYIRIKRTYTFAGKTHTEEKLVHRESAEAKLYLAETGQDPSAPAVSPSDADQQQKRMP
RKAFRSAFEPVPADGTGVSQRSDLNLGIAKLREQREKAEQAKKLNTVEKSRMDWAGYVDREGIKDELEMAGKSKHS
FVARQEFLARSEAVREEEVRRLRLAGKV

>Arp6
MAGGSRKSRSVAGGPAAPSTTLVLDNGADTIKAGLIPTGSSGGDDNTPTQLQPRIIPNCIARDRHKKIYVASELDK
CRDFGEMAFRRPVEKGYIVNWEAQKEIWEREFFDEKAPLRCDPAETRLVLAEQPNALPALQTHCDQMVFEEFGFAS
YVRGIGPAFNAYQDIQSIFQTPCDQNMPVTLPAEIMLLIDSGYSHTTITPLLQGRPLHPAIRRLDVGGKLMTNYLT
RLISMRHFDMRNETYVVNEMKEAACYVSLDFKGDLEKTWKGTRDEKRKDWLTGAGIAKDYVLPDSHTRFHGVLRDY
DPNASRARRGIPNTEDVLTLRNERFVVPELLFHPSDIGLRQPGLADLIMQSLSVLPIGLWPGLLANVVVVGGNAKF
ENFIQRLQMELLERVPTECVVRVARPENPIISTWLGASNFARHEHVERLAVTKQQYEEYGAGWVARKFAAGLGVDS

>Swc6
MTAEFEDQDDNYVSEEDSDFAPDDAAEEDESLSSDDDEAAEGESAPAKPDRKRPAADSSADDVGFENSGDEAIIER
GKKNKRDIRKKLAPTRIGRKRTGHMRAAEKEGKYAKIPKGPVTIDVDKLWQDLISAPVVRVPGQDTTATTQPQPNG
DAQKPSQSKPVPNAADESEYIRIKRTYTFAGKTHTEEKLVHRESAEAKLYLAETGQDPSAPAVSPSDADQQQKRMP
RKAFRSAFEPVPADGTGVSQRSDLNLGIAKLREQREKAEQAKKLNTVEKSRMDWAGYVDREGIKDELEMAGKSKHS
FVARQEFLARSEAVREEEVRRLRLAGKV
```

4.2 Cloning

4.2.1 *C. t.* SWR1

C. t. SWR1 was recombinantly cloned making use of the MultiBac system [291] for insect cell expression. Codon optimized (for Hi5 insect cell expression) DNA sequences of Swr1, Swc2, Swc6, Bdf1, Yaf9, Swc3, Swc4 and Swc5 were ordered from Invitrogen (ThermoFischer). Swr1 was ordered with a 2xFLAG tag for affinity purification (amino acid sequence: `DYKDDDDKGTDYKDDDDK`) Arp6 was amplified from cDNA that was kindly provided by Olga Kolesnikova. Rvb1, Rvb2, Arp4 and actin were already available in the lab from the recombinant protein production of INO80. In a first step, each subunit was PCR (polymerase chain reaction) amplified and cloned into the gene cassette of one of the following empty vectors: pIDK (kanamycin resistance), pIDS (spectinomycin resistance), pIDC (chloramphenicol resistance) or pACEBac1 (gentamycin resistance). While the latter is a donor vector required for recombination with the baculovirus genome, the other represent acceptor vectors. They comprise a polyhedrin promotor and SV40 terminator signal for over-expression in insect cells. First, the vectors were linearized by PCR and purified from a 1 % agarose gel with the QIAEX®II Gel extraction kit (Qiagen). Linearized vectors were stored at −20 °C. Next, subunits were amplified with forward and reverse primers, both of which contained a 15 bp overhang resembling the two ends of the linearized vectors. In more detail, 20 ng of template DNA were amplified using the commercially available Polymerase from Thermo Scientific in the respective HF or GC buffer depending on the GC content of the PCR product. PCR products longer than 2000 bp were amplified with the specialised KOD hot start Polymerase. PCR products were analysed on a 1 % agarose gel and PCR purified with the QIAquick® PCR purification kit (Qiagen) when only one specific band was visible. If that was not the case, PCR conditions were optimized by changing the annealing temperature and extension time or if possible the desired band was cut from agarose gel and gel purified with the QIAEX®II Gel extraction kit (Qiagen). Subsequently, vector and subunit insert were combined utilizing the In-Fusion kit from Takara Bio US with a three-fold molar excess of insert DNA. Constructs of pACEBac1 were transformed by heat shock at 42 °C into XL1blue cells (Stratagene) while the others were transformed into pIR cells (Geneva Biotech). Plasmids were miniprepped with the QIAprep®Spin Miniprep kit (Qiagen) and sequenced with the Eurofins GATC sequencing service. In this manner Rvb1, Rvb2, Bdf1 and Yaf9 were cloned into pACEBac1; Arp4, actin, Swc3 and Swc4 into pIDK; Swr1, Swc2 and Swc5 into pIDS and lastly Swc6 and Arp6 into pIDC respectively. Subsequently, plasmids containing two subunits were prepared by PCR and In-Fusion reaction. For this purpose the plasmid of one subunit was linearized while the second subunit was amplified only together with its promotor and terminator. Again, primers with a 15 bp overhang were used allowing subsequent In-Fusion reaction. Likewise, minipreps and sequencing were performed. The final plasmids were assembled through Cre recombination (New England Biolabs, NEB) by combining one of the donor vectors with one acceptor vector resulting in a plasmid that contained in total four gene cassettes of the respective subunits. Plasmids were miniprepped and analysed by PCR to check if the desired subunits were present. The product was used again as a donor plasmid for the next round of Cre recombination. In this way, two plasmids were generated containing all subunits. The first plasmid contained: Rvb1, Rvb2, Swr1, Swc2, Arp4, actin, Swc6 and Arp6, while the second plasmid comprised the subunit of Bdf1, Yaf9, Swc3, Swc4 and Swc5. One plasmid would have been to large for handling. Of note is, that typically XL1blue cells (Stratagene) were used for transformation but also DH5α cells (Life Technologies) were tested and worked. With each Cre recombination step also the number of used antibiotics for colony selection grew, however cells tended not to grow on more than two antibiotics.

In the later stages, I therefore plated on a maximum of two antibiotics and streaked the picked clones onto the remaining antibiotics to make sure that they had all desired antibiotics resistance genes on the plasmid.

4.2.2 *S. c.* SWR1

Genes were ordered from Invitrogen (Thermofisher) except for Rvb1, Rvb2, actin and Arp4, which were amplified from cDNA from the Müller lab at EMBL. The same cloning strategy and procedure was applied like for *C. t.* SWR1. Subunits were cloned into the following vectors: pACEBac1: Swr1, Swc5, Swc7; pIDS: Swr1, Arp4, actin; pIDK: Swc6, Arp6, Swc3; pIDC: Swc2, Swc4, Bdf1, Yaf9. It was necessary to place Swr1 onto the donor vector pACEBac1 due to its size. Having it on pIDS, prevented successful Cre recombination. Swr1 subunit contained a double FLAG-tag on its C-terminus. Lastly, the following two final plasmids were generated through consecutive rounds of Cre recombinations: Plasmid 1: Swr1, Rvb1, Rvb2, Swc6, Arp6, Swc2, Swc4; Plasmid 2: Swc5, Swc7, Arp4, actin, Swc3, Bdf1, Yaf9. Final plasmids were sequenced via the Plasmidsaurus service to check for potential point mutations.

4.2.3 Arp module constructs

C. t. Arp module constructs were cloned based on the MultiBac system introduced in Section 4.2.1. Briefly, three constructs of Swr1 were cloned onto pACEBac1 including a C-terminal double FLAG tag: 1) residues 386-458 for $Arp^{HSA\ minimal}$ construct 2) residues 386-535 for $Arp^{HSA\ FL}$ construct 3) residues 1-535 for $Arp^{N\text{-}HSA\ FL}$ construct (FL: full-length, N for N-terminal region). The different Arp module constructs were assembled by Cre recombination (from NEB) by adding Swc4-pIDC and Arp4-actin-pIDK to the Swr1-pACEBac1 plasmids respectively.

4.2.4 Cloning of other constructs

The single subunits Swc5, Bdf1 and eGFP were required for *Chaetomium thermophilum* pulldowns and expressed in insect cells. The single subunits were cloned onto a pACEBac1 vector to enable baculovirus production and subsequent protein expression in insect cells. In addition, a C-terminal double FLAG tag was added. The double FLAG tag was inserted into the Swc5-pACEBac1 vector by amplifying the tag from another construct and using it as a megaprimer. As the megaprimer was designed with complementary overhangs to the Swc5-pACEBac vector, the tag could be inserted by PCR. The original vector was cleaved with DpnI (NEB) prior to PCR clean up and transformation. For Bdf1 and eGFP, the new Swc5-2xFLAG-pACEBac1 vector was linearized with primers occluding the Swc5 sequence and gel-purified with the QIAEX®II Gel extraction kit (Qiagen). In a second PCR, Bdf1 or eGfP were amplified with primers creating 15 bp overhang to the linearized 2xFLAG-pACEBac vector. Finally, 2xFLAG-pACEBac1 vector and Bdf1 or eGFP inserts were combined using the In-Fusion reaction kit (Takara Bio US).

4.3 Protein expression and purifications

4.3.1 Baculovirus generation and protein expression

Swc5, Bdf1, eGFP, Arp module constructs, SWR1 and INO80 were expressed in insect cells. The latter two were expressed by co-infecting insect cells with two separately generated baculoviruses. For

this purpose, Bacmid DNA was prepared by transformation of the respective construct into electro-competent DH5α embacY cells (Geneva Biotech) that contain a plasmid with the genetic blueprint for the baculovirus. Cells were picked according to blue-white selection and cultivated overnight in 5 ml LB medium. Cells were lysed using the Qiagen miniprep buffers: P1, P2 and N3. The supernatant of the first centrifugation step (about 700 µl) was mixed with 700 µl of 100 % isopropanol and kept on ice for 30 min allowing the DNA to precipitate. Next, the DNA was pelleted for 30 min at 20 000 g and 4 °C. The resulting DNA pellet was washed in 70 % ethanol once, followed by another centrifugation step of 15 min. Finally, the pellet was air-dried under a sterile hood for about 15 min and dissolved in 30 µl deionized water. The procedure resulted in DNA stocks of about 4000 ng/µl and were kept in the fridge until usage. Presence of subunits was checked performing PCR reactions with 10 ng of Bacmid DNA in a volume of 25 µl. Baculoviruses were generated (V0) using the X-treme transfection reagent (Roche). Therefore, 15 µg of Bacmid DNA were mixed with 100 µl of Sf900^{TM} SFM (1x) Serum free medium complete (Gibco). At the same time 5 µl of X-treme reagent were mixed with 50 µl medium. The two components were combined and left for 15 min while 3 ml of Sf9 or Sf21 cells (ThermoFisher Scientific Cat. 11497013) were prepared at 0.3×10^6 cells/ml in a six-well plate. The transfection mix was added to the cells drop-wise and incubated in the dark for about 60 h at 27 °C. Resulting viruses were checked for eGFP fluorescence, collected and used for virus amplification (V1). For this purpose 500 µl and/or 1 ml of V0 viruses were added to 25 ml of Sf9 or Sf21 cells in an Erlenmeyer flask. Viruses were harvested after 65 h and cell parameters were checked. Ideally, the cells should have doubled about once – otherwise the virus was too strong. In later virus generations, Sf9 cells were used for both V0 and V1 preparations as this results in higher protein expression yields. Final protein expressions were performed in Hi5 cells (ThermoFisher Scientific Cat. B85502) for about 60 h at 27 °C at virus dilutions ranging from 1:100 (INO80 and SWR1; alternatively 1:200 or 1:500 used) to 1:1000 (single subunits). Insect cells were harvested by centrifugation at 800 g for 30 min. Resulting pellets were resuspended in PBS buffer (phosphate buffered saline) and centrifuged once more for 15 min in a table top centrifuge. Pellets were flash frozen in liquid nitrogen and stored at −80 °C until usage.

4.3.2 INO80

C. t. INO80

C. t. INO80 was expressed and purified likewise to Eustermann et al. [77]. Constructs from the Hopfner lab were used in which the eleven INO80 subunits are divided among two Multibac vectors [291]. In more detail, the N-terminal truncated INO80-complex (Ino80 residues 718-1848) comprises a double FLAG tag on the C-terminus of the Ino80 subunit that was used for an anti-FLAG affinity purification step followed by ion exchange chromatography. Two litres of Hi5 cells were lysed by sonication (Beckmann, JA 25.5 rotor; at 40 %) in 30 mM HEPES pH 7.5, 500 mM NaCl, 10 % glycerol, 0.25 mM DDT and EDTA-free protease inhibitor tablets (Roche, one tablet per 50 ml of lysis buffer, 3 ml of lysis buffer per 1 g of pellet). In the following centrifugation step, the lyaste was cleared at 20 000 g for 1 h. Next, the supernatant was kept and incubated with 1 ml of anti-FLAG M2 affinity gel (GenScript, Cat. L00432) for 2 h. Beads were captured in a flow through column (Biorad) and washed with 25 column volumes of lysis buffer followed by 15 volumes of wash buffer (30 mM HEPES pH 7.5, 150 mM NaCl, 5 % glycerol, 0.25 mM DDT and 0.25 mM $CaCl_2$). Elution was performed in six column volumes of wash buffer supplemented with 0.2 mg/ml single FLAG peptide. For this purpose, three times two column volumes of elution buffer were incubated for 20 min with the beads before capturing the eluted protein. Finally, INO80 was purified over a MonoQ 5/50 GL column (GE Healthcare) or CaptoQ^{TM}HiRes Q5/50 (Cytiva) to remove DNA bound complex. The

sample was loaded in presence of low salt buffer (30 mM HEPES pH 7.5, 150 mM NaCl, 0.25 mM DDT and 0.25 mM $CaCl_2$). INO80 was eluted through a salt gradient that increased to 1 M NaCl over one hour. Due to the collection of 50 µl fractions and the usage of an ÄKTA micro setup, no protein spin concentration was required. Instead, INO80 fractions were pooled and flash frozen in liquid nitrogen in the presence of 25 % glycerol. For cryoEM purposes, INO80 was stored over night in the cold room (4 °C) and used for fresh grid preparations. INO80 was analysed on SDS-PAGE NuPAGE^{TM} 4-12 % Bis-Tris Gel 1 mmx 10 or 12 wells in MOPS buffer (invitrogen).

INO80-Arp mutants

The INO80-Arp5 mutants (Arp5-heel: R501E-K502E, Arp5-heel+foot:R501E-K502E-Q507A-R509A-M510S-K511A-I513S) were generated by Franziska Kunert applying the quick change mutagenesis technology (NEB) on the Arp5-Ies2-Ies6-pIDK contruct. First the heel mutant was generated to use the resulting construct for the heel plus foot mutant. The Arp5-Ies2-Ies6-pIDK construct was added to the Ino80(2xFLAG)-Rvb1-Rvb2-pACEBac plasmid using Cre recombination (NEB). Expression and purification was done likewise to wild type *C. t.* INO80.

Yeast INO80

Yeast INO80 was purified by Franziska Kunert according to [150].

4.3.3 *C. t.* SWR1

SWR1 was purified likewise to INO80 with some minor changes in the buffer compositions. As SWR1 contains the zinc finger binding subunit Swc6, 20 µM $ZnCl_2$ were added in the purification buffers. Otherwise, similar buffers where used compared to [323]. As KCl seemed to stabilize Bdf1 within the SWR1 complex, it was used in the wash buffers and for the ion exchange chromatography. This results in the following buffer compositions: Lysis Buffer: 50 mM HEPES pH 7.0, 500 mM NaCl, 10 % glycerol, 0.25 mM DDT, 20 µM $ZnCl_2$; Wash Buffer: 50 mM HEPES pH 7.0, 150 mM KCl, 5 % glycerol, 0.25 mM DDT, 20 µM $ZnCl_2$, 0.5 µM $CaCl_2$; Elution Buffer: Wash Buffer supplemented with 0.2 mg/ml single FLAG pepetide; MonoQ Buffer A: 50 mM HEPES pH 7.0, 150 mM KCl, 0.25 mM DDT, 20 µM $ZnCl_2$, 0.5 µM $CaCl_2$, MonoQ Buffer B: 50 mM HEPES pH 7.0, 1 M KCl, 0.25 mM DDT, 20 µM $ZnCl_2$, 0.5 µM $CaCl_2$. SWR1 was analysed on SDS-PAGE NuPAGE^{TM} 4-12 % Bis-Tris Gel 1 mmx 10 or 12 wells in MOPS buffer (invitrogen).

4.3.4 Arp module constructs

The Arp module constructs Arp$^{\text{HSA minimal}}$, Arp$^{\text{HSA FL}}$, Arp$^{\text{N-HSA FL}}$ as well as Arp$^{\text{N-HSA FL}}$-Yaf9 were purified likewise to SWR1 as described above. For the expression of Arp$^{\text{N-HSA FL}}$ with Yaf9, Hi5 insect cells were co-infected with viruses for the Arp module and the Bdf1-Yaf9-Swc3-Swc4-Swc5 construct. The HSA helix of the Arp module was double FLAG tagged and could hence pull down the unlabeled Yaf9 subunit. Hi5 cells were co-infected with virus ratios of 1:300 respectively. In one trial Arp$^{\text{N-HSA FL}}$construct was purified in presence of LatB (Latrunculin B) and ATP. The protocol from the SWR1 purifications only differed in this point: After washing the anti-FLAG affinity beads two times with 10 ml lysis buffer, the beads were incubated with lysis buffer supplemented with 10 µM of LatB and 1 mM of ATP.

4.3.5 Yeast SWR1

Yeast SWR1 was similarly purified compared to *C. t.* SWR1 but with minor adaptations. Sonication was performed at milder conditions at 20 % intensity for two times 20 s (Sonifier® W-250 D). For cell lysis, a buffer including EDTA and magnesium chloride was used to resemble the buffer used for yeast SWR1 purifications [221]. The buffer composition was therefore: 50 mM HEPES pH 7.0, 500 mM NaCl, 10 % glycerol, 0.25 mM DDT, 20 µM $ZnCl_2$, 1 mM EDTA and 2 mM $MgCl_2$. Washes of FLAG beads was performed in 50 mM HEPES pH 7.0, 300 mM KCl, 5 % glycerol, 0.25 mM DDT, 20 µM $ZnCl_2$, 1 mM EDTA, 2 µM $MgCl_2$ meaning that a higher salt concentration compared to the *C. t.* SWR1 purifications was used for washing. Eluted yeast SWR1 was diluted to 150 mM KCl prior to loading on a CaptoQ column. The CaptoQ chromatography step was performed likewise to *C. t.* SWR1.

4.3.6 Swc5 and Bdf1

Swc5

Swc5 was purified likewise to SWR1 using the same buffers. However, the incubation of the lysate with the anti-FLAG beads was performed overnight and the purification procedure was continued on the next day. Also, a more efficient bead washing protocol was used, which started with three 10 ml lysis buffer washes of the beads in a Falcon tube with subsequent 800 g centrifugation steps. Next, the beads were transferred in 15 ml wash buffer to the flow through column (Poly Prep Chromatography gravity flow column, Bio-Rad). Here, the beads were only washed with an additional 10 ml of wash buffer prior to sample elution. Anion exchange chromatography was performed with a CaptoQ^{TM}HiRes Q5/50 (Cytiva) column which is equivalent to a MonoQ column.

Bdf1

Purification of Bdf1 was performed likewise to *C. t.* SWR1 and Swc5. Binding to the anti-FLAG beads was enabled overnight. Capto Q buffers were adjusted to Buffer A: 50 mM HEPES pH 7.0, 150 mM KCl, 0.25 mM DDT and Buffer B: 50 mM HEPES pH 7.0, 1 M KCl, 0.25 mM DDT.

4.3.7 eGFP

The protein eGFP was purfied via its C-terminal double-FLAG tag followed by protein spin-concentration. The protocol for cell lysis and the anti-FLAG purification was adapted from SWR1 purifications. However, the buffer composition and washing procedure differed. Cell lysis was performed in 50 mM HEPES pH 7.0, 500 mM NaCl, 10 % glycerol 0.25 mM (lysis buffer). Initial washing of FLAG beads was done three times in 10 ml of wash buffer 1 (50 mM HEPES pH 7.0, 150 mM NaCl, 5 % glycerol, 0.25 mM DDT) by centrifugation prior to transferring the beads in 10 ml of wash buffer 1 into a Biorad flow-through column. Washing was continued with 20 ml of wash buffer 1 and wash buffer 2 (wash buffer 1 minus glycerol). Elution was performed in wash buffer 2 supplemented with 0.2 mg/ml single FLAG peptide (4*x* 2 ml). FLAG elutions were concentrated in a Millipore spin concentrator (3 kDa cut-off, Amicon® Ultra Centrifugal Filters). Aliquots were flash frozen in liquid nitrogen in presence of 15 % glycerol and stored at −80 °C until further usage.

4.3.8 *C. t.* H2A.Z and H2B

Expression in E.coli

C. t. HA2.Z and H2B were expressed in BL21 Rosetta cells (Novagene). For this purpose, bacteria were plated on kanamycin and chloramphenicol LB-Agar plates on the day prior to protein expression. A pre-culture of 20 ml per one litre expression culture was prepared by rinsing off the colonies from one plate per three litre of expression culture (only used kanamycin for pre-culture). Pre-culture was grown at 37 °C for 2.5 h. The 1 l expression culture was inoculated with 20 ml of the pre-culture and grown to an OD_{600} of 0.6 at 37 °C and 190 rpm. Expression was started with 1 mM IPTG and was enabled for 3 h. Cultures were pelleted at 3000 g and stored at −80 °C until usage.

Inclusion body purification

Histones were purified from inclusion bodies as follows: E.coli pellets from 3 l were resuspended in 30 ml wash buffer (50 mM Tris-HCL pH 8.0, 100 mM NaCl, 1 mM DDT) that was supplemented with 1 mg/ml lysozyme (BioChemica, PanReac AppliChem). Cells were further lysed by sonication for three times three minutes and pelleted at 15 000 rpm (Beckmann, JA 25.5 rotor) and 4 °C for 15 min. Pellets were washed four times in 40 ml wash buffer (first two washes in presence of 1 % Triton X-100) using a glass homogenizer. Inclusion bodies were pelleted in between the washing steps at 15 000 rpm for 15 min (Beckmann, JA-25.5 rotor). Final pellet was frozen and stored at −80 °C.

Dialysis against trifluoroacetic acid

In brief, the purified histones are unfolded in this step by guanidinium HCl. Next, the guanidinium was removed in a dialysis against TFA (trifluoroacetic acid) which in turn causes the precipitation of nucleic acids. In detail, inclusion bodies were disolved in 1 ml of DMSO at room temperature for 30 min. Afterwards, 30 ml of unfolding buffer (50 mM Tris-HCL pH8.0, 100 mM NaCl, 6 M Guanidinium HCl, 1 mM EDTA, 10 mM DDT) was added for 1 h. The solution was centrifuged at 12 000 rpm for 1 h at 4 °C. The resulting supernatant was dialysed three times against 2 l of 0.1 % (v/v) trifluoroacetic acid (TFA) for 3 h for respectively. Dialysed material was centrifuged for 1 h at 12 000 rpm (Beckmann, JA-25.5 rotor) and 4 °C to remove precipitated nucleic acids. Histones were lyophilized for at least 48 h. Samples were stored at −20 °C until histone purificaton.

Histone purification

Lyophilized histones were dissolved in 7 ml of buffer A (20 mM HEPES pH7.0, 7 M Urea, 10 mM lysine) and cleared by a short centrifugation. H2A.Z and H2B were purified on an ÄKTA purifier on a HiTrap SP (5 ml) column. A two-fold washing step was performed at 10 % and 20 % buffer B (20 mM HEPES pH 7.0, 7 M Urea, 1 M NaCl, 10 mM lysine). Subsequently, histones were eluted on a 20-60 %B gradient within 30 min. Fractions of 2 ml (better to use 1 ml) were collected and carefully analysed on a SDS-PAGE NuPAGE^{TM} 4-12 % Bis-Tris Gel 1 mm, 10 or 12 wells (invitrogen) in MES buffer (invitrogen) before pooling similar fractions. Purified histones were dialyzed (3.5 kDa cut-off) against 15 mM Tris-HCl pH 7.5-7.7, aliquoted in 1 mg portions and lyophilized over night.

4.4 Nucleosome, hexasome and Histone dimer reconstitutions

4.4.1 Histone octamer reconstitution

Individual canonical human histones (H2A, H2B, H3.1, H4) were purchased from The Histone Source (Colorado State University, USA) and dissolved in unfolding buffer (25 mM Tirs-HCl, pH7.5, 7 M guanidiniumchloride, 5 mM β-mercaptoethanol). Histones were mixed with the following molar ratios: H2A:H2B:H3.1:H4 1.2:1.2:1.0:1.0. Histones were diluted to 1 mg/ml and dialysed four times against 1 l of refolding buffer (10 mM Tris-HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM β-Mercaptoethanol) for 1 h respectively but the last step over night. Formed octamers were concentrated to a final volume of 1-1.5 ml and purified on a Superdex 200 increase 10/300 (Cytiva) column in the presence of refolding buffer. Octamer fractions were concentrated to at least 1 mg/ml and stored in 50 % of glycerol at −20 °C.

4.4.2 Synthesis and purification of nucleosomal DNA

Nucleosomal DNA was produced by amplifying the DNA template from a pEX vector containing the 601 sequence with adjacently available DNA that could serve as linker DNA. Either 0N80, 25N80, 60N80 or 80N80 were produced. For the hexasome project 0N80, 0H80, 0H40, 0N20 and 0H20 were produced. The forward primer was designed to determine the linker on the 5' end of the 601 sequence, while the reverse primer set the linker on the 3' end. Therefore, the following primers were use:

- forward primer for 0 bp linker: `5' CTGGAGAATCCCGGTGCCGAGG 3'`
- forward primer for 25 bp linker: `5' GTGATGGACCCTATACGCG 3'`
- forward primer for 60 bp linker: `5' AAAGCATGATTCTTCACACCGAG 3'`
- forward primer for 80 bp linker: `5' GGGATCCTAATGACCAAGGAAAGC 3'`
- reverse primer for 0 bp linker: `5' ACAGGATGTATATATCTGACACGTGC 3'`
- reverse primer for 25 bp linker: `5' GGTCGCTGTTCAATACATGCACAGGATG 3'`
- reverse primer for 80 bp linker: `5' CGGTACCCGGGGATCCTCTAG 3'`

This resulted in the following DNA sequences for 0H80 and 0N80 used for the structural studies of INO80 (601 sequence in bold):

5'-**CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTACGC
GCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGT**G
CATGTATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAGCTCCCACTCTAGAGGATCCCCGGGTAC
CG-3'

DNA sequence of 80N25 nucleosomes:

5'-GGGATCCTAATGACCAAGGAAAGCATGATTCTTCACACCGAGTTCATCCCTTATGTGATGGACCCTATACGC
GGCCGCC**CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTA
CGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCT
GT**GGCAAGGTCGCTGTTCAATACATGC-3'

For PCR amplification, 2x 96-well plates with a PCR volume of 100 µl per well were prepared. PCRs were performed in the presence of 50 % glycerol and HF buffer (Thermo Scientific). Extension time was set to 15 s. PCR products were pooled for subsequent anion exchange chromatography with a Source15Q 4.6/100 (Cytiva) column. The column was equilibrated in 25 mM Tris-HCl pH 7.5 and bound DNA was eluted through a gradient with 25 mM Tris-HCl pH 7.5 and 2 M NaCl. Fractions were analysed on a 1 % agarose gel to differentiate between the desired PCR product and later eluting primers. Fractions were pooled and dialysed overnight against 1 l of 5 mM Tirs-HCl at pH7.5. Subsequently, DNA was ethanol precipitated and resolved in water.

4.4.3 Nucleosome reconstitution

Octamers were dialysed against N2000 buffer (25 mM Tris-HCl pH 7.5, 2 M NaCl, 0.5 mM β-Mercaptoethanol) to remove the glycerol. DNA and octamers were mixed at a molar ratio of 0.9:1. DNA-octamer mixture was transferred into a slide-A-Lyzer® (0.1-0.5 ml, 3.5 kDa cutoff, Thermo Scientific) and placed into a bucket of 300 ml of N2000 buffer. In an overnight procedure, the salt concentration was lowered by titration with 3 l of N50 buffer (25 mM Tris-HCl, pH7.5, 50 mM NaCl, 0.5 mM β-Mercaptoethanol) through the usage of a peristaltic pump (Minipuls®3 from Gilson). The next day, nucleosomes were purified on a Source 15Q 4.6/100 (Cytiva) column in the presence of N50 buffer. Nucleosomes were eluted over a gradient with N2000 buffer. Typically, the chromatogram profile showed three peaks: the first one being the nucleosomes, then hexasomes eluted and finally free DNA. Nucleosomes were dialysed over night to buffer N50 and were concentrated in a 0.5 ml spin-concentrator with a cut-off of 10 kDa (Amicon® Ultra Centrifugal Filters).

4.4.4 Preparation of hexasomes

Hexasomes were prepared likewise to [68]. In brief, human H2A/H2B and human H3.1/H4 tetramers were reconstituted likewise to the octamer preparation described above. Next, dimer, tetramers and DNA were mixed at a ratio of 1.2:1.1:1 in N2000 buffer and purified likewise to nucleosomes.

4.4.5 *C. t.* H2A.Z/H2B dimer reconsitution

C. t. H2A.Z/H2B dimers were reconstituted by unfolding the individual histones followed by refolding in presence of the other histone binding partner. Finally, dimers were purified by size exclusion chromatography. In more detail, 1 mg of H2A.Z and H2B were dissolved in 250 µl unfolding buffer (25 mM Tris-HCl pH 7.5, 7 M guanidinium chloride, 1 mM EDTA, 5 mM β-mercaptoethanol) respectively. Equimolar amounts of H2A.Z and H2B were mixed and diluted to 1 mg/ml by adding 5 ml of unfolding buffer. Histones were dialysed (SpectraPor3 dialysis membrane, 3.5 kDa) four times against refolding buffer (10 mM Tris-HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) for 1.5 h each, while the last step was performed over night. Formed dimers were concentrated to approximately 1-1.5 ml (3 kDa cut-off, Amicon® Ultra Centrifugal Filters) before loading them on a superdex 200 16/600 column which had been pre-equilibrated in refolding buffer via a 2 ml loop. Fractions of 750 µl were collected and checked for stoichiometry on SDS-PAGE NuPAGE™ 4-12 % Bis-Tris Gel 1 mm (MES buffer, invitrogen). Dimers were spin-concentrated and stored in 50 % glycerol at −20 °C.

4.5 Structural studies of INO80 on 0N80 nucleosomes

INO80 was mixed 1:1 with 0N80 at a concentration of 600 nM in a buffer containing 20 mM HEPES pH8, 50 mM KCl, 0.25 mM, $CaCl_2$ and 0.25 mM DDT. INO80-nucleosome sample was plunge frozen on a Vitrobot Mark IV (ThermoFisher) in liquid ethane. For this purpose, the sample was mixed with 0.05 % of the detergent β-octyl-glycoside. Further, Quantifoil copper mesh 200 R2/1 grids were plasma cleaned with the FISHIONE Instrument Nanoclean in the presence of 90 % argon and 10 % oxygen for 15 s. Vitrobot was set to 100 % humidity, 4 °C, wait time 0 s, blot time of 2 or 3 s and a blot force of 2 or 3. A total of 11000 micrographs was collected on a Titan Krios transmission electron microscope (TEM) from FEI at a magnification of 130 k which corresponds to a pixel size of 1.04 Å. Furthermore, images were illuminated with a total dose of 41.6 e/Å^2.
Data processing was handled throughout in RELION 4.0 [263] where micrographs were subjected to motion correction with the RELION implemented MotionCor2 followed by ctf estimation with Gctf. Particles were automatically picked in RELION with a threshold of 0.05 and an inter-particle distance of 250 Å. Particles were extracted with a diameter of 360 pixel whereas a binning factor of 2 was applied (new box size 180 pixel). Subsequently, particles were 2D classified and INO80-complexes were selected for the creation of an initial model. Next, the initial model was used in order to perform 3D classifications on the full dataset. In consecutive rounds of 3D classifications, apo INO80 classes were separated from the desired classes of nucleosome bound INO80. Particles were further 3D refined and subjected to per-micrograph ctf refinement, Bayesian polishing as well as per-particle ctf refinement. Finally, before performing the final 3D refinement, particles were re-extracted without binning. A total amount of 70091 particles were utilized for the final 3D refinement and subsequent post-processing.

4.6 Structural studies on INO80-hexasome

4.6.1 CryoEM sample preparation, data collection and data processing

INO80-hexasome samples were prepared by Min Zhang. Details on the sample preparation, data collection and performed data processing steps are outlined in the methods section of [281].

4.6.2 Model building

EM densities for state 1 were used for initial model building due to the better quality compared to state 2. Models of state 1 were then used as a starting point for state 2. Model building was performed in both DeepEMhancer [324] sharpened maps and unsharpened maps from the different focused refinements. As starting models, already published data was used: PDB 6FML for Arp5 core and Ies6; PDB 8AVG for Ino80 ATPase and insert as well as Rvb1/2, Arp5 grappler and Arp5 foot; PDB 7OHC for histone core and DNA; PDB 3I62 for the ADP•AlF_4^- configuration within the active site of the Ino80 motor lobes. Structures from the PDB were rigid body fitted in ChimeraX (version 1.4) [325] and next flexibly fitted in COOT (version 0.9.7) [275]. Within COOT, the protein models were rebuilt using self-restraints while I built the DNA beyond SHL−3 using German McClure restraints. Side chains were built except for the Arp5 grappler which was only built as poly-alanine model from residues 311-602 in state 1 (including the Arp5 foot) and residues 311-603 in state 2 (excluding Arp5 foot). The Arp5 foot interaction with the H4α2 helix was modelled using RosettaES [326] (done by Frank Dimaio) for analysis and visualization but was truncated in the PDB deposited model. The ADP•AlF_4^- within the active site was built according to PDB 3I62 with Ino80 residue threonine 1004 located at the water molecule position that coordinates the magnesium ion in the PDB. Next,

iterative runs of ISOLDE 1.4 [276] with self-restraints, followed by real-space refinements in PHENIX (version 1.20.1) [277] were performed. DNA restraints were separately generated in ACHESYM [327]. First, PHENIX refinements were done into the unsharpened maps. In the last step, sharpened maps from LocScale (version 2.1.5.) [328] were used instead. Lastly, composite models of both state 1 and 2 were generated from the individually refined local models by initially rigid body fitting the respective models in the composite EM densities. Then, chains were connected and side chains at map interfaces adapted. Composite models were used to generate a composite map in PHENIX for molecular dynamics adaptations in ISOLDE. Models were real-spaced refined with self-restraints one last time to generate the refinement statistics for the composite models. Statistics of the refined maps and built models is outlined in Table 4.1 for state 1 and in Table 4.2 for state 2.

INO80-core module hexasome (state1)	whole	Ino80 ATPase-hexasome	composite	Ino80 ATPase-DNA	Hexasome	Rvb1/Rvb2 Ino80 insert, Ies2	Arp5 core, Ies6	Arp5 grappler
PDB			8OO7	8OO9	8OOA	8OOC	8OOF	8OOK
EMDB	EMD-17019	EMD-17023	EMD-17006	EMD-17007	EMD-17008	EMD-17010	EMD-17012	EMD-17017
For each reconstruction								
Refined particles (no.)	72,400	72,400		72,400	72,400	72,400	72,400	72,400
Final particles (no.)	72,400	72,400		72,400	72,400	72,400	72,400	72,400
Symmetry imposed	C1	C1		C1	C1	C1	C1	C1
Resolution (global, Å, 3DFSC)	3.15	3.57		3.22	3.18	2.93	2.90	5.69
FSC 0.5 (unmasked/masked)	7.59/3.75	8.22/4.42		8.00/3.75	8.00/3.89	7.59/3.40	7.59/3.36	8.70/7.59
FSC 0.143 (unmasked/masked)	3.89/3.15	4.29/3.57		3.75/3.22	3.89/3.18	3.05/2.93	3.70/2.90	3.79/5.69
Resolution range (local, Å)	2.9-10.9	3.3-11.9		3.2-7.3	3.1-6.5	2.9-4.7	3.1-4.5	3.7-8.6
Map sharpening B factor ($Å^2$)	−65	−91		−89	−70	−51	−64	−155
3DFSC sphericity	0.881	0.774		0.795	0.795	0.923	0.875	0.744
Initial model used for model building (PDB code)				8AV6 3I62	7OHC	8AV6	6FML	8AV6
Model composition (for each model)								
Protein			4693	439	529	3626	556	199
Ligands (ADP, Mg, ATP)			7 ADP 1 ATP 2 Mg 1 AlF4	1 ADP 1 Mg 1 AlF4	0	6 ADP 1 ATP 1 Mg	1 ATP 1 Mg	0
DNA			226	42	204	–	–	–
Model refinement (for each model)								
Refinement package			Phenix 1.20.1-4487	Phenix 1.20.1-4487	Phenix 1.20.1-4487	Phenix 1.20.1-4487	Phenix 1.20.1-4487	Phenix 1.20.1-4487
real or reciprocal space			real	real	real	real	real	real
resolution cutoff			2.40	2.92	2.80	2.50	2.55	5.00
Model-Map-scores								
CC (mask)			0.81	0.78	0.76	0.86	0.84	0.71
CC (peaks)			0.73	0.70	0.67	0.81	0.75	0.46
CC (volume)			0.80	0.78	0.75	0.85	0.83	0.66
Map versus Model FSC 0.5 (Å)			2.8	3.2	3.1	2.7	2.7	8.5
B factors ($Å^2$)								
Protein residues			64.34	61.99	37.73	66.43	58.74	495.92
Ligands			46.52	57.92	–	47.65	39.05	–
DNA			65.94	41.69	51.87	–	–	–
R.m.s deviations from ideal values								
Bond lengths (Å)			0.005	0.010	0.008	0.003	0.005	0.005
Bond angles (°)			1.098	1.260	1.194	0.694	1.038	0.942
Validation (for each model)								
MolProbity score			1.10	1.19	0.99	1.09	1.21	0.72
CaBLAM outliers (%)			1.16	1.42	0.59	1.18	0.94	0
Clashscore			2.54	3.29	2.17	2.46	2.60	0.69
Poor rotamers (%)			0	0.0	0.0	0.0	0.0	–
C-beta outliers			0.0	0.0	0.0	0.0	0.0	0.0
EMRinger score (if better than 4 Å resolution)			3.75	3.72	3.22	3.73	4.92	–
Ramachandran plot								
Favored (%)			97.74	97.68	98.45	97.72	97.06	100
Allowed (%)			2.26	2.32	1.55	2.28	2.94	0
Outliers (%)			0	0	0	0	0	

Table 4.1: **Table with refinement and model building statistics for state 1 maps and models**. Table was adapted from the INO80-hexasome publication [281].

INO80-core module hexasome (state2)	whole	Arp5 grappler	composite	Ino80 ATPase-hexasome	Rvb1/Rvb2 Ino80 insert, Ies2	Arp5 core Ies6
PDB	–	–	8OOP	8OOS	8OOR	8OOT
EMDB	EMD-17029	EMD-17676	EMD-17025	EMD-17027	EMD-17026	EMD-17028
For each reconstruction						
Refined particles (no.)	98,967	98,967		98,967	98,967	98,967
Final particles (no.)	98,967	98,967		98,967	98,967	98,967
Symmetry imposed	C1	C1		C1	C1	C1
Resolution (global, Å, 3DFSC)	3.08	7.59		3.29	2.87	2.85
FSC 0.5 (unmasked/masked)	6.58/3.61	8.00/8.70		7.79/4.05	7.59/3.40	6.88/3.29
FSC 0.143 (unmasked/masked)	3.75/3.08	4.23/7.59		4.05/3.29	3.05/2.93	3.61/2.85
Resolution range (Å)	2.9-10	3.5-10.3		3.2-8.0	2.9-4.4	2.9-5.1
Map sharpening B factor ($Å^2$)	−62	−688		−83	−51	−65
3DFSC sphericity	0.907	0.871		0.828	0.944	0.852
Initial model used for model building (PDB code)				8AV6 3I62	8AV6	6FML
Model composition (for each model)						
Protein			4677	966	3626	556
Ligands (ADP, Mg, ATP)			7 ADP 2 Mg 1 ATP 1 AlF4	1 ADP 1 Mg 1 AlF4	6 ADP 1 Mg 1 ATP	1 Mg 1 ATP
DNA			217	217	0	0
Model refinement (for each model)						
Refinement package			Phenix 1.20.1-4487	Phenix 1.20.1-4487	Phenix 1.20.1-4487	Phenix 1.20.1-4487
real or reciprocal space			real	real	real	real
resolution cutoff			2.50	2.70	2.50	2.60
Model-Map scores						
CC (mask)			0.83	0.73	0.88	0.87
CC (peaks)			0.75	0.59	0.84	0.77
CC (volume)			0.82	0.72	0.88	0.87
Map versus Model FSC 0.5 (Å)			2.7	3.2	2.5	2.9
B factors ($Å^2$)						
Protein residues			68.62	67.24	63.79	93.26
Ligands			51.99	79.40	46.60	81.73
DNA			71.42	66.41	–	–
R.m.s deviations from ideal values						
Bond lengths (Å)			0.006	0.007	0.011	0.004
Bond angles (°)			1.190	1.167	0.959	1.005
Validation (for each model)						
MolProbity score			0.94	1.29	1.19	1.20
CaBLAM outliers (%)			1.31	1.84	1.35	1.13
Clashscore			1.28	3.08	2.70	2.03
Poor rotamers (%)			0	0	0	0
C-beta outliers			0	0	0	0
EMRinger score (if better than 4 Å resolution)			3.82	2.66	4.13	5.43
Ramachandran plot						
Favored (%)			97.58	96.83	97.27	96.51
Allowed (%)			2.42	3.17	2.73	3.49
Outliers (%)			0	0	0	0

Table 4.2: **Table with refinement and model building statistics for state 2 maps and models**. Table was adapted from the INO80-hexasome publication [281].

4.6.3 INO80 activity assays – ATP-dependent nucleosome sliding assays

Sliding assays were used to probe the INO80 activity on various nucleosome and hexasome constructs [77][67]. For this purpose 6-Carboxyfluorescein (6-FAM) labeled 0N80 nucleosomes, 0H80 or 0H40 hexasomes were used. 100 nM of nucleosomes or hexasomes were mixed with 30 nM of *C. t.* or *S. c.* INO80 in the presence of sliding buffer (25 mM HEPES at pH 8.0, 60 mM KCl, 0.1 mg/ml BSA

(bovine serum albumin), 7 % glycerol,0.25 mM DDT, 2 mM $MgCl_2$). Components were pre-incubated at 26 °C for 10 min before starting the sliding reaction with a final concentration of 2 mM $MgCl_2$ and 1 mM ATP. Probes of 10 µl were taken for each time point and mixed with 5 µl of Lambda DNA (Bio-Labs 500 µg/ml) to compete off INO80 and stop the reaction. Time points were taken at 30 s, 1 min, 2 min, 5 min, 10 min, 30 min, 60 min and 90 min. In addition, one sample prior to ATP and $MgCl_2$ incorporation was kept. Samples were analysed on 10-well Native PAGEs (Invitrogen NativePAGE^{TM} 4-16 % Bis-Tris Gel, 10-well 1.0 mm) and mixed for this purpose with 1 µl of 20 % Ficol solution. 15 µl per sample were loaded and gels ran at 100 V for 1.5 h , followed by another 0.5 h at 150 V. Gels with hexasome samples were prepared as follows: 100 V for 1.5 h then 1.5 h at 150 V to achieve good separation of remodeled versus non-remodeled hexasome bands. Gels were scanned on a Typhoon FLA9500 machine and analysed using ImageJ [331]. Corrections were applied to account for small differences in terms of sample loading and for hexasomes being present in nucleosome samples and reversely. To achieve this, the total sample amount was calculated by adding up the band intensities of non-remodeled, intermediate and remodeled bands. Next, the percentage of the non-remodeled substrate (%S) in relation to the total amount was determined. The minus ATP samples represents 0 % of remodeled nucleosome or hexasome samples. Based on its value, the percentage (%) of remodeled samples were calculated for all time points. Graphs were plotted in GraphPad Prism version 9.5.1 (Dotmatics). Figure 4.1 shows comparison of activity of INO80 on 0N80, 0H80 and 0H40 substrates.

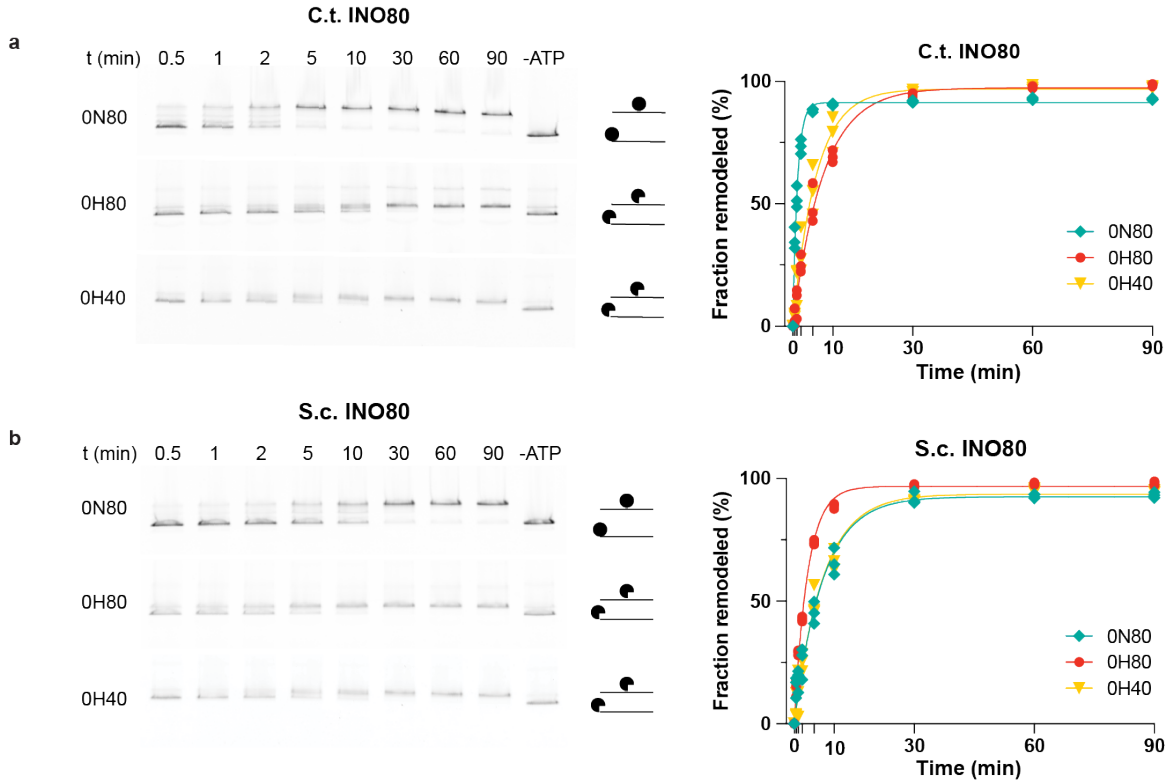


Figure 4.1: **Comparison of sliding activity of *C. t.*and *S. c.* INO80 on 0N80, 0H80 and 0H40. a** *Left:* Native PAGEs of sliding assays. *Right:* Graphs showing sliding activity of *C. t.* on 0N80, 0H80 and 0H40. 0N80 and 0H40 comprise the same amount of extra-nucleosomal DNA. Yet, INO80 slides 0N80 with the fastest kinetic. **b** The same representation and sliding assay setup but for *S. c.* INO80.

4.6.4 Nucleosome/Hexasome competition assays

Competition assays were performed and analysed likewise to the sliding assays. However, 0N80 nucleosomes and 0H40 hexasomes were used in the same reaction tube to determine the INO80 substrate preference. 0N80 nucleosomes were 6-FAM labeled, while 0H40 contained a Cy5.5 label. Substrates were pre-analysed on Native PAGE to precisely calculate molarities in relation to each other. The assay was run by mixing 100 nM of 0N80 and 0H40 substrates with 30 nM of yeast or *C. t.* INO80. Gels were run at 100 V for 1.5 h followed by another 1.5 h at 150 V. Finally, gels were scanned for 6-FAM and Cy5.5 to detect nucleosomes and hexasomes from the same gel.

4.7 Structural studies on Apo SWR1

4.7.1 Sample preparation

Apo SWR1 sample was neither concentrated nor crosslinked for cryoEM. Instead a fraction was directly taken from the anion exchange chromatography step that was performed at an ÄKTA micro setup. Quantifoil mesh 200 R2/1 grids were plasma cleaned (FISHIONE Instruments, Nanoclean) at a mixture of 10 % of oxygen and 90 % of argon for 15 s. Subsequently, 3 µl of Apo SWR1 at 600 nM were applied to the grid inside the humidity controlled FEI Vitrobot Mark IV (ThermoFisher) at 4 °C and 100 % humidity. Apo SWR1 was mixed directly prior to sample vitrification with β-octyl-glycoside at a final concentration of 0.05 %. Grids were blotted with Whatman blotting paper (Cytiva, CAT No. 10311807) without wait time, blot force of 2 and blot time of 2 s and plunge frozen in liquid ethane.

4.7.2 Data collection

Micrographs were collected on a Titan Krios electron microscope (FEI) harboring a K2 direct electron detector (from Gatan). Six shots per hole were recorded using beam tilt in an automated way within the SerialEM [329] resulting in a total number of 5064 micrographs. Images were generated at a spot size of 8, an exposure of 10 s and at a magnification of 130k which corresponds to a pixel size of 1.04 Å/pixel. The microscope settings resulted in a total electon dose of 47.34 e/A^2 that was distributed over 40 frames.

4.7.3 Data processing and model building

Particle picking was performed on the fly during data acquisition with the software WARP [330]. Pre-processing of the micrographs was done in RELION 4.0 [263] which included motion correction with the RELION implemented MotionCor2 as well as ctf estimation with Gctf. Particle coordinates of WARP were used to extract 619519 particles with a box size of 360 Å. Subsequently, motion corrected micrographs as well as extracted particles were transferred to cryoSPARC [264] to perform 2D classifications. Classes showed mainly the Core complex of SWR1 (168877 particles) while a sub-fraction represented the Arp module (35000 particles). In order to pick more particles, I trained separately the neural network based picking tool TOPAZ [268] for both modules. Training and picking was performed in three consecutive rounds. This resulted in 356180 Core and 183000 Arp module particles. A subset of 2D classes was utilized to generate ab-initio reconstructions within cryoSPARC. Particles of the Core module were further processed as follows: All TOPAZ picked particles from the last job were subjected to a heterogeneous refinement which created two garbage classes and one volume showing features of the Core module (422267 particles). All TOPAZ picked particles were utilized as an input to avoid removal of rare views at this stage. Another round of

heterogeneous refinement separated one class with the overall Core module (268443 particles) from a class that missed the Swc3 subunit. This step was repeated which resulted again in one class without Swc3 and one with the intact Core module. The resulting 166166 particles were transferred back to RELION by creating a star file with the new particle coordinates and refinement parameters. In addition, the 3D class of the previous heterogeneous refinement (268443 particles) was imported into RELION as an initial model. Next, particles were re-extracted in RELION and classified into one 3D class using the map from cryoSPARC as initial model. Particles were then subjected to 3D refinement using the 3D class as a reference (initial low pass filtered to 40 Å) and applying a mask. The resulting 3D map was of good quality except for the rim region of Swc3 and the Swc6/Arp6 arm. To generate an overall map of high quality first, I proceeded with post-processing, which included per-micrograph ctf refinement, Bayesian polishing as well as per-particle refinement yielding a map of an overall resolution of 2.97 Å. In order to improve the map density around the Swc6/Arp6 arm, particle re-centering onto the Swc6/Arp6 subunit was applied. For this purpose, a mask around the Swc6/Arp6 arm was generated and used to calculate the offset from the box center with help of the RELION image handler tool in the terminal. Based on these x, y and z values, particles were re-extracted and subjected to the same procedure of 3D refinement and post-processing workflow using masks for the whole Core module. However, the final 3D refinement was re-started at iteration 13. The resulting map showed a better connectivity at the rim of Arp6 and Swc6 and had an overall resolution of 3.04 Å.

Model building was performed based on published structures (Rvb1/Rvb2 from PDB 8AV6) or AlphaFold Multimer predictions [273]. The following AlphaFold predictions were generated: Swc2-Swc3, Swr1 as well as Swr1 anchor-Arp6-Swc6. The focused map around the Swc6/Arp6 arm was utilized for model building of Swc6, Arp6, the Swr1 anchoring helices and one of the Rvb2 OB-folds that sits close to Swc6. The remaining subunits and domains were built into the overall map. First, pdb models were rigid body fitted in ChimeraX [325], followed by flexible fitting in COOT [275] applying molecule self-restraints. Models were carefully inspected, Ramachandran and Rotamer outliers resolved. Due to time restrictions, I did not continue to improve the models in Chimera Isolde [276] and did not perform phenix refinement. However, model building will be continued for later publication. Likewise to the Core module of SWR1, TOPAZ picked particles of the Arp module as well as the initial model were exported to RELION. The initial model was used as a reference to classify the TOPAZ picked particles into one 3D class to generate a RELION internal 3D reference. The resulting density of both the initial model and the 3D class was of low resolution. It comprises the Arp module core as well as two arms. The map could not be improved by neither 3D classifications nor local refinements in cryoSPARC as well as RELION. Instead, one 3D refinement was performed. In order to apply an integrative approach of cryoEM and structure predictions, two AlphaFold multimer models were created: 1) Swr1 HSA helix, two copies of actin, Arp4 and Swc4 2) Swr1 N-term and HSA helix, one copy of actin, Arp4, Swc4 and Yaf9. Density of N-terminal region of Swr1, Yaf9 and Swc4 fitted well into the foot that extends from the Arp module core. However, AlphaFold did not predict the angle of the foot correctly. The round blob on top of the other site of the Arp module harbors the second copy of actin. In order to improve the fitting, three more AlphaFold models were generated that contained: 1) The Arp module core of Swr1 HSA helix (α1, N-terminal region of Swc4 including Swc4 joint and 2W hairpin, actin and Arp4 2) C-terminal region of Swc4, N-terminal region of Swr1 and Yaf9 3) actin and α2 helix of HSA. The three models were separately fitted into the 3D refined Arp module map. The interfaces do not fit perfectly as AlphaFold did not predict the conformations correctly but the fitting can only be improved when an EM map of higher resolution is obtained.

4.7.4 Single molecule analysis of Core and Arp module pairs

The single molecule pipeline was set up by Luis Hauptmann and performed as described in [281]. In short, Luis Hauptmann mapped nearest neighbor pairs of the final set of Core module (166166 particles) and Arp modules (183000 particles). Distances were plotted and showed a preferred distance of 173 Å. A selection of Core-Arp module pairs was created that comprise a distance of 90-210 Å. The outputted star file was imported into RELION and utilized to re-extract Core-Arp module pairs with a box size of 480 Å. Particles were subjected to 2D classification asking for 20 classes.

4.8 Structural studies on SWR1 Arp module

4.8.1 Sample preparation

The Arp module$^{\text{N-HSA FL}}$-Yaf9 sample was diluted to 4 µM and 50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM $CaCl_2$, 0.25 mM DDT. From this sample, 3 µl were applied to Quantifoil copper mesh 200 R2/1 grids in the presence of 0.05 % of the detergent β-octyl-glycoside and plunge frozen with the following settings at a Vitrobot Mark IV (ThermoFisher): 100 % humidity, 4 °C, wait time 0 s, blot time of 2 s and a blot force of 2. Grids were before plasma cleaned using the FISHIONE Instrument Nanoclean.

4.8.2 Data collection

Date was acquired at 300 keV on a Titan Krios transmission electron microscope (TEM) from FEI. The microscope was equipped with a K3 direct detector and an energy filter from Gatan set to a width of 20 eV. A total of 3315 micrographs with a defocus range of 0.8 to 1.8 were collected by imaging five shots per hole. I made use of automated data acquisition based on a beam-tilt acquisition scheme and hole references within SerialEM [329]. Furthermore, images were acquired at a magnification of 105000 which corresponds to an equilibrated pixel size of 0.822 Å. Micrographs were collected by an exposure time of 1.57 s and a dose of 46.39 e/A^2 spread over 40 frames. Finally, WARP [330] was utilized for pre-processing and particle picking.

4.8.3 Data processing and model building

Micrographs were again pre-processed in RELION 4.0 [263] making use of the RELION implemented motion correction tool MotionCor2 as well as Gctf for ctf estimation. WARP picking parameters of 1163571 particles were imported and used for particle extraction of the Arp module with a box size of 256 px. Subsequently, extracted particles and motion-corrected micrographs were imported into the data processing software cryoSPARC [264]. Arp modules were subjected to 2D classification which revealed three type of classes 1) Arp module 2) Yaf9 3) junk particles. Only two classes represented Yaf9 with a total number of 24717 particles. the identity of Yaf9 classes was clear from AlphaFold predictions. Yaf9 particles were used for three rounds of TOPAZ training and subsequent particle picking respectively. However, the number of Yaf9 particles could not be increased. Instead, initial model creation and non-uniform refinement were performed that were utilized to fit the AlphaFold predicted model. Back to the Arp module classes, 499450 particles resembling the Arp module were selected and re-imported together with the initial model to RELION. Based on an AlphaFold multimer prediction of Swr1$^{\text{HSA}}$-Arp4-actin-Swc4, it was apparent that the initial model was mirrored. Therefore, the underlying volume was volume flipped in ChimeraX [325]. Particles were re-extracted in RELION this time with an adapted smaller box size of 180 px which corresponds

to 148 Å. Using the initial model from cryoSPARC, Arp modules were classified into one 3D class to obtain a RELION internal reference. During this process RELION 5.0 [252] was released which has a novel implementation of Blush regularisation and is based on a denoising convolutional neuronal network [252]. Thereby, 3D classification and 3D refinements of smaller particles, like in the presented case, can be improved. To make use of the blush option, I imported the particles and 3D class for reference to RELION 5.0 were the remaining steps were performed. I created a soft mask (volume for mask creation lowpass filtered to 15 Å, extended binary map and soft-edge by 3 px respectively) for subsequent 3D classification (with blush option) asking for three classes. One class was showing a weaker density around the Swc4-joint, therefore the other two classes were combined and subjected to another round of 3D classification again asking for three classes. The best two classes were combined resulting in a total number of 89998 Arp module particles. During the inspection of 3D classes, I carefully compared the volumes to the AlphaFold model as in my hands the blush function seemed to hallucinate density in weaker resolved areas. The final set of particles was refined applying the same mask and reference (initial lowpass filter to 15 Å, use blush regularisation). Next, post-processing, per micrograph ctf refinement, Bayesian polishing as well as per particle ctf refinement were applied in this order. The final post-processed map is at 3.44 Å resolution. However, I preferred to use the final 3D refinement after per particle ctf refinement which was at 3.79 Å. For model building, I fitted the mentioned AlphaFold Multimer predictions within ChimeraX [325] into the final density. Next, I used the flexible fitting option of COOT [275] to refine the model applying self-restraints. Resulting ramachandran and rotamer outliers were fixed and the model was carefully inspected. For time restraints, I did not perform further model improvement or refinements yet, but will do so for later publication of the results.

4.9 Towards a SWR1-nucleosome structure

CryoEM samples were prepared with either 60N25, 80N25 or 25N80 nucleosomes. I aimed at complex concentrations from 600 to 900 nM. For this purpose, SWR1, nucleosomes as well as H2A.Z/H2B dimers were mixed at a 1:1:1 molar ratio. In cases where dialysis of the complex was performed, SWR1 and H2A.Z/H2B were premixed at a high salt concentration (about 500 mM KCl). Next, dilution buffer (50 mM HEPES pH 7.5, 0.25 mM DDT, 0.5 mM $CaCl_2$, 20 µM $ZnCl_2$) and finally the nucleosomes were added. Absorption was measured prior and post dialysis to calculate the concentration. Typically the dialysis of the complex was prepared with the highest complex concentration possible and diluted after dialysis if necessary. Dialysis was performed in 500 µl Slide-A-Lyzer MINI Dialysis Units (Thermo Scientific) against 300 ml of dialysis buffer (50 mM HEPES pH 7.5, 75 mM KCl, 20 µM $ZnCl_2$, 0.5 mM $CaCl_2$, 0.25 mM DDT) for 0.5-1 h. Calciumchloride was left out when nucleotide analogues were added. If SWR1 and H2A.Z/H2B were pre-dialysed in the absence of nucleosomes, they were mixed in dilution buffer and subjected to the same dialysis procedure. If the nucleotide analogues ADP•AlF_4^- or ADP•BeF_3^- were tested, they were added post-dialysis and right before sample preparation for cryoEM. In both cases 10x pre-mixtures of 20 mM ADP/33 mM $MgCl_2$ and 100 mM NaF/20 mM $AlCl_3$ or 160 mM NaF/20 mM BeF_2 were prepared and added to the complex sample to reach a final 1x dilution. Further, 0.05 % of the detergent β-ctyl-glycoside or Tween20 at 0.005 % was added right before sample vitrification. Samples were prepared on 1) Quantifoil copper mesh 200 R2/1 2) Quantifoil copper mesh 200 R2/1 2 nm Carbon or 3) Quantifoil gold mesh 200 R2/1 grids. Grids were plasma cleaned in the FISHIONE Instrument Nanoclean with the following settings: Gold grids: 75 % argon, 25 % oxygen, 45 s, Copper grids: 90 % argon, 10 % oxygen, 15 s. Grids were plunge frozen with the help of the Vitrobot Mark IV (ThermoFisher) set to 100 % humidity, 4 °C, wait time 0 s, blot time of 2 s and a blot force of 2 or 3. Copper grids with carbon support were pre-

pared at lower complex concentration (80 nM but with a wait time prior to vitrification of 1-2 min). Grids were inspected at a FEI Talos Arctica TEM (200 kV, Falcon 3 camera) and small datasets of a few thousand micrographs were collected on a Titan Krios transmission electron microscope (TEM) from FEI equipped with a K3 direct detector from Gatan. Pre-processing of micrographs (motion correction and ctf estimation) was performed in RELION 4.0 [263] likewise to particle extraction of WARP picked [330] particles. 2D classifications were performed in cryoSPARC [264]. These revealed that complexes of SWR1 and nucleosomes were falling apart in the tested conditions.

4.10 AlphaFold Multimer predictions

AlphaFold Multimer predictions were performed with the publicly available version 2.3.2 [273]. Thomas Hoffmann created scripts for running subsequent MSAs (multiple sequence alignments), feature prediction, a finalizing job as well as model relaxation. I created 25 models per prediction and compared them, however I used only the model with the best scores ($ranked_0$). PAE plots were inspected as quality control and the models were colored in ChimeraX based on their confidence score of AlphaFold (stored and accessible as B factor value).

4.11 Chaetomium pulldowns

In brief, the purified bait proteins eGFP (negative control), SWR1-complex, Swc5 or Bdf1 were used to pull down interacting subunits from *Chaetomium thermophilum*. Experiments were prepared in triplicates and TMT-labeled by the mass spectrometry facility for subsequent analysis. The following three experiments were performed: 1) eGFP and Swc5 2) eGFP and Bdf1 3) eGFP, Bdf1 and SWR1-complex.
In the initial setup, the bait proteins were purified on the fly. Per replicate 0.5 g of insect cell pellets of the respective bait protein were lysed and immobilized on 50 µl of anti-FLAG beads (GenScript, Cat. L00432). After washing of the beads, Chaetomium lysate was added. In an optimized setup, bait proteins were purified to homogeneity and then immoblized in equimolar amounts on anti-FLAG beads. For experiments 1) and 2) 9.15 µM and for experiment 3) 2 µM bait proteins were added each to 50 µl anti-FLAG affinity beads that had been equilibrated in lysis buffer (50 mM HEPES pH7, 500 mM NaCl, 10 % glycerol, 0.25 mM DDT and 20 µM $ZnCl_2$). Bait proteins were immobilized for about 1 hour on a rolling wheel at 4 °C. Beads where washed twice in lysis buffer. In the meantime, 10 g of freeze-milled *Chaetomium thermophilum* lysate (prepared by the Häring lab in Würzburg) was dissolved in 25 ml of lysis buffer supplemented with 0.05 % NP-40. Samples were fully dissolved on a rolling wheel for about one hour. Samples were then mildly sonicated at 40 % intensity for 2 min (Sonifier® W-250 D). Lysate was cleared at 30 000 g at 4 °C for 45 min in a Ti-40 rotor before adding it to the prepared FLAG beads. The lysate was incubated with the FLAG beads for 3 h in the cold room on a rolling wheel. FLAG beads were transferred into Eppendorf tubes and washed ten times in 1 ml wash buffer (50 mM HEPES pH7, 150 mM NaCl, 5 % glycerol, 0.25 mM DDT and 20 µM $ZnCl_2$, 0.5 mM $CaCl_2$). Proteins were eluted off the beads with 60 µl of 2*x* SDS buffer. While 50 µl were handed over to Jennifer Schwarz from the EMBL proteomics core facility, 10 µl were analysed on SDS-PAGE. Mass spectrometry data was analysed as follows. Enriched hits and enriched candidates (compared to the eGFP control) were further analysed by checking their identity using the uniprot database. Hits for the transcription and translation machinery were not followed up. Remaining interesting hits were checked on the AlphaFold database and compared to the AlphaFold prediction of Swc7. Plots of results were generated in R studio using and adapting scripts provided by Thomas Dahlet.

4.12 INO80-Swc5 crosslinking studies

Testing of different crosslinking conditions
To test different BS3 concentrations, 5 µg of ΔN-INO80 was mixed with an equimolar amount of Swc5 in crosslinking buffer (30 mM HEPES pH 7.8, 150 mM NaCl, 0.5 mM $CaCl_2$ and 0.25 mM DDT). A stock of 50 mM BS3 was prepared in water. The following BS3 conditions were tested: 0 mM, 0.025 mM, 0.05 mM, 0.075 mM, 0.11 mM, 0.125 mM, 0.25 mM, 0.5 mM and 1 mM. The BS3 was added and crosslinking allowed for 30 min shaking at 350 rpm at room temperature. The reaction was stopped by adding Tris-HCl pH8 at a final concentration of 100 mM and incubation at 600 rpm and 35 °C for 10 min. Samples were then mixed with 5 µl 5*x* SDS buffer and analysed on a SDS-PAGE NuPAGE^{TM} 4-12 % Bis-Tris Gel 1 mm.

Large Scale Crosslinking experiment and mass spectrometry
For the final crosslinking mass spectrometry experiment, 55 µg of full-length INO80 (received from the Hopfner lab in Munich) were mixed with an equimolar amount of Swc5 and crosslinking buffer (see above). Crosslinking was enabled by addition of 0.25 mM BS3 for 30 min at room temperature while shaking at 350 rpm. The reaction was quenched with a final concentration of 100 mM Tris-HCl pH 8.0. An amount of 5 µg of the reaction was analysed on SDS-PAGE. The remaining sample was submitted to mass spectrometry analysis by Mandy Rettel of the EMBL proteomics core facility. Crosslinking data was visualized with xiView [293].

4.13 SWR1 crosslinking mass spectrometry

Crosslinking was performed likewise to the INO80-Swc5 crosslinking. However, a BS3 concentration of 0.15 mM was chosen for crosslinking. Peptides were separated on a 120 min gradient to enable detection of more peptides in the mass spectrometer.

4.14 Mass photometry measurements

Mass photometry measurements were performed on a Refeyn Two MP mass photometer machine. Samples or complexes were prepared at an initial concentration of 200 nM. SWR1 samples and complexes were mixed in filtered buffer (through 0.2 µm filter) of 20 mM HEPES pH7.5, 50-200 mM KCl, 20 µM $ZnCl_2$, 0.25 mM DDT. A respective buffer volume of 19 µl (without the DDT) was applied to the Refeyn for calibration before adding 1 µl of the protein sample. Depending on the used amounts, the protein sample was thereby diluted 1:18 or 1:19. SWR1-nucleosome and SWR1-nucleosome-H2A.Z/H2B complexes were analysed at 1:1:1 ratios.
Arp module constructs were prepared likewise at 200 nM but in a buffer of 50 mM HEPES pH7.0, 75 or 200 mM NaCl and 0.25 mM DDT. The effect of 1 mM of either ADP or ATP in the presence of 0.25 mM DDT was tested.

4.15 Applied software

Software	Publisher
UCSF ChimeraX, version 1.5-1.7, ISOLDE plugin	UCSF Resource for Biocomputing, Visualization and Informatics [325] [276]
WARP	[330]
SerialEM	[329]
CryoSPARC, version 3.3 and 4.0	Structura Biotechnology [264]
RELION, version 3.0, 4.0, 5.0	[263] [252]
COOT (Crystallographic Object-Oriented Toolkit, version 0.9.7)	Medical Research Council Laboratory of Molecular Biology (MRC-LMB) [275]
AlphaFold2 Multimer (2.3.2)	[273] [272]
PHENIX (version 1.20.1)	[277]
Jalview	[284][285]
Clustal Omega	European Molecular Biology Laboratory (EMBL-EBI), [287]
xiView	Rappsilber lab, [293]
AcquireMP, DiscoverMP, version 2.3	Refeyn
Prism	Graphpad, Dotmatics
RStudio	Posit Software, PBC
SnapGene	GSL Biotech
Fiji ImageJ	[331]
Adobe Illustrator®	Adobe

Table 4.3: Applied software and respective publishers.

Bibliography

[1] Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).

[2] Zhou, C. Y., Johnson, S. L., Gamarra, N. I. & Narlikar, G. J. Mechanisms of ATP-Dependent Chromatin Remodeling Motors. *Annual Review of Biophysics* **45**, 153–181 (2016).

[3] Shrader, T. E. & Crothers, D. M. Effects of DNA sequence and histone-histone interactions on nucleosome placement. *Journal of Molecular Biology* **216**, 69–84 (1990).

[4] Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Quarterly Reviews of Biophysics* **34**, 269–324 (2001).

[5] Millán-Zambrano, G., Burton, A., Bannister, A. J. & Schneider, R. Histone post-translational modifications — cause and consequence of genome function. *Nature Reviews Genetics* **23**, 563–580 (2022).

[6] Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **21**, 381–395 (2011).

[7] Garcia-Ramirez, M., Rocchini, C. & Ausio, J. Modulation of chromatin folding by histone acetylation. *Journal of Biological Chemistry* **270**, 17923–17928 (1995). URL `http://dx.doi.org/10.1074/jbc.270.30.17923`.

[8] Görisch, S. M., Wachsmuth, M., Tóth, K. F., Lichter, P. & Rippe, K. Histone acetylation increases chromatin accessibility. *Journal of Cell Science* **118**, 5825–5834 (2005).

[9] Lai, W. K. & Pugh, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature Reviews Molecular Cell Biology* **18**, 548–562 (2017). URL `http://dx.doi.org/10.1038/nrm.2017.47`.

[10] Martire, S. & Banaszynski, L. A. The roles of histone variants in fine-tuning chromatin organization and function (2020).

[11] Giaimo, B. D., Ferrante, F., Herchenröther, A., Hake, S. B. & Borggrefe, T. The histone variant H2A.Z in gene regulation. *Epigenetics and Chromatin* **12**, 1–22 (2019).

[12] McGinty, R. K. & Tan, S. Nucleosome structure and function (2015).

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

[14] McGinty, R. K. & Tan, S. Principles of nucleosome recognition by chromatin factors and enzymes. *Current Opinion in Structural Biology* **71**, 16–26 (2021).

[15] Skrajna, A. *et al.* Comprehensive nucleosome interactome screen establishes fundamental principles of nucleosome binding. *Nucleic Acids Research* **48**, 9415–9432 (2020).

[16] Kalashnikova, A. A., Porter-Goff, M. E., Muthurajan, U. M., Luger, K. & Hansen, J. C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *Journal of the Royal Society Interface* **10** (2013).

[17] McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014). URL `http://dx.doi.org/10.1038/nature13890`.

[18] Baldi, S., Korber, P. & Becker, P. B. Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nature Structural and Molecular Biology* **27**, 109–118 (2020).

[19] Eustermann, S., Patel, A. B., Hopfner, K. P., He, Y. & Korber, P. Energy-driven genome regulation by ATP-dependent chromatin remodellers. *Nature Reviews Molecular Cell Biology* (2023).

[20] Lai, B. *et al.* Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. *Nature* **562**, 281–285 (2018). URL `http://dx.doi.org/10.1038/s41586-018-0567-3`.

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

[22] Valouev, A. *et al.* Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516–522 (2011).

[23] Zhang, T., Zhang, W. & Jiang, J. Genome-wide nucleosome occupancy and positioning and their impact on gene expression and evolution in plants. *Plant Physiology* **168**, 1406–1416 (2015).

[24] Krietenstein, N. *et al.* Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709–721.e12 (2016).

[25] 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).

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

[27] Bao, Y. & Shen, X. SnapShot: Chromatin Remodeling Complexes. *Cell* **129**, 632.e1–632.e2 (2007).

[28] Wang, J. *et al.* Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Research* **22**, 1798–1812 (2012).

[29] Iyer, V. & Struhl, K. Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO Journal* **14**, 2570–2579 (1995).

[30] Lorch, Y., Maier-Davis, B. & Kornberg, R. D. Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes and Development* **28**, 2492–2497 (2014).

[31] Kubik, S. *et al.* Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. *Molecular Cell* **71**, 89–102 (2018).

[32] Haberle, V. *et al.* Two independent transcription initiation codes overlap on vertebrate core promoters. *Nature* **507**, 381–385 (2014).

[33] Ibarra-Morales, D. *et al.* Histone variant H2A.Z regulates zygotic genome activation. *Nature Communications* **12**, 1–14 (2021).

[34] Reyes, A. A., Marcum, R. D. & He, Y. Structure and Function of Chromatin Remodelers. *Journal of Molecular Biology* **433**, 166929 (2021). URL `https://doi.org/10.1016/j.jmb.2021.166929`.

[35] Yin, Y. *et al.* A basal-level activity of ATR links replication fork surveillance and stress response. *Molecular Cell* **81**, 4243–4257.e6 (2021). URL `https://doi.org/10.1016/j.molcel.2021.08.009`.

[36] Oberbeckmann, E. *et al.* Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nature Communications* **12** (2021).

[37] Yen, K., Vinayachandran, V. & Pugh, B. F. XSWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246 (2013).

[38] Jiang, C. & Pugh, B. F. Nucleosome positioning and gene regulation: Advances through genomics. *Nature Reviews Genetics* **10**, 161–172 (2009).

[39] Ocampo, J., Chereji, R. V., Eriksson, P. R. & Clark, D. J. The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. *Nucleic Acids Research* **44**, 4625–4635 (2016).

[40] Gkikopoulos, T. *et al.* A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. *Science* **333**, 1758–1760 (2011).

[41] Compton, J. L., Bellard, M. & Chambon, P. Biochemical evidence of variability in the DNA repeat length in the chromatin of higher eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 4382–4386 (1976).

[42] Godde, J. S. & Widom, J. Chromatin structure of Schizosaccharomyces pombe. A nucleosome repeat length that is shorter than the chromatosomal DNA length. *Journal of Molecular Biology* **226**, 1009–1025 (1992).

[43] Horowitz, R. A., Agard, D. A., Sedat, J. W. & Woodcock, C. L. The three-dimensional architecture of chromatin in situ: Electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. *Journal of Cell Biology* **125**, 1–10 (1994).

[44] Fyodorov, D. V., Zhou, B. R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nature Reviews Molecular Cell Biology* **19**, 192–206 (2018). URL `http://dx.doi.org/10.1038/nrm.2017.94`.

[45] Robinson, P. J., Fairall, L., Huynh, V. A. & Rhodes, D. EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6506–6511 (2006).

[46] Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016). URL `http://dx.doi.org/10.1038/nature16496`.

[47] Ricci, M. A., Manzo, C., García-Parajo, M. F., Lakadamyali, M. & Cosma, M. P. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* **160**, 1145–1158 (2015).

[48] Suto RK, MJ, C., DJ, T. & K., L. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nature structural & molecular biology* **19**, 31–38 (2014).

[49] Faast, R. *et al.* Histone variant H2A.Z is required for early mammalian development. *Current Biology* **11**, 1183–1187 (2001).

[50] Clarkson, M. J., Wells, J. R., Gibson, F., Saint, R. & Tremethick, D. J. Regions of variant histone His2AvD required for Drosophila development. *Chemtracts* **13**, 683–688 (2000).

[51] Zhang, H., Roberts, D. N. & Cairns, B. R. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123**, 219–231 (2005).

[52] Albert, I. *et al.* Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature* **446**, 572–576 (2007).

[53] Raisner, R. M. *et al.* Histone variant H2A.Z Marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**, 233–248 (2005).

[54] Lewis, T. S., Sokolova, V., Jung, H., Ng, H. & Tan, D. Structural basis of chromatin regulation by histone variant H2A.Z. *Nucleic Acids Research* **49**, 11379–11391 (2021).

[55] Ramachandran, S., Ahmad, K. & Henikoff, S. Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates. *Molecular Cell* **68**, 1038–1053.e4 (2017).

[56] Arimura, Y., Tachiwana, H., Oda, T., Sato, M. & Kurumizaka, H. Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome. *Biochemistry* **51**, 3302–3309 (2012).

[57] Zlatanova, J., Bishop, T. C., Victor, J. M., Jackson, V. & van Holde, K. The Nucleosome Family: Dynamic and Growing. *Structure* **17**, 160–171 (2009). URL `http://dx.doi.org/10.1016/j.str.2008.12.016`.

[58] Baer, B. W. & Rhodes, D. Eukaryotic RNA polymerase II binds to nucleosome cores from transcribed genes. *Nature* **301**, 482–488 (1983).

[59] Bilokapic, S., Strauss, M. & Halic, M. Histone octamer rearranges to adapt to DNA unwrapping. *Nature Structural and Molecular Biology* **25**, 101–108 (2018). URL `http://dx.doi.org/10.1038/s41594-017-0005-5`.

[60] Rhee, H. S., Bataille, A. R., Zhang, L. & Pugh, B. F. Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell* **159**, 1377–1388 (2014). URL `http://dx.doi.org/10.1016/j.cell.2014.10.054`.

[61] Henikoff, S. Mechanisms of nucleosome dynamics in vivo. *Cold Spring Harbor Perspectives in Medicine* **6** (2016).

[62] Ehara, H., Kujirai, T., Shirouzu, M., Kurumizaka, H. & Sekine, S. I. Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. *Science* **377** (2022).

[63] Farnung, L., Ochmann, M., Garg, G., Vos, S. M. & Cramer, P. Structure of a backtracked hexasomal intermediate of nucleosome transcription. *Molecular Cell* **82**, 3126–3134.e7 (2022).

[64] Bortvin, A. & Winston, F. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* **272**, 1473–1476 (1996).

[65] Liu, Y. *et al.* FACT caught in the act of manipulating the nucleosome. *Nature* **577**, 426–431 (2020). URL `http://dx.doi.org/10.1038/s41586-019-1820-0`.

[66] Mayanagi, K. *et al.* Structural visualization of key steps in nucleosome reorganization by human FACT. *Scientific Reports* **9**, 1–14 (2019).

[67] Hsieh, L. J. *et al.* A hexasome is the preferred substrate for the INO80 chromatin remodeling complex, allowing versatility of function. *Molecular Cell* **82**, 2098–2112.e4 (2022).

[68] Levendosky, R. F., Sabantsev, A., Deindl, S. & Bowman, G. D. The Chd1 chromatin remodeler shifts hexasomes unidirectionally. *eLife* **5**, 1–24 (2016).

[69] Lorch, Y., Maier-Davis, B. & Kornberg, R. D. Histone Acetylation Inhibits RSC and Stabilizes the +1 Nucleosome. *Molecular Cell* **72**, 594–600.e2 (2018).

[70] Singleton, M. R., Dillingham, M. S. & Wigley, D. B. Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual Review of Biochemistry* **76**, 23–50 (2007).

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

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

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

[74] Lessard, J. A. & Crabtree, G. R. Chromatin regulatory mechanisms in pluripotency. *Annual Review of Cell and Developmental Biology* **26**, 503–532 (2010).

[75] 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* **18**, 407–422 (2017).

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

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

[78] Gurard-Levin, Z. A., Quivy, J. P. & Almouzni, G. Histone chaperones: Assisting histone traffic and nucleosome dynamics. *Annual Review of Biochemistry* **83**, 487–517 (2014).

[79] Torigoe, S. E., Urwin, D. L., Ishii, H., Smith, D. E. & Kadonaga, J. T. Identification of a Rapidly Formed Nonnucleosomal Histone-DNA Intermediate that Is Converted into Chromatin by ACF. *Molecular Cell* **43**, 638–648 (2011). URL `http://dx.doi.org/10.1016/j.molcel.2011.07.017`.

[80] Fei, J. *et al.* The prenucleosome, a stable conformational isomer of the nucleosome. *Genes and Development* **29**, 2563–2575 (2015).

[81] Corona, D. F. *et al.* ISWI is an ATP-dependent nucleosome remodeling factor. *Molecular Cell* **3**, 239–245 (1999).

[82] Varga-Weisz, P. D. *et al.* Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**, 598–602 (1997).

[83] Lusser, A., Urwin, D. L. & Kadonaga, J. T. Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nature Structural and Molecular Biology* **12**, 160–166 (2005).

[84] Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R. & Kadonaga, J. T. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**, 145–155 (1997).

[85] Baldi, S., Krebs, S., Blum, H. & Becker, P. B. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. *Nature Structural and Molecular Biology* **25**, 894–901 (2018). URL `http://dx.doi.org/10.1038/s41594-018-0110-0`.

[86] Singh, A. K., Schauer, T., Pfaller, L., Straub, T. & Mueller-Planitz, F. The biogenesis and function of nucleosome arrays. *Nature Communications* **12** (2021). URL `http://biorxiv.org/cgi/content/short/2021.02.10.429500v1?rss=1{&}utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound`.

[87] Kaplan, C. D., Laprade, L. & Winston, F. Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301**, 1096–1099 (2003).

[88] Aalfs, J. D., Narlikar, G. J. & Kingston, R. E. Functional Differences between the Human ATP-dependent Nucleosome Remodeling Proteins BRG1 and SNF2H. *Journal of Biological Chemistry* **276**, 34270–34278 (2001). URL `http://dx.doi.org/10.1074/jbc.M104163200`.

[89] Yang, J. G., Madrid, T. S., Sevastopoulos, E. & Narlikar, G. J. The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nature Structural and Molecular Biology* **13**, 1078–1083 (2006).

[90] Kagalwala, M. N., Glaus, B. J., Dang, W., Zofall, M. & Bartholomew, B. Topography of the ISW2-nucleosome complex: Insights into nucleosome spacing and chromatin remodeling. *EMBO Journal* **23**, 2092–2104 (2004).

[91] Vary Jr., J. C. *et al.* Yeast Isw1p Forms Two Separable Complexes In Vivo. *Molecular and Cellular Biology* **23**, 80–91 (2003).

[92] Tsukiyama, T., Palmer, J., Landel, C. C., Shiloach, J. & Wu, C. Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes and Development* **13**, 686–697 (1999).

[93] Stockdale, C., Flaus, A., Ferreira, H. & Owen-Hughes, T. Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *Journal of Biological Chemistry* **281**, 16279–16288 (2006).

[94] Gangaraju, V. K. & Bartholomew, B. Dependency of ISW1a Chromatin Remodeling on Extranucleosomal DNA. *Molecular and Cellular Biology* **27**, 3217–3225 (2007).

[95] Xiao, H. *et al.* Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Molecular Cell* **8**, 531–543 (2001).

[96] Eberharter, A. *et al.* Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *EMBO Journal* **20**, 3781–3788 (2001).

[97] Fyodorov, D. V., Blower, M. D., Karpen, G. H. & Kadonaga, J. T. Acf1 confers unique activities to ACF/CHRAC and promotes the formation rather than disruption of chromatin in vivo. *Genes and Development* **18**, 170–183 (2004).

[98] Wallrath, L. L. & Elgin, S. C. Position effect variegation in Drosophila is associated with an altered chromatin structure. *Genes and Development* **9**, 1263–1277 (1995).

[99] Whitehouse, I., Rando, O. J., Delrow, J. & Tsukiyama, T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031–1035 (2007).

[100] Parnell, T. J., Schlichter, A., Wilson, B. G. & Cairns, B. R. The chromatin remodelers RSC and ISW1 display functional and chromatin-based promoter antagonism. *eLife* **2015**, 1–21 (2015).

[101] Clapier, C. R. *et al.* Critical Role for the Histone H4 N Terminus in Nucleosome Remodeling by ISWI. *Molecular and Cellular Biology* **21**, 875–883 (2001).

[102] Clapier, C. R., Nightingale, K. P. & Becker, P. B. A critical epitope for substrate recognition by the nucleosome remodeling ATPase ISWI. *Nucleic Acids Research* **30**, 649–655 (2002). NIHMS150003.

[103] Hamiche, A., Kang, J. G., Dennis, C., Xiao, H. & Wu, C. Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14316–14321 (2001).

[104] Grüne, T. *et al.* Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Molecular Cell* **12**, 449–460 (2003).

[105] Leonard, J. D. & Narlikar, G. J. A Nucleotide-Driven Switch Regulates Flanking DNA Length Sensing by a Dimeric Chromatin Remodeler. *Molecular Cell* **57**, 850–859 (2015). URL http://dx.doi.org/10.1016/j.molcel.2015.01.008.

[106] Armache, J. P. *et al.* Cryo-EM structures of remodeler-nucleosome intermediates suggest allosteric control through the nucleosome. *eLife* **8** (2019).

[107] Yan, L., Wu, H., Li, X., Gao, N. & Chen, Z. Structures of the ISWI–nucleosome complex reveal a conserved mechanism of chromatin remodeling. *Nature Structural and Molecular Biology* **26**, 258–266 (2019).

[108] Li, L. *et al.* Structure of the ISW1a complex bound to the dinucleosome. *Nature Structural and Molecular Biology* **31**, 266–274 (2024). URL https://www.biorxiv.org/content/10.1101/2023.01.02.522444v1?rss=1{&}utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound.

[109] 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). URL http://dx.doi.org/10.1016/j.molcel.2010.08.012.

[110] Tran, H. G., Steger, D. J., Iyer, V. R. & Johnson, A. D. The chrome domain protein Chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *EMBO Journal* **19**, 2323–2331 (2000).

[111] Kunert, N. & Brehm, A. Novel Mi-2 related ATP-dependent chromatin remodelers. *Epigenetics* **4**, 209–211 (2009).

[112] Denslow, S. A. & Wade, P. A. The human Mi-2/NuRD complex and gene regulation. *Oncogene* **26**, 5433–5438 (2007).

[113] Murawska, M. & Brehm, A. CHD chromatin remodelers and the transcription cycle. *Transcription* **2**, 244–253 (2011).

[114] Konev, A. Y. *et al.* CHD1 Motor Protein Is Required for. *Science* **1087**, 1–5 (2007).

[115] Allen, H. F., Wade, P. A. & Kutateladze, T. G. The NuRD architecture. *Cellular and Molecular Life Sciences* **70**, 3513–3524 (2013).

[116] Qiu, Y. *et al.* The Chd1 Chromatin Remodeler Shifts Nucleosomal DNA Bidirectionally as a Monomer. *Molecular Cell* **68**, 76–88.e6 (2017). URL https://doi.org/10.1016/j.molcel.2017.08.018.

[117] Rando, O. J. & Winston, F. Chromatin and transcription in yeast. *Genetics* **190**, 351–387 (2012).

[118] Fan, H. Y., He, X., Kingston, R. E. & Narlikar, G. J. Distinct strategies to make nucleosomal DNA accessible. *Molecular Cell* **11**, 1311–1322 (2003).

[119] Whitehouse, I. *et al.* Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400**, 784–787 (1999).

[120] Schubert, H. L. *et al.* Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 3345–3350 (2013).

[121] Han, Y., Reyes, A. A., Malik, S. & He, Y. Cryo-EM structure of SWI/SNF complex bound to a nucleosome. *Nature* **579**, 452–455 (2020). URL http://dx.doi.org/10.1038/s41586-020-2087-1.

[122] Ye, Y. *et al.* Structure of the RSC complex bound to the nucleosome. *Science* **366**, 838–843 (2019).

[123] He, S. *et al.* Structure of nucleosome-bound human BAF complex. *Science (New York, N.Y.)* **9761**, 1–12 (2020).

[124] Yuan, J., Chen, K., Zhang, W. & Chen, Z. Structure of human chromatin-remodelling PBAF complex bound to a nucleosome. *Nature* **605**, 166–171 (2022). URL http://dx.doi.org/10.1038/s41586-022-04658-5.

[125] Chen, K., Yuan, J., Sia, Y. & Chen, Z. Mechanism of action of the SWI/SNF family complexes. *Nucleus* **14** (2023). URL https://doi.org/10.1080/19491034.2023.2165604.

[126] Euskirchen, G., Auerbach, R. K. & Snyder, M. SWI/SNF chromatin-remodeling factors: Multiscale analyses and diverse functions. *Journal of Biological Chemistry* **287**, 30897–30905 (2012). URL http://dx.doi.org/10.1074/jbc.R111.309302.

[127] 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* **80**, 712–725.e5 (2020). URL `https://doi.org/10.1016/j.molcel.2020.09.024`.

[128] Cairns, B. R. *et al.* RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87**, 1249–1260 (1996). URL `http://dx.doi.org/10.1016/S0092-8674(00)81820-6`.

[129] Hartley, P. D. & Madhani, H. D. Mechanisms that Specify Promoter Nucleosome Location and Identity. *Cell* **137**, 445–458 (2009). URL `http://dx.doi.org/10.1016/j.cell.2009.02.043`.

[130] Michel, B. C. *et al.* A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nature Cell Biology* **20**, 1410–1420 (2018). URL `http://dx.doi.org/10.1038/s41556-018-0221-1`.

[131] Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463**, 474–484 (2010).

[132] Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Science Advances* **1**, 1–17 (2015).

[133] Watanabe, S., Radman-Livaja, M., Rando, O. J. & Peterson, C. L. A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme. *Science* **340**, 195–199 (2013).

[134] Wang, F., Ranjan, A., Wei, D. & Wu, C. Comment on "A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme". *Science* **353**, 358a (2016).

[135] Krogan, N. J. *et al.* A Snf2 Family ATPase Complex Required for Recruitment of the Histone H2A Variant Htz1. *Molecular Cell* **12**, 1565–1576 (2003).

[136] Mizuguchi, G. *et al.* ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. *Science* **303**, 343–348 (2004).

[137] Kobor, M. S. *et al.* A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biology* **2** (2004).

[138] Hsu, C. C. *et al.* Recognition of histone acetylation by the GAS41 YEATS domain promotes H2A.Z deposition in non-small cell lung cancer. *Genes and Development* **32**, 58–69 (2018).

[139] Lashgari, A., Millau, J. F., Jacques, P. É. & Gaudreau, L. Global inhibition of transcription causes an increase in histone H2A.Z incorporation within gene bodies. *Nucleic Acids Research* **45**, 12715–12722 (2017).

[140] Dijkwel, Y. & Tremethick, D. J. The Role of the Histone Variant H2A.Z in Metazoan Development. *Journal of Developmental Biology* **10** (2022).

[141] Ayala, R. *et al.* Structure and regulation of the human INO80-nucleosome complex. *Nature* **556**, 391–395 (2018).

[142] Willhoft, O. *et al.* Structure and dynamics of the yeast SWR1-nucleosome complex. *Science* **362** (2018).

[143] Yu, J. *et al.* Structural insights into histone exchange by human SRCAP complex. *Cell Discovery* **10**, 1–18 (2024).

[144] 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).

[145] Ji, L. *et al.* Structure of the NuA4 histone acetyltransferase complex. *Proceedings of the National Academy of Sciences of the United States of America* **119** (2022).

[146] Zukin, S. A. *et al.* Structure and flexibility of the yeast NuA4 histone acetyltransferase complex. *eLife* **11** (2022).

[147] Willhoft, O. & Wigley, D. B. INO80 and SWR1 complexes: the non-identical twins of chromatin remodelling. *Current Opinion in Structural Biology* **61**, 50–58 (2020).

[148] Lakomek, K., Stoehr, G., Tosi, A., Schmailzl, M. & Hopfner, K. P. Structural basis for dodecameric assembly states and conformational plasticity of the full-length AAA+ at pases rvb1·Rvb2. *Structure* **23**, 483–495 (2015). URL `http://dx.doi.org/10.1016/j.str.2014.12.015`.

[149] Zhou, C. Y. *et al.* Regulation of Rvb1/Rvb2 by a Domain within the INO80 Chromatin Remodeling Complex Implicates the Yeast Rvbs as Protein Assembly Chaperones. *Cell Reports* **19**, 2033–2044 (2017). URL `http://dx.doi.org/10.1016/j.celrep.2017.05.029`.

[150] Kunert, F. *et al.* Structural mechanism of extranucleosomal DNA readout by the INO80 complex (2022).

[151] Chen, L., Conaway, R. C. & Conaway, J. W. Multiple modes of regulation of the human Ino80 SNF2 ATPase by subunits of the INO80 chromatinremodeling complex. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 20497–20502 (2013).

[152] Cairns, B. R., Erdjument-Bromage, H., Tempst, P., Winston, F. & Kornberg, R. D. Two Actin-Related Proteins Are Shared Functional Components of the Chromatin-Remodeling Complexes RSC and SWI/SNFTwo Actin-Related Proteins Are Shared Functional Components of the Chromatin-Remodeling Complexes RSC and SWI/SNF. *Molecular Cell* **2**, 639–651 (1998).

[153] Peterson, C. L., Zhao, Y. & Chait, B. T. Subunits of the Yeast SWI/SNF Complex Are Members of the Actin-related Protein (ARP) Family. *Journal of Biological Chemistry* **273**, 23641–23644 (1998).

[154] Knoll, K. R. *et al.* The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nature Structural and Molecular Biology* **25**, 823–832 (2018).

[155] 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. *Communications Biology* **1**, 1–11 (2018). URL `http://dx.doi.org/10.1038/s42003-017-0002-6`.

[156] Shen, X., Ranallo, R., Choi, E. & Wu, C. Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Molecular Cell* **12**, 147–155 (2003).

[157] Gerhold, C. B. & Gasser, S. M. INO80 and SWR complexes: Relating structure to function in chromatin remodeling. *Trends in Cell Biology* **24**, 619–631 (2014). URL `http://dx.doi.org/10.1016/j.tcb.2014.06.004`.

[158] 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).

[159] Min, J. N. *et al.* The mINO80 chromatin remodeling complex is required for efficient telomere replication and maintenance of genome stability. *Cell Research* **23**, 1396–1413 (2013).

[160] Zhou, B. *et al.* INO80 governs superenhancer-mediated oncogenic transcription and tumor growth in melanoma. *Genes and Development* **30**, 1440–1453 (2016).

[161] Ebbert, R., Birkmann, A. & Schüller, H. J. The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. *Molecular Microbiology* **32**, 741–751 (1999).

[162] Klopf, E., Schmidt, H. A., Clauder-Münster, S., Steinmetz, L. M. & Schüller, C. INO80 represses osmostress induced gene expression by resetting promoter proximal nucleosomes. *Nucleic Acids Research* **45**, 3752–3766 (2017).

[163] Van Attikum, H., Fritsch, O., Hohn, B. & Gasser, S. M. Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell* **119**, 777–788 (2004).

[164] Udugama, M., Sabri, A. & Bartholomew, B. The INO80 ATP-Dependent Chromatin Remodeling Complex Is a Nucleosome Spacing Factor. *Molecular and Cellular Biology* **31**, 662–673 (2011).

[165] Basu, A. *et al.* Measuring DNA mechanics on the genome scale. *Nature* **589**, 462–467 (2021). URL `https://www.nature.com/articles/s41586-020-03052-3?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound`.

[166] Basu, A. *et al.* Deciphering the mechanical code of the genome and epigenome. *Nature Structural and Molecular Biology* **29**, 1178–1187 (2022).

[167] Oberbeckmann, E. *et al.* Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nature Communications* **12** (2021).

[168] Klopf, E. *et al.* Cooperation between the INO80 Complex and Histone Chaperones Determines Adaptation of Stress Gene Transcription in the Yeast Saccharomyces cerevisiae. *Molecular and Cellular Biology* **29**, 4994–5007 (2009).

[169] Oberbeckmann, E. *et al.* Ruler elements in chromatin remodelers set nucleosome array spacing and phasing. *Nature Communications* **12**, 1–17 (2021).

[170] Yao, W. *et al.* The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation. *Molecular and Cellular Biology* **36**, 979–991 (2016).

[171] Marquardt, S. *et al.* A chromatin-based mechanism for limiting divergent noncoding transcription. *Cell* **157**, 1712–1723 (2014). URL `http://dx.doi.org/10.1016/j.cell.2014.04.036`.

[172] Alcid, E. A. & Tsukiyama, T. ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. *Genes and Development* **28**, 2348–2360 (2014).

[173] Papamichos-Chronakis, M. & Peterson, C. L. The Ino80 chromatin-remodeling enzyme regulates replisome function and stability. *Nature Structural and Molecular Biology* **15**, 338–345 (2008).

[174] Shen, X., Mizuguchi, G., Hamiche, A. & Wu, C. A chromatin remodelling complex involved in transcription and. *DNA processing. Nature* **406**, 541–544 (2000).

[175] Van Attikum, H., Fritsch, O. & Gasser, S. M. Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *EMBO Journal* **26**, 4113–4125 (2007).

[176] Tsukuda, T., Fleming, A. B., Nickoloff, J. A. & Osley, M. A. Chromatin remodelling at a DNA double-strand break site in Saccharomyces cerevisiae. *Nature* **438**, 379–383 (2005).

[177] Hauer, M. H. *et al.* Histone degradation in response to DNA damage enhances chromatin dynamics and recombination rates. *Nature Structural and Molecular Biology* **24**, 99–107 (2017).

[178] Adam, S. *et al.* Real-Time Tracking of Parental Histones Reveals Their Contribution to Chromatin Integrity Following DNA Damage. *Molecular Cell* **64**, 65–78 (2016).

[179] Lafon, A. *et al.* INO80 Chromatin Remodeler Facilitates Release of RNA Polymerase II from Chromatin for Ubiquitin-Mediated Proteasomal Degradation. *Molecular Cell* **60**, 784–796 (2015). URL `http://dx.doi.org/10.1016/j.molcel.2015.10.028`.

[180] Poli, J. *et al.* Mec1, INO80, and the PAF1 complex cooperate to limit transcription replication conflicts through RNAPII removal during replication stress. *Genes and Development* **30**, 337–354 (2016).

[181] Hamperl, S. & Cimprich, K. A. Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell* **167**, 1455–1467 (2016).

[182] Seeber, A., Dion, V. & Gasser, S. M. Checkpoint kinases and the INO80 nucleosome remodeling complex enhance global chromatin mobility in response to DNA damage. *Genes and Development* **27**, 1999–2008 (2013).

[183] Strecker, J. *et al.* DNA damage signalling targets the kinetochore to promote chromatin mobility. *Nature Cell Biology* **18**, 281–290 (2016).

[184] Seeber, A. & Gasser, S. M. Chromatin organization and dynamics in double-strand break repair. *Current Opinion in Genetics and Development* **43**, 9–16 (2017). URL `http://dx.doi.org/10.1016/j.gde.2016.10.005`.

[185] Tramantano, M. *et al.* Constitutive turnover of histone H2A.Z at yeast promoters requires the preinitiation complex. *eLife* **5**, 1–30 (2016).

[186] Weber, C. M., Ramachandran, S. & Henikoff, S. Nucleosomes are context-specific, H2A.Z-Modulated barriers to RNA polymerase. *Molecular Cell* **53**, 819–830 (2014). URL `http://dx.doi.org/10.1016/j.molcel.2014.02.014`.

[187] Tosi, A. *et al.* Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* **154**, 1207 (2013).

[188] 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). URL http://dx.doi.org/10.1074/jbc.M111.222505.

[189] Watanabe, S. *et al.* Structural analyses of the chromatin remodelling enzymes INO80-C and SWR-C. *Nature Communications* **6**, 1–10 (2015).

[190] Yao, W. *et al.* Assembly of the Arp5 (actin-related protein) subunit involved in distinct ino80 chromatin remodeling activities. *Journal of Biological Chemistry* **290**, 25700–25709 (2015). URL http://dx.doi.org/10.1074/jbc.M115.674887.

[191] Willhoft, O., Bythell-Douglas, R., McCormack, E. A. & Wigley, D. B. Synergy and antagonism in regulation of recombinant human INO80 chromatin remodeling complex. *Nucleic Acids Research* **44**, 8179–8188 (2016).

[192] Brahma, S., Ngubo, M., Paul, S., Udugama, M. & Bartholomew, B. The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling. *Nature Communications* **9**, 1–10 (2018).

[193] Cai, Y. *et al.* YY1 functions with INO80 to activate transcription. *Nature Structural and Molecular Biology* **14**, 872–874 (2007).

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

[195] Shanle, E. K. *et al.* Association of Taf14 with acetylated histone H3 directs gene transcription and the DNA damage response. *Genes and Development* **29**, 1795–1800 (2015).

[196] Kim, J. M. *et al.* Single-molecule imaging of chromatin remodelers reveals role of atpase in promoting fast kinetics of target search and dissociation from chromatin. *eLife* **10** (2021). URL https://elifesciences.org/articles/69387?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound.

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

[198] Long, H. *et al.* H2A.Z facilitates licensing and activation of early replication origins. *Nature* **577**, 576–581 (2020).

[199] Xu, Y. *et al.* Histone H2A.Z Controls a Critical Chromatin Remodeling Step Required for DNA Double-Strand Break Repair. *Molecular Cell* **48**, 723–733 (2012).

[200] Gursoy-Yuzugullu, O., House, N. & Price, B. D. Patching Broken DNA: Nucleosome Dynamics and the Repair of DNA Breaks. *Journal of Molecular Biology* **428**, 1846–1860 (2016).

[201] Berta, D. G. *et al.* Deficient H2A.Z deposition is associated with genesis of uterine leiomyoma. *Nature* **596**, 398–403 (2021).

[202] Ranjan, A. *et al.* Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* **154**, 1232 (2013).

[203] Yen, K., Vinayachandran, V. & Pugh, B. F. XSWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246 (2013). URL http://dx.doi.org/10.1016/j.cell.2013.08.043.

[204] Jin, C. *et al.* H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nature Genetics* **41**, 941–945 (2009).

[205] Ranjan, A. *et al.* Live-cell single particle imaging reveals the role of RNA polymerase II in histone H2A.Z eviction. *eLife* **9** (2020). URL https://elifesciences.org/articles/55667?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound.

[206] Mylonas, C., Lee, C., Auld, A. L., Cisse, I. I. & Boyer, L. A. A dual role for H2A.Z.1 in modulating the dynamics of RNA polymerase II initiation and elongation. *nature structural & molecular biology* **28**, 435–442 (2021). URL http://dx.doi.org/10.1038/s41594-021-00589-3.

[207] Santisteban, M. S., Hang, M. & Smith, M. M. Histone Variant H2A.Z and RNA Polymerae II Transcription Elongation (2011).

[208] Neves, L. T. *et al.* The histone variant H2A.Z promotes efficient cotranscriptional splicing in s. cerevisiae. *Genes and Development* **31**, 702–717 (2017).

[209] Klein, B. J. *et al.* Yaf9 subunit of the NuA4 and SWR1 complexes targets histone H3K27ac through its YEATS domain. *Nucleic Acids Research* **46**, 421–430 (2018).

[210] Wang, A. Y. *et al.* Asf1-like structure of the conserved Yaf9 YEATS domain and role in H2A.Z deposition and acetylation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21573–21578 (2009).

[211] Ladurner, A. G., Inouye, C., Jain, R. & Tjian, R. Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Molecular Cell* **11**, 365–376 (2003).

[212] Hsu, C. C. *et al.* Gas41 links histone acetylation to H2A.Z deposition and maintenance of embryonic stem cell identity. *Cell Discovery* **4** (2018). URL http://dx.doi.org/10.1038/s41421-018-0027-0.

[213] Carcamo, C. C. *et al.* ATP binding facilitates target search of SWR1 chromatin remodeler by promoting one-dimensional diffusion on DNA. *eLife* **11** (2022). URL https://elifesciences.org/articles/77352?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound.

[214] Sun, L., Pierrakeas, L., Li, T. & Luk, E. Thermosensitive Nucleosome Editing Reveals the Role of DNA Sequence in Targeted Histone Variant Deposition. *Cell Reports* **30**, 257–268.e5 (2020).

[215] Converso, C. *et al.* Nucleic acid sequence contributes to remodeler-mediated targeting of histone H2A.Z. *bioRxiv* (2023).

[216] Krietenstein, N., Wippo, C. J., Lieleg, C. & Korber, P. *Genome-wide in vitro reconstitution of yeast chromatin with in vivo-like nucleosome positioning*, vol. 513 (Elsevier Inc., 2012), 1 edn.

[217] Van, C., Williams, J. S., Kunkel, T. A. & Peterson, C. L. Deposition of histone H2A.Z by the SWR-C remodeling enzyme prevents genome instability. *DNA Repair* **25**, 9–14 (2015). URL http://dx.doi.org/10.1016/j.dnarep.2014.10.010.

[218] Van Attikum, H., Fritsch, O. & Gasser, S. M. Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *EMBO Journal* **26**, 4113–4125 (2007).

[219] Rangasamy, D., Greaves, I. & Tremethick, D. J. RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nature Structural and Molecular Biology* **11**, 650–655 (2004).

[220] Ruhl, D. D. *et al.* Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry* **45**, 5671–5677 (2006).

[221] Lin, C. L. *et al.* Functional characterization and architecture of recombinant yeast SWR1 histone exchange complex. *Nucleic Acids Research* **45**, 7249–7260 (2017).

[222] Nguyen, V. Q. *et al.* Molecular architecture of the ATP-dependent chromatin-remodeling complex SWR1. *Cell* **154**, 1220–1231 (2013). URL http://dx.doi.org/10.1016/j.cell.2013.08.018.

[223] Ranjan, A. *et al.* H2A histone-fold and DNA elements in nucleosome activate SWR1-mediated H2A.Z replacement in budding yeast. *eLife* **4**, 1–11 (2015).

[224] Wu, W. H. *et al.* Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nature Structural and Molecular Biology* **12**, 1064–1071 (2005).

[225] Cao, T. *et al.* Crystal structure of a nuclear actin ternary complex. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 8985–8990 (2016).

[226] Huang, Y. *et al.* Role of a DEF/Y motif in histone H2A-H2B recognition and nucleosome editing. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 3543–3550 (2020).

[227] Baier, A. S., Gioacchini, N., Eek, P., Tan, S. & Peterson, C. L. Dual engagement of the nucleosomal acidic patches is essential for deposition of histone H2A.Z by SWR1C. *bioRxiv* (2023). URL https://www.researchsquare.com/article/rs-3050911/v1?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound.

[228] Wu, W. H. *et al.* N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex. *Journal of Biological Chemistry* **284**, 6200–6207 (2009).

[229] Poyton, M. F. *et al.* Coordinated DNA and histone dynamics drive accurate histone H2A.Z exchange. *Science Advances* **8** (2022).

[230] Fan, J., Moreno, A. T., Baier, A. S., Loparo, J. J. & Peterson, C. L. H2A.Z deposition by SWR1C involves multiple ATP-dependent steps. *Nature Communications* **13** (2022). URL https://www.researchsquare.com/article/rs-1280538/v1?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound.

[231] Luk, E. *et al.* Chz1, a Nuclear Chaperone for Histone H2AZ. *Molecular Cell* **25**, 357–368 (2007).

[232] Wang, Y. *et al.* Structural insights into histone chaperone Chz1-mediated H2A.Z recognition and histone replacement. *PLoS Biology* **17**, 1–20 (2019).

[233] Liang, X. *et al.* Structural basis of H2A.Z recognition by SRCAP chromatin-remodeling subunit YL1. *Nature Structural and Molecular Biology* **23**, 317–323 (2016).

[234] Hong, J. *et al.* The Catalytic Subunit of the SWR1 Remodeler Is a Histone Chaperone for the H2A.Z-H2B Dimer. *Molecular Cell* **53**, 498–505 (2014).

[235] Luk, E. *et al.* Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell* **143**, 725–736 (2010).

[236] Aguilar-Gurrieri, C. *et al.* Structural evidence for Nap1-dependent H2A–H2B deposition and nucleosome assembly. *The EMBO Journal* **35**, 1465–1482 (2016).

[237] Altaf, M. *et al.* NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex. *Journal of Biological Chemistry* **285**, 15966–15977 (2010).

[238] Bird, A. W. *et al.* Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* **419**, 411–415 (2002).

[239] Keogh, M. C. *et al.* The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. *Genes and Development* **20**, 660–665 (2006).

[240] Downey, M. *et al.* Acetylome profiling reveals overlap in the regulation ofdiverse processes by sirtuins, gcn5, and esa1. *Molecular and Cellular Proteomics* **14**, 162–176 (2015).

[241] Lin, Y.-y. *et al.* Protein Acetylation Microarray Reveals NuA4 Controls Key Metabolic Target Regulating Gluconeogenesis. *Cell* **136**, 1073–1084 (2009).

[242] Fazzio, T. G., Huff, J. T. & Panning, B. An RNAi Screen of Chromatin Proteins Identifies Tip60-p400 as a Regulator of Embryonic Stem Cell Identity. *Cell* **134**, 162–174 (2008).

[243] Babiarz, J. E., Halley, J. E. & Rine, J. Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae. *Genes and Development* **20**, 700–710 (2006).

[244] Durant, M. & Pugh, F. B. NuA4-Directed Chromatin Transactions throughout the Saccharomyces cerevisiae Genome. *Molecular and Cellular Biology* **27**, 5327–5335 (2007).

[245] Bruzzone, M. J., Grünberg, S., Kubik, S., Zentner, G. E. & Shore, D. Distinct patterns of histone acetyltransferase and mediator deployment at yeast protein-coding genes. *Genes and Development* **32**, 1252–1265 (2018).

[246] Doyon, Y. & Côté, J. The highly conserved and multifunctional NuA4 HAT complex. *Current Opinion in Genetics and Development* **14**, 147–154 (2004).

[247] Doyon, Y., Selleck, W., Lane, W. S., Tan, S. & Côté, J. Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans. *Molecular and Cellular Biology* **24**, 1884–1896 (2004).

[248] Rossetto, D. *et al.* Eaf5/7/3 form a functionally independent NuA4 submodule linked to RNA polymerase II -coupled nucleosome recycling . *The EMBO Journal* **33**, 1397–1415 (2014).

[249] Wang, X., Ahmad, S., Zhang, Z., Côté, J. & Cai, G. Architecture of the Saccharomyces cerevisiae NuA4/TIP60 complex. *Nature Communications* **9**, 1–11 (2018).

[250] Auger, A. *et al.* Eaf1 Is the Platform for NuA4 Molecular Assembly That Evolutionarily Links Chromatin Acetylation to ATP-Dependent Exchange of Histone H2A Variants. *Molecular and Cellular Biology* **28**, 2257–2270 (2008).

[251] Kühlbrandt, W. The resolution revolution (2014).

[252] Kimanius, D. *et al.* Data-driven regularisation lowers the size barrier of cryo-EM structure determination. *bioRxiv* 2023.10.23.563586 (2023). URL `https://www.biorxiv.org/content/10.1101/2023.10.23.563586v1{%}0A`.

[253] Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020). URL `http://dx.doi.org/10.1038/s41586-020-2833-4`.

[254] Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).

[255] Young, L. N. & Villa, E. Bringing Structure to Cell Biology with Cryo-Electron Tomography. *Annual Review of Biophysics* **52**, 573–595 (2023).

[256] Nogales, E. & Mahamid, J. Bridging structural and cell biology with cryo-electron microscopy. *Nature* **628**, 47–56 (2024).

[257] Dubochet, J. & McDowall, A. W. Vitrification of Pure Water for Electron Microscopy. *Journal of Microscopy* **124**, 3–4 (1981).

[258] Glaeser, R. M. & Han, B.-G. Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophysics Reports* **3**, 1–7 (2017).

[259] Xu, Y. & Dang, S. Recent Technical Advances in Sample Preparation for Single-Particle Cryo-EM. *Frontiers in Molecular Biosciences* **9**, 1–12 (2022).

[260] Pantelic, R. S., Suk, J. W., Hao, Y., Ruoff, R. S. & Stahlberg, H. Oxidative doping renders graphene hydrophilic, facilitating its use as a support in biological tem. *Nano Letters* **11**, 4319–4323 (2011).

[261] Liu, Y., Meng, X. & Liu, Z. Deformed grids for single-particle cryo-electron microscopy of specimens exhibiting a preferred orientation. *J Struct Biol* **182**, 255–258 (2013). `NIHMS150003`.

[262] D'Imprima, E. *et al.* Protein denaturation at the air-water interface and how to prevent it. *eLife* **8**, 1–18 (2019).

[263] Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal* **478**, 4169–4185 (2021).

[264] 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).

[265] Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single image processing. *eLife* (2018).

[266] Saibil, H. R. Cryo-EM in molecular and cellular biology. *Molecular Cell* **82**, 274–284 (2022). URL `https://doi.org/10.1016/j.molcel.2021.12.016`.

[267] Vilas, J. L., Carazo, J. M. & Sorzano, C. O. S. Emerging Themes in CryoEM - Single Particle Analysis Image Processing. *Chemical Reviews* **122**, 13915–13951 (2022).

[268] Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods* **16**, 1153–1160 (2020).

[269] Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *bioRxiv* 1–18 (2018).

[270] Punjani, A. & Fleet, D. J. 3DFlex: determining structure and motion of flexible proteins from cryo-EM. *Nature Methods* **20**, 860–870 (2023).

[271] Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nature Methods* **18**, 176–185 (2021).

[272] Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

[273] Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv* (2022). URL `https://doi.org/10.1101/2021.10.04.463034`.

[274] Jamali, K., Kimanius, D. & Scheres, S. H. W. A Graph Neural Network Approach to Automated Model Building in Cryo-EM Maps. *conference paper at ICLR* 1–27 (2022).

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

[276] 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).

[277] Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213–221 (2010).

[278] Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallographica Section D: Structural Biology* **77**, 1282–1291 (2021).

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

[280] Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of Molecular Biology* **276**, 19–42 (1998).

[281] Zhang, M. *et al.* Hexasome-INO80 complex reveals structural basis of noncanonical nucleosoe remodeling. *Science* (2023).

[282] Hanson, P. I. & Whiteheart, S. W. AAA+ proteins: Have engine, will work. *Nature Reviews Molecular Cell Biology* **6**, 519–529 (2005).

[283] Jansen, A. & Verstrepen, K. J. Nucleosome Positioning in Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews* **75**, 301–320 (2011).

[284] Troshin, P. V., Procter, J. B. & Barton, G. J. Java bioinformatics analysis web services for multiple sequence alignment-JABAWS:MSA. *Bioinformatics* **27**, 2001–2002 (2011).

[285] Troshin, P. V. *et al.* JABAWS 2.2 distributed web services for Bioinformatics: Protein disorder, conservation and RNA secondary structure. *Bioinformatics* **34**, 1939–1940 (2018).

[286] Amlacher, S. *et al.* Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell* **146**, 277–289 (2011). URL `http://dx.doi.org/10.1016/j.cell.2011.06.039`.

[287] Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 1–6 (2011).

[288] Holm, L. DALI and the persistence of protein shape. *Protein Science* **29**, 128–140 (2020).

[289] van Kempen, M. *et al.* Fast and accurate protein structure search with Foldseek. *Nature Biotechnology* (2023).

[290] Dubos, C. *et al.* MYB transcription factors in Arabidopsis. *Trends in Plant Science* **15**, 573–581 (2010).

[291] Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nature Methods* **3**, 1021–1032 (2006).

[292] Young, G. *et al.* Quantitative mass imaging of single biological macromolecules. *Science* **360**, 423–427 (2018).

[293] Graham, M., Combe, C., Kolbowski, L. & Rappsilber, J. xiView: A common platform for the downstream analysis of Crosslinking Mass Spectrometry data. *bioRxiv* 561829 (2019).

[294] Faini, M., Stengel, F. & Aebersold, R. The Evolving Contribution of Mass Spectrometry to Integrative Structural Biology. *Journal of the American Society for Mass Spectrometry* **27**, 966–974 (2016). URL `http://dx.doi.org/10.1007/s13361-016-1382-4`.

[295] Seidel, M. *et al.* Co-translational binding of importins to nascent proteins. *Nature Communications* **14**, 1–15 (2023).

[296] Michaelis, A. C. *et al.* The social and structural architecture of the yeast protein interactome. *Nature* **624**, 192–200 (2023).

[297] Seidel, M. *et al.* Co-translational assembly orchestrates competing biogenesis pathways. *Nature Communications* **13**, 1–15 (2022).

[298] Dominguez, R. & Holmes, K. C. Actin structure and function. *Annual Review of Biophysics* **40**, 169–186 (2011).

[299] Drulyte, I. *et al.* Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallographica Section D: Structural Biology* **74**, 560–571 (2018).

[300] Winger, J., Nodelman, I. M., Levendosky, R. F. & Bowman, G. D. A twist defect mechanism for ATP-dependent translocation of nucleosomal DNA. *eLife* **7**, 1–23 (2018).

[301] Brahma, S. & Henikoff, S. RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. *Molecular Cell* **73**, 238–249.e3 (2019).

[302] Nodelman, I. M. *et al.* Interdomain Communication of the Chd1 Chromatin Remodeler across the DNA Gyres of the Nucleosome. *Molecular Cell* **65**, 447–459.e6 (2017).

[303] Sundaramoorthy, R. *et al.* Structure of the chromatin remodelling enzyme Chd1 bound to a ubiquitinylated nucleosome. *eLife* **7**, 1–28 (2018).

[304] Chittori, S., Hong, J., Bai, Y. & Subramaniam, S. Structure of the primed state of the ATPase domain of chromatin remodeling factor ISWI bound to the nucleosome. *Nucleic Acids Research* **47**, 9400–9409 (2019).

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

[306] Wagner, F. R. *et al.* Structure of SWI/SNF chromatin remodeller RSC bound to a nucleosome. *Nature* **579**, 448–451 (2020).

[307] Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440–445 (2017).

[308] Sun, L. & Luk, E. Dual function of Swc5 in SWR remodeling ATPase activation and histone H2A eviction. *Nucleic Acids Research* **45**, 9931–9946 (2017).

[309] Huang, Y. *et al.* Role of a DEF/Y motif in histone H2A-H2B recognition and nucleosome editing. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 3543–3550 (2020).

[310] Baker, R. W. *et al.* Structural insights into assembly and function of the RSC chromatin remodeling complex. *Nature Structural and Molecular Biology* **28**, 71–80 (2021).

[311] Peterson, C. L., Zhao, Y. & Chait, B. T. Subunits of the yeast SWI/SNF complex are members of the actin-related protein (ARP) family. *Journal of Biological Chemistry* **273**, 23641–23644 (1998). URL `http://dx.doi.org/10.1074/jbc.273.37.23641`.

[312] Zhang, J., Gundu, A. & Strahl, B. D. Recognition of acetylated histone by Yaf9 regulates metabolic cycling of transcription initiation and chromatin regulatory factors. *Genes & Development* **35**, 1678–1692 (2021).

[313] Klebl, D. P., Aspinall, L. & Muench, S. P. Time resolved applications for Cryo-EM; approaches, challenges and future directions. *Current Opinion in Structural Biology* **83**, 102696 (2023). URL `https://doi.org/10.1016/j.sbi.2023.102696`.

[314] Amann, S. J., Keihsler, D., Bodrug, T., Brown, N. G. & Haselbach, D. Frozen in time: analyzing molecular dynamics with time-resolved cryo-EM. *Structure* **31**, 4–19 (2023). URL `https://doi.org/10.1016/j.str.2022.11.014`.

[315] Mäeots, M. E. *et al.* Modular microfluidics enables kinetic insight from time-resolved cryo-EM. *Nature Communications* **11**, 1–14 (2020). URL `http://dx.doi.org/10.1038/s41467-020-17230-4`.

[316] Kontziampasis, D. *et al.* A cryo-EM grid preparation device for time-resolved structural studies. *IUCrJ* **6**, 1024–1031 (2019).

[317] Frank, J. Time-resolved cryo-electron microscopy: Recent progress. *Journal of Structural Biology* **200**, 303–306 (2017). URL `https://doi.org/10.1016/j.jsb.2017.06.005`.

[318] Dandey, V. P. *et al.* Time-resolved cryo-EM using Spotiton. *Nature Methods* **17**, 897–900 (2020).

[319] Carbone, C. E. *et al.* Time-resolved cryo-EM visualizes ribosomal translocation with EF-G and GTP. *Nature Communications* **12**, 1–13 (2021).

[320] Deindl, S. *et al.* ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. *Cell* **152**, 442–452 (2013).

[321] Yudkovsky, N., Logie, C., Hahn, S. & Peterson, C. L. Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes and Development* **13**, 2369–2374 (1999).

[322] Jian, Y. *et al.* Interplay of two transcription factors for recruitment of the chromatin remodeling complex modulates fungal nitrosative stress response. *Nature Communications* **12** (2021). URL `http://dx.doi.org/10.1038/s41467-021-22831-8`.

[323] Lin, C. L. *et al.* Functional characterization and architecture of recombinant yeast SWR1 histone exchange complex. *Nucleic Acids Research* **45**, 7249–7260 (2017).

[324] Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology* **4**, 1–8 (2021). URL `http://dx.doi.org/10.1038/s42003-021-02399-1`.

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

[326] Frenz, B., Walls, A. C., Egelman, E. H., Veesler, D. & Di Maio, F. RosettaES: A sampling strategy enabling automated interpretation of difficult cryo-EMmaps. *Nature Methods* **14**, 797–800 (2017).

[327] Kowiel, M., Brzezinski, D., Gilski, M. & Jaskolski, M. Conformation-dependent restraints for polynucleotides: The sugar moiety. *Nucleic Acids Research* **48**, 962–973 (2020).

[328] Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *eLife* **6**, 1–26 (2017).

[329] Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* **152**, 36–51 (2005).

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

[331] Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682 (2012).